Economics Working Paper Series
Faculty of Business, Economics and Law, AUT

Category-Dependent Preferences and Stochastic Choice

Matthew Ryan

2024/06*

* supersedes 2019/03

Category-Dependent Preferences and Stochastic Choice

Matthew Ryan*
School of Economics and Finance
Auckland University of Technology

October 4, 2024

Abstract

We introduce a generalisation of Aguiar's (2017) *random categorisation rule (RCR)* that allows preferences to be category dependent. Our *generalised random categorisation rule (GRCR)* requires the preferences associated with two different categories to agree on their intersection. We show that this is equivalent to relaxing the completeness and transitivity requirement on the preference relation in the RCR model. The GRCR model is therefore characterised by relaxing Aguiar's (2017) *Acyclicity* axiom to an *Asymmetry* axiom. We also provide a characterisation of the model in terms of Block-Marschak polynomials. Finally, we show that a random choice function has a Manzini and Mariotti (2014) model iff it has both a GRCR representation and a Brady and Rehbeck (2016) model.

*I would like to gratefully acknowledge the School of Economics and Finance at Queen Mary University of London for its generous hospitality during the writing of this paper. Earlier drafts circulated under the title "Generalised Random Categorisation Rules". Special thanks to Marco Mariotti for illuminating discussions of the ideas in this paper. I have also benefitted from audience comments at the NZ Microeconomics Study Group (Auckland) and the BRIC conference (Aarhus). Email: `mryan@aut.ac.nz`

1 Introduction

Consumption items are elements of categories before they are elements of budget sets. Outside laboratory settings, consumers rarely "experience" budget sets.[1] Rather, consumers experience wants, needs and desires. It is therefore natural that they should organise the world into different categories that meet such needs, wants and desires. Categorisation simplifies choice by focussing attention on a salient subset of the budget set – the *consideration set*.

We imagine the following scenario. A consumer experiences a desire for a particular category of good. She then determines which, if any, goods in that category are within her budget set. If none is available, she does not make a purchase – which we formally represent as choosing the "default" option. Otherwise, she chooses her most preferred member of the category within her budget set. Aguiar's (2017) *random categorisation rule (RCR)* models just such a scenario. Our model will differ in one important respect. We allow that categorisation may affect how alternatives are evaluated: preferences may be category-dependent, but with an important restriction.

The motivation behind our generalisation of Aguiar's RCR is the idea that preferences are partially formed rather than entirely given. The mechanism of preference formation is *purposive comparison* – comparing alternatives for the purpose of making a choice. In particular, this is what enforces "rational" consistency on preferences. Only alternatives within the same category are ever directly compared for the purposes of making a choice. Moreover, alternatives that do not co-exist in any single category may be intrinsically more difficult to compare – apples versus oranges, rather than different varieties of apples. It is therefore natural that preference discipline may be weaker across categories than within.

Our *generalised random categorisation rule (GRCR)* relaxes the restriction that the same preference order (on the universal domain of alternatives) guides choice within every category. However, in line with the motivation above, we restrict this category-dependence by requiring agreement on category intersections. If two goods co-exist in two different categories, they must be ranked the same way by the preference orders associated with each of the categories. Conflicting category-dependent rankings of two goods cannot survive their careful comparison for the purpose of making a choice.

As a consumer gains more experience, *indirect* comparisons may gradually eliminate preference conflict *across* categories. Her behaviour will approxi-

[1] By "budget set" we simply mean the set of alternatives available for choice; the alternatives need not be bundles of goods and the available set need not be determined by a fixed budget and fixed prices for each good.

mate more and more closely to an RCR. However, a fully dynamic model of preference formation is a task for future research.[2]

Our main results offer two characterisations of the GRCR. The first is obtained by relaxing the Acyclicity axiom in Aguiar's (2017) axiomatisation of the RCR to an Asymmetry condition. The second characterisation uses the fact that the RCFs with a GRCR representation form a subclass of the RCFs with a *random utility model (RUM).* We identify the restrictions on the Block-Marschak polynomials that determine this subclass.

2 Random categorisation

Let X be a finite set of alternatives. We use 2^X to denote the power set of X. A *choice set* will be an element of 2^X. Let $a^* \notin X$ denote the default (or outside) option. Note that a^* is excluded from X by assumption: if the choice set is empty, then a^* must be chosen. For each $A \subseteq X$ let $A^* = A \cup \{a^*\}$. A consumer facing choice set A may choose any element of A^*; the latter is called the consumer's *budget set.* Let $\overline{A} = X \diagdown A$ for each $A \subseteq X$. *Thus, $\overline{A}$ is the complement of A in X, not in X^*.* Throughout, we omit brackets around singleton sets whenever convenient, provided no confusion is likely to arise.

A *random choice function (RCF)* is a mapping $p : X^* \times 2^X \to [0,1]$ that satisfies $p(x, A) = 0$ if $x \notin A^*$ and

$$\sum_{x \in A^*} p(x, A) = 1$$

for all $A \subseteq X$. We interpret $p(x, A)$ as the probability of choosing x given choice set A. If $E \subseteq X^*$ we write $p(E, A)$ as shorthand for

$$\sum_{x \in E} p(x, A).$$

We interpret p as the stochastic choice behaviour of some individual and we seek conditions for its consistency with some convenient model of behaviour.

To describe the models of interest we need some additional notation. Let Σ denote the set of reflexive and antisymmetric binary relations on X, and let $\Lambda \subseteq \Sigma$ denote the elements of Σ which are also complete and transitive (i.e., linear orders $\equiv$ antisymmetric weak orders). If $\succsim \in \Sigma$ then $\succ$ and $\sim$ are defined from $\succsim$ in the usual way. Given $x \in X$ and $\succsim \in \Sigma$, let $\succsim(x)$ denote the $\succsim$-upper contour of x and $\succ(x)$ the $\succ$-upper contour. That is:

$$\succsim(x) \equiv \{y \in X \mid y \succsim x\}$$

[2]For one attempt at a dynamic model in similar spirit, see Barokas (2021).

and

$$\succ(x) \equiv \{y \in X \mid y \succ x\}.$$

Note that $\succsim(x) = \succ(x) \cup \{x\}$ when $\succsim \in \Lambda$. Moreover, if $\succsim \in \Lambda$ then x is $\succsim$-maximal in $E \subseteq X$ iff $E \cap \succsim(x) = \{x\}$. We employ analogous notation for binary relations on X^*. In particular: Λ^* is the set of linear orders on X^* and

$$\succsim^*(x) \equiv \{y \in X^* \mid y \succsim x\}$$

for any $x \in X$ and any $\succsim^* \in \Lambda^*$. The set of all probability mass functions on the finite set Λ^* is denoted by Δ. If $\pi \in \Delta$ and $E \subseteq \Lambda^*$ then we write $\pi(E)$ for

$$\sum_{\succsim^* \in E} \pi(\succsim^*).$$

Finally, an *attention index* (Aguiar *et al.*, 2023) is a mapping $m : 2^X \to [0,1]$ that satisfies

$$\sum_{E \subseteq X} m(E) = 1.$$

Let $\mathcal{A}$ denote the set of all attention indices. If $m \in \mathcal{A}$ we use supp(m) to denote the support of m. We interpret $m(E)$ as the probability that the decision-maker pays attention (only) to the category of goods defined by the set E.

The following model goes back (at least) to Block and Marschak (1960):[3]

Definition 1 *If p is a random choice function, then p has a* ***random utility model (RUM)*** *if there exists some $\pi \in \Delta$ such that the following holds for any $(x, A) \in X^* \times 2^X$ with $x \in A^*$:*

$$p(x, A) = \pi(\{\succsim^* \in \Lambda^* \mid A^* \cap \succsim^*(x) = \{x\}\})$$

In this case, we say that π is a RUM for p.

Next, we have Aguiar's (2017) random categorisation rule:

Definition 2 (Aguiar, 2017) *If p is a random choice function, then p can be represented by a* ***random categorisation rule (RCR)*** *iff there exist $m \in \mathcal{A}$ and $\succsim \in \Lambda$ such that*

$$p(x, A) = \sum_{E : (E \cap A) \cap \succsim(x) = \{x\}} m(E) \quad (1)$$

[3]Definition 1 adapts the classical RUM to a choice environment with default. In the classical version, there is no requirement that budget sets contain a^*; the domain of p is $X^* \times \left(2^{X^*} \setminus \{\emptyset\}\right)$, with $p(x, A)$ interpreted as the probability of choosing x when facing budget set (not choice set) A.

for all $A \subseteq X$ *and every* $x \in X$*, and*

$$p(a^*, A) = \sum_{E: E \cap A = \emptyset} m(E) \qquad (2)$$

for any $A \subseteq X$*. In this case, we say that* $(m, \succsim)$ *is an RCR for* p*.*

The RCR captures the spirit of the scenario in the Introduction. Desires to consume a particular category of goods arrive randomly according to the attention index, m. If no member of the category is available in the current choice set, the consumer chooses the default – see (2) – which we interpret as choosing not to do any shopping. Otherwise, she chooses the $\succsim$-most-preferred alternative in the intersection of the category and the choice set: see (1). We refer to this intersection as the consumer's *consideration set*.

Aguiar (2017) shows that every RCF with an RCR representation also has a RUM. This is not surprising. In the presence of a default option, random preference can replicate random categorisation by randomly demoting alternatives below a^* in the ranking. Rather than choosing category E with probability $m(E)$, we choose (with the same probability) an order $\succsim^*_E \in \Lambda^*$ that ranks everything in E above a^*, everything in $\overline{E}$ below a^* and matches the RCR order, $\succsim$, on E.

A feature of the RCR, which is in common with many models of choice based on consideration sets,[4] is that the default is implicitly inferior to anything in X, since it is chosen only if nothing else is considered. While the assumption that all budget sets contain a common default is a *structural* assumption which can (in principle) be checked, the additional assumption that the default is undesirable is a restriction on *preferences*, which are endogenous to the model. Moreover, this restriction does not have the flavour of a "rationality" requirement, like transitivity for example. How can it be justified?

Consider the following variation on the RCR, which relaxes this preference requirement: there exist $m \in \mathcal{A}$ and $\succsim^* \in \Lambda^*$ such that

$$p(a, A) = \sum_{E: (E \cap A)^* \cap \succsim^*(a) = \{a\}} m(E)$$

for all $A \subseteq X$ and every $a \in X$, and

$$p(a^*, A) = 1 - \sum_{a \in A} p(a, A) = \sum_{E: (E \cap A)^* \cap \succsim^*(a^*) = \{a^*\}} m(E).$$

[4]This trend was established in the pioneering work of Manzini and Mariotti (2014).

If $a^* \succ^* a$ for some $a \in X$, then this model implies that we must have $p(a, A) = 0$ for all $A \subseteq X$. Options that are ranked below the default are irrelevant to choice, and this will be empirically evident. Therefore, provided we add one further *structural* assumption – that $p(a, \{a\}) > 0$ for all $a \in X$ – we must have a^* ranked below every $a \in X$ in this model. Under this additional assumption, we effectively have an RCR representation. If this structural assumption is *not* met, then it does no harm to remove the offending elements of X until it does, since the removed options would never be chosen anyway. This logic suggests that the preference restrictions in an RCR are indeed innocuous (or rather, can be replaced by defensible structural asusmptions).

The following generalises the RCR notion in a way that preserves this logic:

Definition 3 *If p is a random choice function, then p can be represented by a* ***generalised random categorisation rule (GRCR)*** *iff there exist $m \in \mathcal{A}$ and $\succsim_E \in \Lambda$ for each $E \in 2^X$, such that: (i) for each $E, F \in \text{supp}(m)$ the binary relations $\succsim_E$ and $\succsim_F$ agree on $E \cap F$; and (ii) for all $A \subseteq X$ and every $a \in X$,*

$$p(a, A) = \sum_{E:(E \cap A) \cap \succsim_E (a) = \{a\}} m(E) \tag{3}$$

and

$$p(a^*, A) = \sum_{E:E \cap A = \emptyset} m(E) \tag{4}$$

Condition (i) says that any two category-dependent preferences in the support of m must agree on the intersection of their respective categories. If (i) is strengthened to a requirement that $\succsim_E = \succsim_F = \succsim$ for each $E, F \in 2^X$ then $(m, \succsim)$ is an RCR for p.[5] Hence, any RCR is a GRCR. However, the converse is false, as the following example illustrates.

Example 1 *Let $X = \{a, b, c\}$, $E = \{a, b\}$, $F = \{b, c\}$ and $G = \{a, c\}$. Define attention index $m \in \mathcal{A}$ as follows:*

$$m(E) = m(F) = m(G) = \frac{1}{3}.$$

Let $\succsim_E \in \Lambda$ satisfy $b \succ_E a$, let $\succsim_F \in \Lambda$ satisfy $c \succ_F b$, and let $\succsim_G \in \Lambda$ satisfy $a \succ_G c$. Defining p using (3)-(4) gives an RCF with a GRCR representation. Suppose $(\hat{m}, \succsim)$ is an RCR that rationalises p. Since

$$p(a, \{a, b\}) = \frac{1}{3} < \frac{2}{3} = p(a, \{a\}) \tag{5}$$

[5]In fact, it suffices that $\succsim_E = \succsim_F$ for each $E, F \in \text{supp}(m)$.

we deduce $b \succ a$. (Note that $p(a, \{a\}) = \sum_{A: a \in A} \hat{m}(A)$, so (5) implies that a is not always chosen when it is considered alongside b.) Likewise, we deduce $c \succ b$ from $p(b, \{b, c\}) < p(b, \{b\})$ and $a \succ c$ from $p(c, \{a, c\}) < p(c, \{c\})$. This contradicts the transitivity of $\succ$.

How can we justify the GRCR assumption that *all* the category-dependent preferences rank the default last? Once again, we start by defining a variant model that relaxes just this constraint. Consider a model in which there exist $m \in \mathcal{A}$ and $\succsim_E^* \in \Lambda^*$ for each $E \in 2^X$, such that (i*) for each $E, F \in 2^X$ with $m(E)\, m(F) > 0$, the binary relations $\succsim_E^*$ and $\succsim_F^*$ agree on $(E \cap F)^*$; and (ii*) for all $A \subseteq X$ and every $a \in X$,

$$p(a, A) \;=\; \sum_{E: (E \cap A)^* \cap \succsim_E^*(a) = \{a\}} m(E)$$

and

$$p(a^*, A) \;=\; 1 - \sum_{a \in A} p(a, A).$$

For any $a \in X$, condition (i*) implies that a^* is ranked the same way against a by any $\succsim_E^*$ with $a \in E$ and $m(E) > 0$. Hence, if $a^* \succ_E^* a$ for some $E \subseteq X$ and some $a \in E$ with $m(E) > 0$ then $a^* \succ_F^* a$ for any $F \subseteq X$ with $a \in F$ and $m(F) > 0$. Therefore, we can justify our preference restriction in the GRCR by exactly the same logic as its justification in the RCR context. Condition (i) in the GRCR definition plays an important role in this justification.

As for an RCR (see Corollary 1 of Aguiar, 2017), the attention index in a GRCR is unique:[6]

Proposition 1 *Let p be an RCF with a GRCR representation. Then the attention index in the representation is unique (i.e., any two GRCR representations for p must share the same attention index). In particular*

$$m(A) \;=\; \sum_{B: B \subseteq A} (-1)^{|A \setminus B|}\, p\left(a^*, \overline{B}\right) \tag{6}$$

for each $A \subseteq X$.

Aguiar (2017, Lemma 1) notes that any RCF with an RCR representation also has a RUM. This result also generalises to the GRCR.

Proposition 2 *Any RCF with a GRCR representation also has a RUM. The converse is false.*

[6] Proofs of all results can be found in the Appendix.

3 Characterisations of the GRCR class

In this section we characterise the class of RCFs that can be represented by a GRCR. In fact, we provide two characterisations: see Theorem 3 and Corollary 1 below.

Our first main result will facilitate the first of these characterisations by allowing us to restate the GRCR definition in a useful equivalent form:

Theorem 1 *If p is a random choice function, then p has a GRCR representation iff there exists $(m, \succsim) \in \mathcal{A} \times \Sigma$ such that*

$$p(a, A) = \sum_{E : E \cap A \cap \succsim(a) = \{a\}} m(E) \tag{7}$$

and

$$p(a^*, A) = \sum_{E : E \cap A = \emptyset} m(E) \tag{8}$$

If p is a random choice function and the pair $(m, \succsim) \in \mathcal{A} \times \Sigma$ satisfies (7) and (8), then we will also refer to $(m, \succsim)$ as a GRCR representation for p. In this formulation, a GRCR has a single preference relation but it need not be linear: reflexivity and antisymmetry are the only (explicit) requirements. However, conditions (7)-(8), *together with the assumption that p is an RCF*, impose additional restrictions on $\succsim$. This is the essence of Theorem 1. Inother words, condition (i) in Definition 3 ensures that category dependence of the linear preference can be replicated by relaxing the linearity requirement on the (single) preference relation.

Note that if $(m, \succsim)$ is a GRCR representation for p in the sense of Theorem 1, then (8) ensures that m continues to be uniquely determined from p by the Möbius inversion formula (6) – see the proof of Proposition 1.

It is easy to construct a GRCR representation, in the sense of Theorem 1, for Example 1. Simply define m as before and let

$$\succsim = \{(a, a), (b, b), (c, c), (b, a), (c, a), (a, c)\}.$$

In a GRCR, the preferences on any *consideration set* that occurs with positive probability are independent of how that consideration set was determined: if $E, F \in \text{supp}(m)$ and $A, B \in 2^X$ are choice sets such that $A \cap E = B \cap F$, then $\succsim_E$ and $\succsim_F$ coincide on this common intersection. Theorem 1 therefore allows us to express the GRCR in the language of "random attention" models (see Cattaneo *et al.,* 2020 and Kovach and Suleymanov, 2023). Let us say that $\mu : 2^X \times 2^X \to [0,1]$ is a *consideration* mapping provided $\mu(B,A) > 0$ only if $B \subseteq A$, and

$$\sum_{B \subseteq A} \mu(B,A) = 1$$

for each $A \subseteq X$. We interpret $\mu(B,A)$ as the probability that the consideration set is B when choosing from A. A consideration mapping, μ, has a *constant random attention (CRA)* representation if there exists an attention index, m, such that

$$\mu(B,A) \;=\; \sum_{E : E \cap A = B} m(E)$$

for every $A, B \in 2^X$ with $B \subseteq A$. It is evident that if μ has a CRA representation then μ is *monotonic* in the sense of Cattaneo *et al.* (2020, Assumption 1). We therefore have the following result, whose straightforward proof is omitted:

Proposition 3 *Let p be an RCF. Then p has a GRCR iff there exists a consideration mapping, μ, with a CRA representation and $\succsim \in \Sigma$ such that: for all $A \subseteq X$ and every $a \in X$,*

$$p(a,A) \;=\; \sum_{E : E \cap \precsim(a) = \{a\}} \mu(E,A) \tag{9}$$

and

$$p(a^*,A) \;=\; \mu(\emptyset, A) \tag{10}$$

The GRCR is thus a generalised form of *random attention model* (Cattaneo *et al.*, 2020, Definition 3) in which preferences are not required to be linear. In fact, the GRCR does not quite "generalise" the random attention model, since the GRCR imposes the additional restriction that the outside option be ranked last in X^*. There is no outside option in the set-up of Cattaneo *et al.* (2020); the set-up of Kovach and Suleymanov (2023) includes an outside option, but its ranking is unrestricted.[7]

[7] We also deviate from the terminology in Cattaneo *et al.* (2020) and Kovach and Suleymanov (2023). They refer to μ as an "attention rule" but we wish to distinguish the category to which *attention* is focussed from the *consideration* set, which also incorporates any budget or other feasibility constraints on choice.

3.1 From Acyclicity to Asymmetry

Using Theorem 1, the GRCR can be axiomatically characterised by adapting Aguiar's (2017) axiomatisation of the RCR. The latter involves two axioms.

The first is the *Weakly Decreasing Marginal Propensity (of Choice)*, or *WDMP*, axiom. This axiom requires *C-total monotonicity* of the function $\varphi : 2^X \to [0,1]$ defined by $\varphi(A) = 1 - p(a^*, A)$. We refer the reader to Aguiar's paper (*ibid.*, Definition 10) for the formal definition of C-total monotonicity. For our purposes, the key feature of the WDMP axiom is that it is equivalent to requiring $m \geq 0$ when $m : 2^X \to \mathbb{R}$ is defined by (6) – see the discussion on p.48 of Aguiar (2017) and Section 7.2.3 of Grabisch (2016).[8] Since m is the Möbius inverse of the mapping $E \to p\left(a^*, \overline{E}\right)$, and since $p(a^*, \emptyset) = 1$, the WDMP axiom also ensures that $m \in \mathcal{A}$.

Axiom 1 (WDMP) *The function $m : 2^X \to \mathbb{R}$ defined by (6) satisfies $m \geq 0$.*

To state his second axiom, Aguiar defines the following *revealed strict preference* relation on X: $a \rhd b$ iff $p(b, A \cup a) \neq p(b, A)$ for some A containing b (where $a, b \in X$).

Axiom 2 (Acyclicity) *The relation $\rhd$ is acyclic.*

Acyclicity of $\rhd$ ensures that it has an extension in Λ by Szpilrajn's Theorem. Aguiar proves that *any* such linear extension, together with the attention function described by (6), provides an RCR representation for p. Conversely, every p with an RCR representation satisfies the WDMP and Acyclicity Axioms.

Theorem 2 (Aguiar, 2017) *Let p be a random choice function. Then p has an RCR representation iff it satisfies WDMP and Acyclicity.*

To characterise the GRCR we weaken Acyclicity to:

Axiom 3 (Asymmetry) *The relation $\rhd$ is asymmetric.*

[8] Aguiar's terminology is somewhat non-standard. What he calls the C-total monotonicity property is better known in the literature as the ∞-*alternating* (or $\cup$-alternating) property. The implied property of the *dual function* $\overline{\varphi} : 2^X \to [0,1]$, defined by $\overline{\varphi}(A) = 1 - \varphi\left(\overline{A}\right)$, is conventionally called "total monotonicity" (equivalently, "∞-monotonicity" or "$\cap$-monotonicity").

In choice-theoretic terms, the Acyclicity axiom has the flavour of a Strong Axiom of Revealed Preference, while our Asymmetry Axiom is analogous to a Weak Axiom of Revealed Preference. Defining $\trianglerighteq$ to be the minimal reflexive extension of $\triangleright$ (that is: $a \trianglerighteq b$ iff $a = b$ or $a \triangleright b$)[9] we have:

Theorem 3 *Let p be an RCF. Then p has a GRCR representation iff it satisfies WDMP and Asymmetry. Moreover, if p has a GRCR representation, then $(m, \trianglerighteq)$ is a GRCR representation for p, where $m : 2^X \to \mathbb{R}$ is defined by (6). (We refer to this as the "canonical" GRCR representation for p.)*

Thus, when p possesses a GRCR representation, both components of the canonical GRCR may be explicitly recovered from p. All components of the canonical representation are revealed by choice data. Example 1 shows that $\trianglerighteq$ may be complete yet not transitive.

The proof of Theorem 3 also clarifies the relationship between the binary relation $\trianglerighteq$ in the canonical representation and the family of category-dependent preferences in a GRCR representation in the sense of Definition 3: if E is in the support of m then the restriction of $\trianglerighteq$ to E coincides with $\succsim_E|_E$ (i.e., the restriction of $\succsim_E$ to E). In fact:

$$\trianglerighteq = \bigcup_{E \in \text{supp}(m)} \succsim_E|_E \,.$$

3.2 A Block-Marschak characterisation

We first recall the characterisation of the classical RUM. For any given RCF, p, define $h_x : 2^X \to [0,1]$ for each $x \in X$ by

$$h_x(E) \;=\; p\left(E, \overline{E} \cup \{x\}\right)$$

and let $m_x : 2^X \to \mathbb{R}$ be the Möbius inverse of h_x (Shafer, 1976, Lemma 2.3). Therefore:

$$h_x(E) \;=\; \sum_{A : A \subseteq E} m_x(A)$$

and

$$m_x(A) \;=\; \sum_{B : B \subseteq A} (-1)^{|A \setminus B|}\, h_x(B)\,.$$

Theorem 4 (McFadden, 2005) *Let p be an RCF. Then p has a RUM iff $m_x \geq 0$ for all $x \in X$ and $m \geq 0$ when $m : 2^X \to \mathbb{R}$ is defined by (6).*

[9] Note that $\trianglerighteq$ need not coincide with $\{(a,b) \in X \times X \mid (b,a) \notin \triangleright\}$.

Theorem 4 is a special case of Theorem 3.3 in McFadden (2005). It adapts the characterisation of a classical RUM (Falmagne, 1978; Barberá and Pattanaik, 1986) to choice problems with a default. This characterisation requires non-negativity of a family of Möbius functions, known as *Block-Marschak polynomials*.[10] In particular, the function m is the Möbius inverse of the mapping $E \to p\left(a^*, \overline{E}\right)$.

It is possible to re-express the Asymmetry axiom in terms of the Block-Marschak polynomials $\{m_x\}_{x \in X}$ defined previously.

Proposition 4 *Suppose that p is an RCF. Then p satisfies Asymmetry iff*

$$\left[\sum_{B:\{a,b\}\subseteq B\subseteq X} |m_a(B)|\right]\left[\sum_{B:\{a,b\}\subseteq B\subseteq X} |m_b(B)|\right] = 0$$

for any $a, b \in X$ with $a \neq b$.

Recall (Proposition 2) that every RCF with a GRCR representation has a RUM. Recall too that if p has a RUM then $m_x \geq 0$ for all $x \in X$ (Theorem 4). The *additional* restriction imposed by the Asymmetry Axiom is therefore:

$$\left[\sum_{B:\{a,b\}\subseteq B\subseteq X} m_a(B)\right]\left[\sum_{B:\{a,b\}\subseteq B\subseteq X} m_b(B)\right] = 0$$

for any distinct $a, b \in X$. It is clear that this restriction is substantive: it determines a "non-generic" subclass.

Corollary 1 *Let p be an RCF and let $m : 2^X \to \mathbb{R}$ be defined by (6). Then p has a GRCR representation iff $m \geq 0$, $m_x \geq 0$ for all $x \in X$, and*

$$\left[\sum_{B:\{a,b\}\subseteq B\subseteq X} m_a(B)\right]\left[\sum_{B:\{a,b\}\subseteq B\subseteq X} m_b(B)\right] = 0$$

for any $a, b \in X$ with $a \neq b$.

[10] In fact, our Möbius functions differ from those appearing in the conventional *Block-Marschak polynomials*. When p has a RUM and $x \in A$, then $m_x(A)$ is the probability of selecting a linear order for which A is the *upper* contour set for x, and $m(A)$ is the probability of selecting an order for which A^* is the *upper* contour set for a^*. The standard Block-Marschak polynomials assign these probabilities instead to the corresponding *lower* contour sets and are obtained by the obvious change of variables. See Block and Marschak (1960).

The following corollary describes this subclass of random utility models.

Corollary 2 *Let p be an RCF and let $\pi \in \Delta$ be a RUM for p. Then p satisfies Asymmetry iff*

$$\pi\left(\{\succsim^* | a \succ^* b \succ^* a^*\}\right)\pi\left(\{\succsim^* | b \succ^* a \succ^* a^*\}\right) \;=\; 0 \tag{11}$$

for any $a, b \in X$.

Condition (11) says that all linear orders in the support of π which rank both a and b above a^* must order a and b the same way. In other words, the Asymmetry Axiom excludes disagreement about the ranking of alternatives which are ranked above the default.[11]

4 Comparison with other random attention models

Aguiar's (2017) RCR is one of several well-known random attention models with an outside option. All these models feature a linear order ($\succsim$) on X and a consideration mapping $\mu : 2^X \times 2^X \to [0, 1]$. In each of these models, the decision-maker, when confronted with a menu, A, first randomly draws a consideration set, B, according to $\mu(\cdot.A)$, then chooses the $\succsim$-maximal element of B, unless $B = \emptyset$ in which case a^* is chosen. The models differ in their specification of μ. If $(m, \succsim)$ is an RCR for p then, as noted above, the associated consideration mapping has the CRA representation

$$\mu(B, A) \;=\; \sum_{E:E\cap A=B} m(E) \tag{12}$$

Two other well-known random attention models are those of Brady and Rehbeck (2016; [BR]) and Manzini and Mariotti (2014; [MM]). The BR model also generates the consideration mapping from an attention index, $m \in \mathcal{A}$, but BR require m to have *full support* (i.e., $m(E) > 0$ for all $E \subseteq X$) and specify μ as follows:

$$\mu(B, A) \;=\; \frac{m(B)}{\sum_{E\subseteq A} m(A)} \tag{13}$$

In this case we say that μ has a *Luce random attention (LRA)* representation. The MM model features a function $\gamma : X \to (0, 1)$ which specifies the

[11] This does not exclude the possibility that the "lower ranked" of the two (say, a) may be chosen with positive probability from a budget set that includes both a and b: the support of π may include an order in which a is ranked above a^* and b below a^*.

probability that any given element of X will receive attention, with these attention probabilities being independent across alternatives. Thus:

$$\mu(A, B) = \prod_{x \in A} \gamma(x) \prod_{y \in B \setminus A} (1 - \gamma(y)) \tag{14}$$

In this case we say that μ has an *independent random attention (IRA)* representation.

The following result is accepted wisdom in the literature:[12]

Proposition 5 *Let p be an RCF. Then p has an MM model iff it has both an RCR representation and a BR model.*

This (correct) claim is usually asserted on the strength of Kovach and Suleymanov (2023).[13] The latter authors prove *(ibid.*, Corollary 1) a version of Proposition 5 for a richer environment in which any element of X^* may serve as default, and random choice behaviour is specified for all possible default options. Their set-up also allows that the default may be chosen even when non-default options are considered. That Proposition 5 is implied by Corollary 1 in Kovach and Suleymanov (2023) is, perhaps, not entirely obvious.[14] We therefore provide a different proof of Proposition 5 in Appendix H. This may be of independent interest.

To understand the relationship between the BR model and the GRCR we make use of the following:

Proposition 7 *Let p be an RCF. If $(m, \trianglerighteq)$ is a canonical GRCR representation for p and $(\hat{m}, \succsim)$ is a BR model for p, then $\trianglerighteq = \succsim$. In particular, $(m, \trianglerighteq)$ is an RCR representation for p.*

[12] See, for example, Figure 2 in Aguiar *et al.* (2023).

[13] Or its unpublished predecessor, Suleymanov (2018).

[14] Corollary 1 in Kovach and Suleymanov (2023) is based on their Proposition 1, which says the following:

Proposition 6 (Suleymanov, 2018; Kovach and Suleymanov, 2023) *Let μ be a consideration mapping. Then μ has an IRA representation iff it has both an LRA representation and a CRA representation.*

Proposition 6 clearly implies the "if" part of Proposition 5. However, to use Proposition 6 to establish the "only if" part we must assume that the RCR representation and BR model for p share the same linear order (as noted by Kovach and Suleymanov: *ibid.*, p.424) and the same consideration mapping. The former assumption is unproblematic (see Proposition 7 below) but we are not aware of an argument that justifies the latter.

Corollary 3 *Let p be an RCF. Then p has an MM model iff it has both a GRCR representation and a BR model.*

Corollary 3 is immediate given Propositions 5 and 7, plus the fact that any RCF with an RCR representation also has a GRCR representation (Theorem 1). Figure 1 summarises the relationships amongst the models.

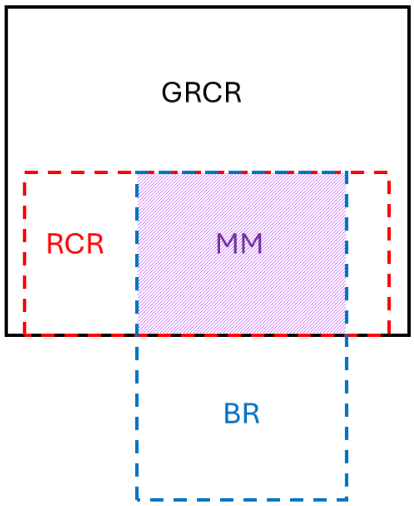

Figure 1: Relationships amongst the models

5 Concluding remarks

In an RCR, choice is guided by a rational (linear) preference relation applied to a random consideration set – the intersection of the choice set with a randomly selected category. We have shown that allowing preference to be *category-dependent*, but with agreement on category intersections, is equivalent to relaxing the *completeness and transitivity* requirements of a category-independent preference relation. It follows that our *generalised* RCR may be characterised by weakening Aguiar's (2017) Acyclicity restriction on his revealed preference relation to an Asymmetry restriction.

The ill-behavedness of category-independent preferences in a GRCR does not undermine the possibility of rational choice, since preferences need only be well-behaved on sets of options which are *considered together*. How much discipline this imposes is endogenous, depending on the individual's consideration process.

If one assumes that preferences are disciplined by the experience of making choices, then it is natural to suppose that preferences will exhibit more consistency on sets of alternatives which are considered together. This is precisely what the GRCR model requires. Of course, the ill-behavedness of the category-independent preference may, in principle (this is outside the model), result in choice cycles over a *sequence* of choice problems (recall Example 1), thereby prompting some further process of preference revision.

We also observed that every GRCR is observationally equivalent to some random utility model, and we showed how to translate the Asymmetry condition into a restriction on Block-Marschak polynomials. This alternative characterisation of the GRCR model raises the prospect of testing the GRCR restriction of the RUM polytope using recently developed econometric methodologies.[15]

As illustrated in Figure 1, the GRCR class of RCFs strictly expands the RCR class, but without encompassing any additional RCFs characterised by the model in Brady and Rehbeck (2016). An auxiliary contribution of the present paper is to provide a new proof of the fact that Manzini and Mariotti's (2014) model characterises the intersection of the RCR and BR classes. The MM model also, it turns out, characterises the intersection of the GRCR and BR classes.

References

[1] Aguiar, V. H. (2017). Random categorization and bounded rationality. *Economics Letters*, **159**, 46-52.

[2] Aguiar, V. H., Boccardi, M. J., Kashaev, N., and Kim, J. (2023). Random utility and limited consideration. *Quantitative Economics*, **14**, 71-116.

[3] Barberá, S., and Pattanaik, P. K. (1986). Falmagne and the rationalizability of stochastic choices in terms of random orderings. *Econometrica*, **54(3)**, 707-715.

[4] Barokas, G. (2021). Dynamic choice under familiarity-based attention. *Social Choice and Welfare*, **57**, 703-720.

[5] Block, H. D., and Marschak, J. (1960). Random orderings and stochastic theories of response, in: *Contributions to Probability and Statistics, I.*

[15] See, for example, Kitamura and Stoye (2018) and McCausland *et al.* (2020). Turansick (2023) has recently proposed a more efficient testing strategy.

Olkin, S. Ghurye, W. Hoeffding, W. Madow, H. Mann (Eds.), Stanford University Press, pp. 97-132.

[6] Brady, R. L., and Rehbeck, J. (2016). Menu-dependent stochastic feasibility. *Econometrica*, **84(3)**, 1203-1223.

[7] Cattaneo, M. D., Ma, X., Masatlioglu, Y., and Suleymanov, E. (2020). A random attention model. *Journal of Political Economy*, **128(7)**, 2796-2836.

[8] Falmagne, J. C. (1978). A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, **18(1)**, 52-72.

[9] Grabisch, M. (2016). *Set Functions, Games and Capacities in Decision Making*. Springer International Publishing.

[10] Kovach, M., and Suleymanov, E. (2023). Reference dependence and random attention. *Journal of Economic Behavior and Organization*, **215**, 421-441.

[11] Manzini, P., and Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, **82(3)**, 1153-1176.

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

Appendices

A Proof of Proposition 1

As for an RCR, the probability of choosing the default option in a GRCR representation depends only on the attention index – the preferences play no role. In particular,

$$p\left(a^{*},A\right) \;=\; \sum_{B:B\cap A=\emptyset} m\left(B\right)$$

for any $A \subseteq X$. Defining $f\left(A\right) \equiv p\left(a^{*},\overline{A}\right)$ we have

$$f\left(A\right) \;=\; \sum_{B\subseteq A} m\left(B\right)$$

for any $A \subseteq X$. It follows by the theory of Möbius inversion (Shafer, 1976, Lemma 2.3) that m is unique and determined by

$$m(A) = \sum_{B:B\subseteq A} (-1)^{|A\setminus B|} f(B)$$

which is (6).

B Proof of Proposition 2

We first show that every GRCR model has an observationally equivalent RUM. Consider a GRCR model with attention index, m, and category-dependent preferences, $\{\succsim_E\}_{E\subseteq X}$. Note that the only part of $\succsim_E$ that matters is the ranking of the elements in E. Now define $\{\succsim_E^*\}_{E\subseteq X}$ by choosing $\succsim_E^* \in \Lambda^*$ so that it agrees with $\succsim_E$ on E, and ranks any element of E (respectively, $\overline{E}$) above (respectively, below) a^*. Note that $\succsim_E^* = \succsim_F^*$ iff $E = F$. For each $\succsim^* \in \Lambda^*$ let

$$\pi(\succsim^*) = \sum_{E:\succsim_E^* = \succsim^*} m(E)$$

(with the convention that a sum over the empty set is zero) so that $\pi \in \Delta$ with support $\{\succsim_E^*\}_{E\subseteq X}$. For any $E, A \in 2^X$ and any $x \in X$:

$$\begin{aligned}
(E\cap A)\cap \succsim_E(x) = \{x\} \quad &\Leftrightarrow \quad E\cap(A\cap \succsim_E(x)) = \{x\} \\
&\Leftrightarrow \quad A^*\cap(E^*\cap \succsim_E^*(x)) = \{x\} \\
&\Leftrightarrow \quad A^*\cap \succsim_E^*(x) = \{x\}
\end{aligned}$$

where the final equivalence uses the following two facts: $\succsim_E^*(x) \subseteq E^*$ whenever $x \in E^*$; and $a^* \succ_E^* y$ for any $y \in \overline{E}$. Hence, for every $A \subseteq X$ and $x \in X$ we have

$$\begin{aligned}
p(x, A) &= \sum_{E:(E\cap A)\cap\succsim_E(x)=\{x\}} m(E) \\
&= \sum_{E:A^*\cap\succsim_E^*(x)=\{x\}} m(E) \\
&= \pi(\{\succsim^* \in \Lambda^* \mid A^*\cap \succsim^*(x) = \{x\}\})
\end{aligned}$$

Since

$$p(a^*, A) = 1 - \sum_{x \in A} p(x, A) = \pi(\{\succsim^* \in \Lambda^* \mid A^* \cap \succsim^* (a^*) = \{a^*\}\})$$

it follows that π gives a RUM for p.

The following example shows that not every RCF with a RUM has a GRCR model.

Example 2 *Let $X = \{a, b\}$ and consider the RUM that chooses each of the following linear orders on X^* (summarised by their asymmetric parts) with probability $\frac{1}{2}$:*

$$a \succ_1 b \succ_1 a^*$$

and

$$b \succ_2 a \succ_2 a^*.$$

Let p denote the RCF generated by this RUM. Suppose p has a RGCR representation. Since $p(x, \{x\}) = 1$ for each $x \in X$ it follows that each alternative in X must be present in every category that is given attention with strictly positive probability. Thus, $m(X) = 1$ in the GRCR, which implies that choice is deterministic. But this contradicts the fact that $p(a, X) = p(b, X) = \frac{1}{2}$.

C Proof of Theorem 1

We first prove:

Lemma C.1 *Let p be a random choice function. Then p has a GRCR representation iff there exists $(m, \succsim) \in \mathcal{A} \times \Sigma$, with $\succsim$ linear on each $E \subseteq X$ for which $m(E) > 0$,[16] such that (7) and (8) hold for all $A \subseteq X$ and every $a \in X$.*

Proof. The "if" part is obvious: for each $E \subseteq X$ for which $m(E) > 0$, define $\succsim_E$ to be any linear extension (to X) of the restriction of $\succsim$ to E, and define $\succsim_E \in \Lambda$ arbitrarily if $m(E) = 0$. Conversely, suppose p has a GRCR representation with attention index m and category-dependent linear preferences $\{\succsim_E\}_{E \subseteq X}$. Let

$$F = \left\{ x \in X \;\middle|\; \sum_{E : x \in E} m(E) = 0 \right\}.$$

[16] That is, the restriction of $\succsim$ to E is linear.

Thus, F contains the alternatives excluded from all categories in the support of m. Define $\succsim \in \Sigma$ as follows: $\succsim \cap F^2 = \{(x,x) \mid x \in F\}$ and for all $E \subseteq X$ with $m(E) > 0$, $\succsim \cap E^2$ is the restriction of $\succsim_E$ to E. This construction is well-defined, since $\succsim_E$ and $\succsim_F$ agree on $E \cap F$ whenever $m(E)\, m(F) > 0$; it is also obvious that $\succsim \in \Sigma$ and that $\succsim$ is linear on E whenever $m(E) > 0$. Since $\succsim$ coincides with $\succsim_E$ on E if $m(E) > 0$, condition (7) is satisfied.

This completes the proof of Lemma C.1. ■

The "only if" part of Theorem 1 follows directly from Lemma C.1. To complete the proof of Theorem 1 it suffices to show that if p is an RCF and $(m, \succsim) \in \mathcal{A} \times \Sigma$ satisfies (7) and (8), then $\succsim$ is linear on each $E \subseteq X$ for which $m(E) > 0$.

Lemma C.2 *Let p be an RCF and suppose $(m, \succsim) \in \mathcal{A} \times \Sigma$ satisfies (7) and (8). Then m assigns zero probability to any $B \subseteq X$ such that (i) there exist $a, b \in B$ with $a \neq b$, $a \notin \succsim(b)$ and $b \notin \succsim(a)$, or (ii) B contains a cycle with respect to $\succ$.*

Proof. Suppose B satisfies (i) and $m(B) > 0$. Let $A = \{a, b\}$. It follows that $\{a,b\} \cap \succsim(a) = \{a\}$ and hence

$$\begin{aligned} p(a, A) &= \sum_{C : C \cap \{a,b\} \cap \succsim(a) = \{a\}} m(C) \\ &= \sum_{C : a \in C} m(C) \\ &= \sum_{E : E \cap \{a,b\} = \{a\}} m(E) + \sum_{F : F \cap \{a,b\} = \{a,b\}} m(F). \end{aligned}$$

Similarly,

$$p(b, A) = \sum_{E : E \cap \{a,b\} = \{b\}} m(E) + \sum_{F : F \cap \{a,b\} = \{a,b\}} m(F).$$

Moreoever,

$$p(a^*, A) = \sum_{C : C \cap \{a,b\} = \emptyset} m(C)$$

so

$$p(A^*, A) = 1 + \sum_{F : F \cap A = \{a,b\}} m(F) \geq 1 + m(B) > 1.$$

This is the desired contradiction.

Next, suppose B satisfies (ii) and $m(B) > 0$. By what we have already established, we may assume that $m(E) = 0$ for any E with $a, b \in E$ such that $a \neq b$, $a \notin \succ (b)$ and $b \notin \succ (a)$. Let $A = \{a_0, a_1, ..., a_n\} \subseteq B$ with $a_0 = a_n$ and $a_i \succ a_{i+1}$ for each $i \in \{0, 1, ..., n-1\}$. Since $\succ$ is asymmetric, $n \geq 3$. It follows that $m(B)$ does not contribute to $p(x, A)$ for any $x \in A^*$. To avoid the conclusion that $p(A^*, A) < 1$ there must be some E with $m(E) > 0$ that contributes to the probability of choosing *more than one* element of A^*. This requires that there exist $a_j, a_k \in A$ with $a_j \neq a_k$,

$$E \cap A \cap \succsim (a_j) = \{a_j\}$$

and

$$E \cap A \cap \succsim (a_k) = \{a_k\}.$$

But this implies $\{a_j, a_k\} \subseteq E \cap A$, and hence $a_j \notin \succsim (a_k)$ and $a_k \notin \succsim (a_j)$. Once again, we have a contradiction. ■

Lemma C.2 means that $\succ$ is connected and acyclic, hence transitive, on any E with $m(E) > 0$. Since $\succsim \in \Sigma$, it follows that $\succsim$ is linear (in particular: complete, antisymmetric and quasi-transitive) on any such E. This completes the proof of Theorem 1.

D Proof of Theorem 3

With Theorem 1 in hand, our proof of Theorem 3 will closely follow the arguments in Aguiar (2017), as readers who are familar with that paper will easily recognise. However, we provide the details for completeness.

It will be useful to refer to the following upper contour sets for the strict and weak revealed preference relations:

$$\rhd (a) \equiv \{b \in X | b \rhd a\}.$$

$$\unrhd (a) \equiv \{b \in X | b \unrhd a\} = \{a\} \cup \rhd (a)$$

for any $a \in X$.

The following is an important observation about $\unrhd$:[17]

Lemma D.1 *Let p be an RCF. Then*

$$p(a, A) = p(a, \unrhd (a) \cap A)$$

for every $A \subseteq X$ and every $a \in A$.

[17] Aguiar (2017, p.51) makes a similar observation.

Proof. If $b \in A$ and $b \notin \trianglerighteq (a)$ then $b \neq a$ and $p(a, B \cup b) = p(a, B)$ for all B containing a. We may therefore successively remove each such b from A without affecting the probability that a is chosen. ■

Next, we have the key implication of Axiom 3:

Lemma D.2 (cf., Aguiar, 2017, Lemma 2.) *Let p be an RCF. If p satisfies Asymmetry, then*

$$p(a, A) \;=\; p(a^*, A \cap \triangleright (a)) \;-\; p(a^*, A \cap \trianglerighteq (a))$$

for every $A \subseteq X$ and every $a \in A$.

Proof. Suppose $a \in A \subseteq X$. For any $b \in \triangleright (a) \cap A$ the asymmetry of $\triangleright$ implies $p(b, \trianglerighteq (a) \cap A) = p(b, \triangleright (a) \cap A)$. Therefore:

$$\begin{aligned}
p(a, \trianglerighteq (a) \cap A) &= p(a, \trianglerighteq (a) \cap A) \;+\; \sum_{b \in \triangleright (a) \cap A} \left[p(b, \trianglerighteq (a) \cap A) - p(b, \triangleright (a) \cap A) \right] \\
&= \left[\sum_{b \in \trianglerighteq (a) \cap A} p(b, \trianglerighteq (a) \cap A) \right] \;-\; \left[\sum_{b \in \triangleright (a) \cap A} p(b, \triangleright (a) \cap A) \right] \\
&= \left[1 - p(a^*, \trianglerighteq (a) \cap A) \right] \;-\; \left[1 - p(a^*, \triangleright (a) \cap A) \right] \\
&= p(a^*, \triangleright (a) \cap A) \;-\; p(a^*, \trianglerighteq (a) \cap A)
\end{aligned}$$

The result now follows by Lemma D.1. ■

In a GRCR representation, the default is chosen iff no other option is available and considered. Hence, if $(m, \succsim)$ is a GRCR for p, $A \subseteq X$ and $a \in A$, then $p(a^*, \succ (a) \cap A)$ is the probability that no consideration is given to anything in A that is preferred to a, and $p(a^*, A \cap \succsim (a))$ is the probability that neither a nor anything in A that is preferred to a is considered. The difference will therefore be the probability of choosing a from A. Lemma D.2 says that the same relationship holds for *any* RCF when we replace $\succ$ with $\triangleright$ and $\succsim$ with $\trianglerighteq$, *provided* $\triangleright$ *is asymmetric.*

Finally, the following fact will be useful for establishing the necessity of our axioms.

Lemma D.3 *Let p be an RCF and let $(m, \succsim)$ be a GRCR for p. Then $\triangleright \subseteq \succ$.*

Proof. Suppose $a, b \in X$ with $a \rhd b$, so $p(b, A) \neq p(b, A \cup a)$ for some $A \subseteq X$ with $b \in A$. Hence $a \notin A$ (and $a \neq b$ in particular) and we may deduce that $a \succ b$ as follows: if $a \notin \succ (b)$ then

$$A \cap \succsim (b) = (A \cup a) \cap \succsim (b)$$

so $p(b, A) = p(b, A \cup a)$, which is a contradiction. ■

Our main result is now within easy reach. We first verify the necessity of the axioms. Let $(m, \succsim)$ be a GRCR for p. The necessity of Axiom 1 follows from Propositions 1-2 and Theorem 4. To see the necessity of Axiom 3, suppose $a, b \in X$ with $a \rhd b$. Hence $a \neq b$ and (Lemma D.3) $a \succ b$. By the asymmetry of $\succ$ we therefore have $b \notin \succ (a)$, from which it follows that

$$C \cap \succsim (a) = (C \cup b) \cap \succsim (a)$$

for any $C \subseteq X$ with $a \in C$. Hence $p(a, C) = p(a, C \cup b)$ for any $C \subseteq X$ with $a \in C$, so $(b, a) \notin \rhd$. Hence $\rhd$ is asymmetric.

Next, we prove sufficiency. Let $m : 2^X \rightarrow \mathbb{R}$ be the function defined by (6). As per the discussion in the paragraph prior to the statement of Axiom 1, the WDMP condition, together with the fact that $p(a^*, \emptyset) = 1$, implies that $m \in \mathcal{A}$ and

$$p(a^*, A) = \sum_{B : B \subseteq \overline{A}} m(B) \tag{15}$$

for any $A \subseteq X$. Using Lemma D.2 and (15) we have, for any $A \subseteq X$ and any $a \in A$:

$$\begin{aligned} p(a, A) &= p(a^*, \rhd (a) \cap A) - p(a^*, \trianglerighteq (a) \cap A) \\ &= [1 - p(a^*, \trianglerighteq (a) \cap A)] - [1 - p(a^*, \rhd (a) \cap A)] \\ &= \left[\sum_{B : B \cap A \cap \trianglerighteq (a) \neq \emptyset} m(B) \right] - \left[\sum_{B : B \cap A \cap \rhd (a) \neq \emptyset} m(B) \right] \\ &= \sum_{B : B \cap A \cap \trianglerighteq (a) = \{a\}} m(B) \end{aligned}$$

Hence $(m, \trianglerighteq)$ is a GRCR for p. This verifies sufficiency and also establishes the final claim in the Theorem.

E Proof of Proposition 4

The Asymmetry Axiom is equivalent to the following: for any $a, b \in X$ with $a \neq b$, either

$$p(a, A \cup \{a\}) = p(a, A \cup \{a, b\}) \tag{16}$$

for all $A \subseteq X \diagdown \{a, b\}$, or

$$p(b, A \cup \{b\}) = p(b, A \cup \{a, b\}) \tag{17}$$

for all $A \subseteq X \diagdown \{a, b\}$. Conditions (16) and (17) may be written as $h_a\left(\overline{A}\right) = h_a\left(\overline{A} \diagdown \{b\}\right)$ and $h_b\left(\overline{A}\right) = h_b\left(\overline{A} \diagdown \{a\}\right)$ respectively, where function h_x is defined in Theorem 4. Recall that $m_x : 2^X \to \mathbb{R}$ is the Möbius inverse of $h_x : 2^X \to \mathbb{R}$ and $m_x(E) = 0$ if $x \notin E$. Hence

$$h_x(A) = \sum_{B : B \subseteq A} m_x(B) = \sum_{B : x \in B \subseteq A} m_x(B).$$

(We note in passing that $0 \leq p(x, X) = h_x(\{x\}) = m_x(\{x\})$, so only the sign of $m_x(E)$ with $|E| \geq 2$ is at issue.) The Asymmetry Axiom may therefore be expressed in terms of m_x as follows: for any $a, b \in X$ with $a \neq b$, either

$$\sum_{B : \{a,b\} \subseteq B \subseteq \overline{A}} m_a(B) = 0 \tag{18}$$

for all $A \subseteq X \diagdown \{a, b\}$, or

$$\sum_{B : \{a,b\} \subseteq B \subseteq \overline{A}} m_b(B) = 0 \tag{19}$$

for all $A \subseteq X \diagdown \{a, b\}$. A simple argument by induction on the cardinality of $\overline{A}$ will convince the reader that (18)-(19) hold iff: for any $a, b \in X$ with $a \neq b$, either

$$\{a, b\} \subseteq B \subseteq X \quad \Rightarrow \quad m_a(B) = 0 \tag{20}$$

or

$$\{a, b\} \subseteq B \subseteq X \quad \Rightarrow \quad m_b(B) = 0 \tag{21}$$

The result follows.

F Proof of Corollary 2

It is well known that

$$m_x(A) = \pi\left(\{\succsim^* \in \Lambda^* | \succsim^*(x) = A\}\right)$$

whenever $x \in A \subseteq X$ (Block and Marschak, 1960; McFadden, 2005). Hence, condition (20) in the proof of Proposition 4 says that

$$\pi\left(\{\succsim^* \in \Lambda^* |\succsim^*(a) = B\}\right) = 0$$

for any $B \subseteq X$ containing $\{a, b\}$. This is evidently equivalent to

$$\pi\left(\{\succsim^* \in \Lambda^* | a \succ^* b \text{ or } a^* \succ^* a\}\right) = 1.$$

Condition (21) may be equivalently expressed as

$$\pi\left(\{\succsim^* \in \Lambda^* | b \succ^* a \text{ or } a^* \succ^* b\}\right) = 1.$$

Asymmetry therefore restricts π as follows: for any $a, b \in X$ with $a \neq b$ there do not exist $\succsim_1^*$ and $\succsim_2^*$ in the support of π such that $b \succ_1^* a \succ_1^* a^*$ and $a \succ_2^* b \succ_2^* a^*$.

G Proof of Proposition 7

If $a \succ b$ then

$$p(b, \{a, b\}) = \frac{\hat{m}(\{b\})}{\sum_{E \subseteq \{a,b\}} \hat{m}(E)} < \frac{\hat{m}(\{b\})}{\sum_{E \subseteq \{b\}} \hat{m}(E)} = p(b, \{b\})$$

since m has full support. Hence $\succ \subseteq \rhd$. It follows that $\unrhd$ is complete and antisymmetric, and contains the linear order, $\succsim$. We now easily deduce that $\succsim = \unrhd$: if $a \rhd b$ we must have $a \succ b$ since otherwise we obtain $b \rhd a$ from $\succ \subseteq \rhd$, which contradicts the asymmetry of $\rhd$.

H Proof of Proposition 5

The "if" part is proved in the text. To prove the "only if" part we first recall the MIDO axiom of Brady and Rehbeck (2016), which requires:[18]

$$\frac{p(a^*, A \setminus \{b\})}{p(a^*, A)} = \frac{p(a^*, B \setminus \{b\})}{p(a^*, B)} \tag{22}$$

for any $\{A, B\} \subseteq 2^X$ and any $b \in A \cap B$.

[18] The MIDO axiom comprises part of the i-Independence condition in Manzini and Mariotti (2014) – the part that restricts the probability of choosing the default.

Lemma H.1 *Let p be an RCF which has a BR model and an RCR representation. Then p satisfies the MIDO axiom.*[19]

Proof. Let $\gamma(x) = p(x, \{x\})$ for each $x \in X$. Note that

$$\gamma(x) = 1 - p(a^*, \{x\}) \in (0, 1).$$

We will show that

$$p(a^*, E) = \prod_{x \in E} (1 - \gamma(x)) \tag{23}$$

for any $E \subseteq X$, from which (22) easily follows. We argue by induction on $|E|$. The case $|E| = 1$ is immediate. Fix some integer $k > 1$ and suppose (23) holds for any E with $|E| < k$. Let $E \subseteq X$ with $|A| = k$. It follows by Proposition 7 that the BR model and RCR representation for p have a common linear order: when $\trianglerighteq$ is complete it is the unique linear order in any RCR representation for p. Let $\succsim$ denote this common linear order and let $x \in E$ be $\succsim$-maximal in E. Then Lemma 3.1 in Brady and Rehbeck (2016) implies that

$$p(a^*, E) \;=\; p(E^* \setminus \{x\}, E)\, p(a^*, E \setminus \{x\}).$$

By the inductive hypothesis we have:

$$\begin{aligned} p(a^*, E) &= p(E^* \setminus \{x\}, E) \prod_{y \in E \setminus \{x\}} (1 - \gamma(y)) \\ &= (1 - p(x, E)) \prod_{y \in E \setminus \{x\}} (1 - \gamma(y)). \end{aligned}$$

Since p has an RCR representation and x is $\succsim$-maximal in E, it follows that $p(x, E) = p(x, \{x\})$ so we have:

$$p(a^*, E) \;=\; (1 - p(x, \{x\})) \prod_{y \in E \setminus \{x\}} (1 - \gamma(y)) \;=\; \prod_{y \in E} (1 - \gamma(y)).$$

This completes the proof of the lemma. ■

Proposition 5 now follows from Theorems 3.1 and 3.3 in Brady and Rehbeck (2016), which imply that any RCR that has a BR model and satisfies the MIDO axiom also has an MM model.

[19] Note that any RCF with a BR model satisfies $p(a^*, E) \in (0, 1)$ for any non-empty $E \subseteq X$, so (22) is always well-defined when p possesses such a model.