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

Generalised Random Categorisation Rules

Matthew Ryan

2019/03

Generalised Random Categorisation Rules

Matthew Ryan*
School of Economics
Auckland University of Technology

February 11, 2019

Abstract

Aguiar's (2017) *random categorisation rule (RCR)* describes random choice behaviour as the maximisation of a linear preference order over the intersection of a random consideration set with the set of available options. A key axiom in Aguiar's (2017) characterisation of the RCR is an *acyclicity* condition on a revealed preference relation derived from the random choice function. We show that this condition may be substantially weakened – to *asymmetry* of the revealed preference relation – without jeopardising the essence of the RCR representation. In our generalisation of the RCR, preferences may be ill-behaved on subsets of alternatives that are never considered together. While these pathologies in preference are masked by the decision-maker's selective attention to an particular choice problem, they may still be revealed by data across different choice problems. Finally, we show that the generalised model remains within the random utility class.

1 Introduction

Block and Marschak (1960) considered probabilistic choice behaviour that may be characterised as the maximisation of a random (injective) utility function. The necessary and sufficient conditions for this characterisation are rather complicated (Falmagne, 1978) and choice data reveal only limited information about the underlying distribution over preferences (Gibbard,

*My thanks to the School of Economics and Finance at Queen Mary University of London for its hospitality during the writing of this paper. Special thanks to Marco Mariotti for many stimulating conversations on random consideration set models.

2017). Manzini and Mariotti (2014) introduced the idea of a *random consideration* respesentation of probabilistic choice, in which a *single* (injective) utility function is maximised over a random subset of the available alternatives. This characterisation is based on the notion that decision-makers assign their attention randomly to a subset of available options. In the model of Manzini and Mariotti, each element is considered randomly and independently. Their model admits an elegant axiomatisation, and all parameters of the model – the attention probabilities and the preferences – can be fully recovered from choice data. Probabilistic choices that satisfy the Manzini and Mariotti axioms also have a random utility characterisation.

Aguiar (2017) provides a generalisation of Manzini and Mariotti's (2014) random consideration model, which he calls a *random categorisation rule*, that relaxes the restrictive independence assumption on consideration. Instead, potential alternatives are grouped into (possibly overlapping) *categories* and one category is randomly attended to when making a decision. Probabilistic choices that follow a random categorisation rule also have a random utility characterisation. The random consideration process can be extracted from choice data, but given the lack of restrictions on this process only limited information on preferences may be revealed.

We offer a generalisation of Aguiar's model that relaxes the assumption that preferences are linear. This is rather natural when consideration is "categorical". If a given collection of alternatives never occur together within any single category that is given positive attention, then the decision-maker is never called upon to "discipline" her preferences over these alternatives: they are never considered *simultaneously*. Preferences only need to be well-behaved within categories that may be considered with positive probability. Nevertheless, ill-behaved portions of the decision-maker's preferences may still be revealed by comparing choice behaviour *across* choice problems.

We show that this generalisation of Aguiar's random categorisation rule may be characterised by relaxing one of Aguiar's axioms. We also prove that our generalisation remains within the random utility class; that is, it is compatible with the maximisation of a randomly chosen *injective* utility function. Finally, we exhibit a random choice function that satisfies our axioms but not Aguiar's, and another that admits a random utility representation but does not satisfy our axioms. In other words, the class of random choice functions characterised by our axioms lies strictly between the random utility class and the those that follow random categorisation rules.

2 Preliminaries

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^* be a "default" option. We interpret a^* as the option of not making a choice, so a^* is excluded from X by assumption: that is, $a^* \notin X$. If the choice set is empty, then a^* is the only option that is available. For each $A \subseteq X$ let $A^* = A \cup \{a^*\}$. Hence, a decision-maker facing choice set A may choose any element of A^*. We also define $\overline{A} = X \diagdown A$ 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 $E \subseteq A^*$ we write $p(E, A)$ as shorthand for

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

Various models of random choice have been proposed, and each such model delineates a particular family of random choice functions. A classical example is the *random utility model (RUM),* most commonly associated (at least amongst economists) with the work of Block and Marschak (1960). More recently, various types of *random consideration (RC)* model have been proposed, beginning with that of Manzini and Mariotti (2014). For technical reasons, the RC models are defined for choice situations with a fixed default – as described here – while RUM's have traditionally been defined for choice situations in which "no choice" is not an option. However, the RUM concept is easily adapted to the present choice-with-default framework.

To describe these models, some further concepts and notation are useful.

- Given a binary relation $R \subseteq E \times E$ on a (finite) set E, we say that R is *asymmetric* if $(x, y) \in R$ implies $(y, x) \notin R$ for any $x, y \in E$. In particular, an asymmetric binary relation is *irreflexive*: $(x, x) \notin R$ for all $x \in E$. We say that R is *connected* if $\{(x, y), (y, x)\} \cap R \neq \emptyset$ for any *distinct* $x, y \in E$ and *transitive* if $(x, z) \in R$ whenever $\{(x, y), (y, z)\} \subseteq R$ for any $x, y, z \in E$. An asymmetric, connected and transitive binary relation is called a *strict linear order (SLO)*. Finally, R is *acyclic* provided there does *not* exist a finite subset $\{a_0, a_1, ..., a_n\}$ of E with $n > 1$ such that $a_0 = a_n$ and $a_i R a_{i+1}$ for all $i \in \{0, 1, ..., n-1\}$. Note that a transitive relation is acyclic and an acyclic relation is asymmetric. As usual, the notation xRy is synonymous with $(x, y) \in R$.

- Let Σ denote the set of strict linear orders on X and Σ^* the set of strict linear orders on X^*. The elements of Σ will be described interchangeably as subsets of $X \times X$ or as bijective mappings $\rho : X \to \{1, 2, ..., |X|\}$: the bijection $\rho \in \Sigma$ is synonymous with the binary relation $\succ^\rho \in \Sigma$ satisfying $a \succ^\rho b$ iff $\rho(a) < \rho(b)$. If $\succ \in \Sigma$ we use $\succsim$ to denote the binary relation $\succ \cup \{(x, x) \mid x \in X\}$. Thus, if $\succ \in \Sigma$ then $\succsim$ is the associated linear order. Given $x \in X$ and $\succ \in \Sigma$, let $\succ(x) \equiv \{y \in X \mid y \succ x\}$ and $\succsim(x) \equiv \{y \in X \mid y \succsim x\}$. Analogous notational conventions will be applied to Σ^*.

- Finally, given a finite set E, we use $\Delta(E)$ to denote the set of probabilities on (the subsets of) E.

3 Random Categorisation

Aguiar (2017) defines a class of random consideration models – those that conform to a *random categorisation rule (RCR)* – which generalises Manzini and Mariotti (2014). Aguiar's RCR also demarcates an important sub-class of random choice functions which have a RUM (*ibid.*, Lemma 1).

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

$$p(x, A) = \pi(\{\rho \in \Sigma^* \mid \rho(x) \leq \rho(a) \text{ for all } a \in A^*\})$$

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

For "classical" choice structures (i.e., those without a default option), there is an analogous definition with π an element of $\Delta(\Sigma)$ rather than $\Delta(\Sigma^*)$. Necessary and sufficient conditions for the existence of a "classical" RUM were first published by Falmagne (1978), based on the famous Block-Marschak inequalities (Block and Marschak, 1960). Necessary and sufficient conditions for the existence of a RUM in the sense of Definition 1 were recently established by Gibbard (2017). Gibbard's characterisation is also based on a set of Block-Marschak inequalities (though the adaptation of Falmagne's logic to the choice-with-default case is non-trivial).

Aguiar's (2017) notion of an RCR is based on what we will call a *consideration function*, which is a mapping $m : 2^X \to [0,1]$ that satisfies

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

The quantity $m(E)$ is the probability that the decision-maker "considers" – that is, resticts attention to – only those available options which lie within E. Let $\mathcal{C}$ denote the set of all consideration functions.

Definition 2 *If p is a random choice function, we say that p has a* random categorisation (RCG) *representation if there exists some $m \in \mathcal{C}$ and some $\succ \in \Sigma$ such that, for each $A \subseteq X$ and every $a \in A$,*

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

and

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

In this case, we say that the pair $(m, \succ)$ is an RCG model for p.

It is important to note that if $(m, \succ)$ is an RCG model then $\succ$ is an element of Σ, not Σ^*. It is implicit in the RCG concept that *the default is strictly worse than anything in X*. In particular, a^* is chosen from A^* if, and only if, no element of A is considered.[1] If something other than a^* is available and considered, then a^* will not be chosen. This fact is what allows us to recover the underlying consideration function from any p with an RCG model – see Gibbard (2017) and the discussion below for more on this important observation.

Our definition of an RCG model differs immaterially from Aguiar's (2017) notion of a random categorisation rule (RCR). An RCR describes preferences via an injective utility function while an RCG uses a strict linear order. Since only ordinal properties of utility are relevant to choice under an RCR, nothing is lost thereby. An RCR also specifies a collection $\mathcal{D}$ of subsets of X in which the support of m is contained. However, since the support of m may be *strictly* contained in $\mathcal{D}$, and since the family $\mathcal{D}$ is *endogenous* to the particular RCR (as opposed to an exogenous constraint upon the notion

[1] This implicit assumption is inherited from Manzini and Mariotti (2014).

of an RCR), its only role is notational convenience. Nothing of substance is lost by omitting any reference to it.

The RCG class is characterised by two conditions on the random choice function. The first is essentially one of the Block-Marschak inequalities.

Axiom 1 (WDMP) *The mapping* $A \longmapsto p\left(a^*, \overline{A}\right)$ *has a Möbius inverse which is non-negative everywhere. That is, there exists some* $m : 2^X \rightarrow \mathbb{R}_+$ *such that*

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

for any $A \subseteq X$*.*

The acronym WDMP stands for "Weakly Decreasing Marginal Propensity (of Choice)". Our statement of the WDMP axiom differs from Aguiar's. His version imposes a *total monotonicity* condition on the mapping $\varphi : 2^X \rightarrow [0, 1]$ defined as follows: $\varphi(A) = 1 - p(a^*, A)$. This is equivalent to our condition (Chateauneuf and Jaffray, 1989), but it is the total monotonicity version that motivates the name of the axiom. Since total monotonicity conditions – like Block-Marschak inequalities – are rather cumbersome to state, and have (arguably) little advantage in intuitive appeal, we adopt the more "direct" version of the WDMP axiom given above. This will help to make the logic of Aguiar's – and our – representation result more transparent.

Since p is an RCF, we have $p(a^*, \emptyset) = 1$ so the function m in the WDMP axiom satisfies

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

It follows that $m \in \mathcal{C}$. The WDMP axiom therefore delivers one of the two elements that make up an RCG representation. Moreover, if p has an RCG representation then choice data "reveal" the underlying consideration function via Möbius inversion of the mapping $A \longmapsto p\left(a^*, \overline{A}\right)$:

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

The fact that the consideration function in an RCG model is observable from choice data is somewhat intuitive. The default is chosen if, and only if, no other available alternative is considered, *irrespective of preferences over* X. Data on the probability of choosing a^* therefore conveys all the information necessary to recover m, albeit in a deeply encoded fashion.

Unlike the consideration function, the strict linear order in an RCG model may not be fully revealed by choice data. The preference $a \succ b$ cannot affect choice unless m assigns positive probability to some E with $\{a, b\} \subseteq E$ (and maybe not even then). Aguiar therefore defines the following *revealed preference* relation on X: $a \rhd b$ iff $p(b, A \cup a) \neq p(b, A)$ for some A containing b (where $a, b \in X$). *Note that $\rhd$ is irreflexive by construction.* If p has an RCG representation $(m, \succ)$, then $\rhd \subseteq \succ$ but the containment may be strict. To ensure that the revealed preference relation is compatible with *some* strict linear order, Aguiar imposes:

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

Acycicity of $\rhd$ ensures that it has an extension in Σ by Szpilrajn's Theorem. The WDMP axiom therefore delivers an element $m \in \mathcal{C}$ and the Acyclicity axiom a strict linear order $\succ \in \Sigma$ that extends $\rhd$. Aguiar proves that this pair constitutes an RCG model for p and, conversely, that every p with an RCG model satisfies his two axioms.

Theorem 1 (Aguiar, 2017) *Let p be a random choice function. Then p has an RCG model iff it satisfies WDMP and Acyclicity.*

Recall that an acyclic binary relation is asymmetric. In the following section we explore the implications of relaxing Axiom 2 to the requirement that $\rhd$ is asymmetric. It turns out that essentially the same representation may be obtained, but the preference relation $\succ$ need no longer be connected nor transitive. These deficiencies in preference may be evident in choice data, but they do not undermine the viability of preference-guided choice due to the decision-maker's selective attention. Cycles are never "noticed" by the decision-maker when confronting any particular choice problem.

4 Generalised Random Categorisation

Consider the following generalisation of the RCG rule:

Definition 3 *If p is a random choice function, we say that p has a* generalised random categorisation (RCG) *representation if there exists some $m \in \mathcal{C}$ and some asymmetric binary relation $\succ \subseteq X \times X$ such that, for each $A \subseteq X$ and every $a \in A$,*

$$p(a, A) = \sum_{B: B \cap A \cap (\succ(a) \cup a) = \{a\}} m(B)$$

and

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

In this case, we say that the pair $(m, \succ)$ is a GRCG model for p.

There are two differences between Definitions 2 and 3: first, the binary relation in a GRCG is required to be asymmetric but not necessarily connected or transitive; second, the specification of $p(a, A)$ in a GRCG model uses "$\succ (a) \cup a$" in place of "$\succsim (a)$" to account for the possibility that $\succ$ may not be connected.

Nevertheless, the basic spirit of an RCG representation is preserved in Definition 3. Element $a \in A$ is chosen from A^* iff it is considered but nothing in $A \cap \succ (a)$ is considered at the same time. The default option is chosen iff nothing in the choice set is considered.

4.1 Representation

We will prove that an RCF has a GRCG representation iff it satisfies the WDMP axiom and:

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

In this section we prove our respresentation result. Readers familiar with Aguiar (2017) will easily recognise that our arguments closely follow Aguiar's proof. After establishing our representation result, we then turn to the interpretation of GRCG models and their relation to RCG models and RUMs.

It will be useful to define the following minimal reflexive extension of the revealed preference relation, $\rhd$: $a \trianglerighteq b$ iff $a = b$ or $a \rhd b$. *Note that $\trianglerighteq$ need **not** coincide with $\{(a, b) \in X \times X \mid (b, a) \notin \rhd\}$.* We also let

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

and

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

for any $a \in X$. The following is an important observation about revealed preference:[2]

Lemma 1 *Let p be an RCF. Then*

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

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

[2] 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. □

Thus, only the elements of A that are revealed preferred to $a \in A$ affect $p(a, A)$. Next, we have the key implication of Axiom 3:

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

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

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

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

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

The result now follows by Lemma 1. □

In a GRCG model, the default is chosen iff no other option is available and considered. Hence, given a GRCG representation $(m, \succ)$, $A \subseteq X$ and $a \in A$, $p(a^*, \succ (a) \cap A)$ is the probability that no consideration is given to anything in A that is preferred to a, and $p(a^*, A \cap (\succ (a) \cup a))$ 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 2 says that the same relationship holds for *any* RCF when we replace $\succ$ with the revealed preference relation, $\vartriangleright$, *provided the latter is asymmetric.*

Our main result is now within easy reach.

Theorem 2 *Let p be an RCF. Then p has a GRCG representation iff it satisfies WDMP and Asymmetry.*

Proof. We first verify the necessity of the axioms. Let $(m, \succ)$ be a GRCG model for p. Then, for any $A \subseteq X$,

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

which implies the WDMP axiom. Now consider Asymmetry of $\rhd$. Suppose $a, b \in X$ with $b \rhd a$, so $p(a, A) \neq p(a, A \cup b)$ for some $A \subseteq X$ with $a \in A$. Hence $b \notin A$ (and $b \neq a$ in particular) and we may deduce that $b \succ a$ as follows: if $b \notin \succ (a)$ then

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

so $p(a, A) = p(a, A \cup b)$, which is a contradiction. By the asymmetry of $\succ$ we therefore have $a \notin \succ (b)$, from which it follows that

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

for any $C \subseteq X$ with $b \in C$. Hence $p(b, C) = p(b, C \cup a)$ for any $C \subseteq X$ with $b \in C$, so $(a, b) \notin \rhd$. This establishes that p satisfies Asymmetry.

Next, we prove sufficiency. The WDMP condition, together with the fact that $p(a^*, \emptyset) = 1$, imply there exists $m \in \mathcal{C}$ such that

$$1 - p(a^*, A) = \sum_{B: B \cap A \neq \emptyset} m(B) \tag{1}$$

and hence

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

for any $A \subseteq X$. Using Lemma 2 and (1) we have, for any $A \subseteq X$ and any $a \in A$:

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

It follows that $(m, \rhd)$ is a GRCG model for p. $\square$

4.2 Interpretation

The formal similarity of a GRCG model with an RCG model obscures an important difference. If an RCF satisfies WDMP and Acyclicity then it has an RCG representation $(m, \succ)$ with $\rhd \subseteq \succ$. This part of the logic is indeed similar in the GRCG case: if an RCF satisfies WDMP and Asymmetry then it has a GRCG representation $(m, \succ)$ with $\rhd \subseteq \succ$. The converse is where a distinction appears. Given a consideration function $m \in \mathcal{C}$ and strict linear order $\succ \in \Sigma$ the RCG model $(m, \succ)$ induces an RCF. That is, the choice probabilities implied by the model will necessarily satisfy $p(A^*, A) = 1$ for every $A \subseteq X$.[3] However, given a pair $(m, \succ)$ with $m \in \mathcal{C}$ and $\succ \subseteq X \times X$ asymmetric, there may not be any RCF for which $(m, \succ)$ is a GRCG model.

Example 1 *Let $|X| = 3$. Define the consideration function m such that it assigns equal probability to each of the eight subsets of X and let $\succ = \emptyset$. If $(m, \succ)$ were a GRCG model for p then*

$$p(a, A) \;=\; \sum_{B : a \in B} m(B) \;=\; \frac{1}{2}$$

for any $A \subseteq X$ and any $a \in A$. In particular, $p(X, X) = 3/2$ so p is not an RCF.

Thus, when interpreting Theorem 2 it is important to bear in mind the *maintained assumption that p is an RCF*. This assumption could be considerably weakened in Aguiar's result but not in ours.

Suppose $m \in \mathcal{C}$ and $\succ \subseteq X \times X$ is asymmetric. In order for $(m, \succ)$ to be a GRCG model for some RCF it is necessary that, for any $A \subseteq X$, each B in the support of m must contribute to $p(x, A)$ for *exactly one* $x \in A^*$. If $\succ$ is linear, this is guaranteed: any B with $B \cap A = \emptyset$ contributes exclusively to $p(a^*, A)$ and for every B with $B \cap A \neq \emptyset$ there is a unique $\succ$-maximal $\hat{a}$ in $B \cap A$, so B contributes exclusively to $p(\hat{a}, A)$ since $\hat{a} \succ a$ for any $a \in (B \cap A) \setminus \{\hat{a}\}$. On the other hand, if $\succ$ has cycles, there may be some B with $B \cap A \neq \emptyset$ that does not contribute to $p(x, A)$ for *any* $x \in A^*$ since there is no undominated element in $B \cap A$; and if $\{(a, b), (b, a)\} \cap \succ = \emptyset$ then $B = \{a, b\} \subseteq A$ contributes to both $p(a, A)$ *and* $p(b, A)$. These observations allow us to rule out any RCF having a GRCG model $(m, \succ)$ in which m assigns positive weight to sets with "unconnected" alternatives or to sets with cycles.

[3] Moreover, if $\rhd$ is the revealed preference relation for the constructed RCF, then $\rhd \subseteq \succ$, as the reader may easily verify.

Lemma 3 *Let p be an RCF and let $(m, \succ)$ be a GRCG model for p. Then m assigns zero probability to any $B \subseteq X$ such that (i) there exist $a, b \in B$ with $a \neq b$, $a \notin \succ (b)$ and $b \notin \succ (a)$, or (ii) B contains a cycle with respect to $\succ$.*

Proof. Suppose B satisfies (i) and $m(B) > 0$. Let $A = \{a, b\} \subseteq B$ with $a \neq b$, $a \notin \succ (b)$ and $b \notin \succ (a)$. It follows that

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

and

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

Therefore

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

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 $j \neq k$,

$$E \cap A \cap (\succ (a_j) \cup a_j) = \{a_j\}$$

and

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

But this implies $\{a_j, a_k\} \subseteq E \cap A$, and hence $a_j \notin \succ (a_k)$ and $a_k \notin \succ (a_j)$, so $m(E) = 0$ for any such E. Once again, we have a contradiction. $\square$

Corollary 1 *If* $(m, \succ)$ *is a GRCG model for some RCF, then any set in the support of* m *is strictly linearly ordered by* $\succ$.

Proof. Lemma 3 means that $m(E) > 0$ only if $\succ$ is connected and acyclic on E. It follows that $\succ$ must be transitive on E. □

In light of Corollary 1, one might reasonably wonder whether the class of GRCG models is really a generalisation of the RCG class at all? The following example verifies that it is.

Example 2 *Let* $X = \{a, b, c\}$ *and consider the RCF described in the following table. Each entry gives the probability of choosing the alternative associated with the row from the choice set associated with the column. (The probabilities of choosing* a^* *are omitted but may be calculated as the column residuals, together with* $p(a^*, \emptyset) = 1$*.)*

	$\{a\}$	$\{b\}$	$\{c\}$	$\{a,b\}$	$\{b,c\}$	$\{a,c\}$	X
a	2/3	0	0	2/3	0	1/3	1/3
b	0	2/3	0	1/3	2/3	0	1/3
c	0	0	2/3	0	1/3	2/3	1/3

Let $\succ = \{(a, b), (b, c), (c, a)\}$ *and define* $m \in \mathcal{C}$ *as follows:*

$$m(\{a, b\}) = m(\{a, c\}) = m(\{b, c\}) = \frac{1}{3}$$

and $m(E) = 0$ *otherwise. It is easily checked that* $(m, \succ)$ *is a GRCG model for* p*. However,* p *does not have any RCG representation since the associated revealed preference relation is cyclic (cf., Aguiar, 2017, Theorem 1):* $a \rhd b$ *since* $p(b, \{a, b\}) < p(b, \{b\})$*,* $b \rhd c$ *since* $p(c, \{b, c\}) < p(c, \{c\})$ *and* $c \rhd a$ *since* $p(a, \{a, c\}) < p(a, \{a\})$*.*

4.3 Every GRCG model is a RUM

Aguiar (2017) shows that any RCF with an RCG representation also has a RUM (*ibid.*, Lemma 1). The same is true of any RCF with a GRCG representation.

Proposition 1 *Let* p *be an RCF. If* p *has a GRCG representation then there exists a RUM for* p*.*

Proof. Let $(m, \succ)$ be a GRCG model for p. For each $E \subseteq X$ with $m(E) > 0$ fix some $\rho^E \in \Sigma^*$ that ranks all elements of E above a^*, all elements of $\overline{E}$

below a^* and ranks the elements of E the same way that $\succ$ does. (Recall that $\succ$ induces a strict linear order on E when $m(E) > 0$: see Corollary 1.) Let $\succ^E \subseteq X \times X$ denote the binary relation described by the restriction of ρ^E to X. For each $E \subseteq X$ with $m(E) = 0$ let $\succ^E$ denote some arbitrary element of Σ.

The key observation is that if $m(E) > 0$ then for any $A \subseteq X$ and any $a \in X$,

$$A \cap E \cap \left(\succ^E (a) \cup a\right) = \{a\} \quad \Leftrightarrow \quad A \cap E \cap (\succ (a) \cup a) = \{a\} \tag{2}$$

This follows since either condition implies $a \in E$, and if $a \in E$ we have

$$E \cap \left(\succ^E (a) \cup a\right) = E \cap (\succ (a) \cup a)$$

using the fact that $\succ^E$ (strictly linearly) orders the elements of E the same way as $\succ$.

Now define $\pi \in \Delta(\Sigma^*)$ by setting $\pi\left(\rho^E\right) = m(E)$ for every E with $m(E) > 0$ and $\pi(\rho) = 0$ otherwise. Let $\hat{p}$ denote the RCF generated by the RUM, π. If $a \in A \subseteq X$ then

$$\begin{aligned}
\hat{p}(a, A) &= \pi(\{\rho \in \Sigma^* \mid \rho(a) \leq \rho(x) \text{ for all } x \in A^*\}) \\
&= \pi\left(\left\{\rho^E \mid m(E) > 0 \text{ and } A \cap E \cap \left(\succ^E (a) \cup a\right) = \{a\}\right\}\right) \\
&= \sum_{E: A \cap E \cap (\succ^E (a) \cup a) = \{a\}} m(E) \\
&= \sum_{E: A \cap E \cap (\succ (a) \cup a) = \{a\}} m(E)
\end{aligned}$$

where the final equality uses (2). Thus, $\hat{p} = p$. □

It is easily checked that a RUM for the RCF in Example 2 is obtained by assigning probability $\frac{1}{3}$ to each element of $\{\rho_1, \rho_2, \rho_3\} \in \Sigma^*$ where:

$$\rho_1(a) = 1,\ \rho_1(b) = 2,\ \rho_1(a^*) = 3,\ \rho_1(c) = 4$$

$$\rho_2(b) = 1,\ \rho_2(c) = 2,\ \rho_2(a^*) = 3,\ \rho_2(a) = 4$$

and

$$\rho_3(c) = 1,\ \rho_3(a) = 2,\ \rho_3(a^*) = 3,\ \rho_3(b) = 4.$$

However, not every RCF with a RUM has a GRCG representation.

Example 3 *Let $X = \{a, b, c\}$ and consider the RUM that chooses each of the following strict linear orders with probability $\frac{1}{2}$:*

$$\rho(a) = 1,\ \rho(b) = 2,\ \rho(c) = 3,\ \rho(a^*) = 4$$

and

$$\hat{\rho}(c) = 1,\ \hat{\rho}(b) = 2,\ \hat{\rho}(a) = 3,\ \hat{\rho}(a^*) = 4.$$

Let p denote the RCF generated by this RUM. Then p may be depicted as follows (as usual, omitting the probabilities of choosing a^*):

	$\{a\}$	$\{b\}$	$\{c\}$	$\{a,b\}$	$\{b,c\}$	$\{a,c\}$	X
a	1	0	0	1/2	0	1/2	1/2
b	0	1	0	1/2	1/2	0	0
c	0	0	1	0	1/2	1/2	1/2

We deduce that p has no GRCG representation from the fact that the revealed preference relation $\rhd$ obtained from p is not asymmetric: $a \rhd b$ since $p(b, X) < p(b, \{b, c\})$ and $b \rhd a$ since $p(a, \{a, b\}) < p(a, \{a\})$.

Thus, the set of RCF's with an RCG representation is a proper subset of those with a GRCG representation, which in turn is a proper subset of those with a RUM. To explore these relationships in more detail, some additional notation will be useful.

Let $\mathcal{P}$ denote the set of RCF's with a GRCG representation and $\mathcal{P}'$ the proper subset of $\mathcal{P}$ with an RCG representation. Next, for every $(E, \succ)$ with $E \subseteq X$ and $\succ \subseteq X \times X$ strictly linear on E, let $\Sigma^*_{(E,\succ)} \subseteq \Sigma^*$ be defined as follows: $\rho \in \Sigma^*_{(E,\succ)}$ iff ρ ranks everything in E above a^*, everything in $\overline{E}$ below a^* and everything in E the same way as $\succ$. Finally, given $\succ \subseteq X \times X$ let

$$\Sigma^*_{\succ} = \bigcup \left\{ \Sigma^*_{(\succ,E)} \mid \succ \text{ is strictly linear on } E \subseteq X \right\}.$$

Note that $\Sigma^*_{(\succ,E)} \cap \Sigma^*_{(\succ,E')} = \emptyset$ for any distinct $E, E' \in 2^X$ on which $\succ$ is strictly linear.

Theorem 3 *Let p be an RCF. Then $p \in \mathcal{P}$ iff there is an asymmetric binary relation $\succ \subseteq X \times X$ and a RUM, $\pi \in \Delta(\Sigma^*)$, for p with $\pi(\Sigma^*_{\succ}) = 1$ and $\pi(\rho) > 0$ for at most one $\rho \in \Sigma^*_{(\succ,E)}$ for every $E \subseteq X$ on which $\succ$ is strictly linear.*

Proof. The "only if" part follows by the construction in the proof of Proposition 1. For the "if" part, let π denote a RUM for p satisfying the indicated support restriction for some asymmetric binary relation $\succ \subseteq X \times X$. If $\succ$ is strictly linear on E and there is some $\rho \in \Sigma^*_{(\succ,E)}$ with $\pi(\rho) > 0$, let

$m(E) = \pi(\rho)$; let $m(E) = 0$ otherwise. In particular, if $m(E) > 0$ then $\succ$ is strictly linear on E and every $\hat{\rho} \in \Sigma^*_{(\succ,E)}$ satisfies $\hat{\rho}(x) < \hat{\rho}(a^*)$ iff $x \in E$. It follows that for any $A \subseteq X$ and any $a \in A$ we have:

$$\begin{aligned} p(a, A) &= \pi(\{\rho \in \Sigma^* \mid \rho(a) \leq \rho(x) \text{ for all } x \in A^*\}) \\ &= \sum_{E: A \cap E \cap (\succ(a) \cup a) = \{a\}} m(E). \end{aligned}$$

Thus, $(m, \succ)$ is a GRCG model for p. □

Corollary 2 *Let p be an RCF. Then $p \in \mathcal{P}'$ iff there is some $\succ \in \Sigma$ and a RUM, $\pi \in \Delta(\Sigma^*)$, for p with $\pi(\Sigma^*_{\succ}) = 1$ and $\pi(\rho) > 0$ for at most one $\rho \in \Sigma^*_{(\succ,E)}$ for every $E \subseteq X$.*

Proof. The "only if" part follows by the construction in the proof of Proposition 1, noting that $\succ \in \Sigma$ if $p \in \mathcal{P}'$. The "if" part follows directly from Theorem 3. □

As Example 2 shows, there exists an RCF p with GRCG model $(m, \succ)$ together with subsets $E, E', E'' \in 2^X$ such that $\succ$ is strictly linear on E, E' and E'' but there is no $\hat{\succ} \in \Sigma$ that orders each of E, E' and E'' the same way as $\succ$.

References

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

[2] Block, H. D., and Marschak, J. (1960). Random Orderings and Stochastic Theories of Response, in: *Contributions to Probability and Statistics, I.* Olkin, S. Ghurye, W. Hoeffding, W. Madow, H. Mann (Eds.), Stanford Univesity Press, pp. 97-132.

[3] 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(3)**, 263-283.

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

[5] Gibbard, P. (2017). Disentangling preferences and limited attention: Random utility models with consideration sets. Unpublished manuscript, Australian National University.

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