

# Gradual Pairwise Comparison and Stochastic Choice\*

Rohan Dutta<sup>†</sup>

January 11, 2020

## Abstract

Guided by evidence from eye-tracking studies of choice, pairwise comparison is assumed to be the building block of the decision-making procedure. A decision-maker with a rational preference may nevertheless consider the constituent pairwise comparisons gradually, easier comparisons preceding difficult ones. Facing a choice problem she may be unable to complete all relevant comparisons and choose with equal odds from alternatives not found inferior. Stochastic choice data consistent with such behaviour is characterized and used to infer the underlying preference relation and the order of pairwise comparisons. The choice procedure offers a novel rationale for behavioural phenomena such as the similarity effect and violations of stochastic transitivity and regularity.

**Keywords:** Revealed preference, bounded rationality, stochastic choice.

---

\*I thank Sean Horan, David Levine, Yusufcan Masatlioglu, Salvatore Modica, Paulo Natenzon, Pietro Ortoleva and Anna Winterbottom for helpful comments. I also thank the editor, Ran Spiegler, and the three anonymous referees for the detailed reports. The paper has been helped immensely by their engagement.

<sup>†</sup>Department of Economics, McGill University, Leacock 531, 855 Sherbrooke St. W, Montreal, QC, H3A 2T7, Canada and CIREQ; 514-398-5611; rohan.dutta@mcgill.ca.

# 1 Introduction

The literature on eye-tracking analysis of multialternative choice, pioneered by Russo and Rosen (1975), offers evidence of the actual choice procedure consisting primarily of a sequence of pairwise comparisons.<sup>1</sup> For a rational agent capable of considering all relevant comparisons (those that shrink the set of options being considered) before making a choice, the simultaneity or sequentiality of such comparisons is irrelevant. This is untrue if the agent is often, for unobservable reasons, unable to complete all relevant comparisons.

I study such a boundedly rational agent who considers the relevant pairwise comparisons of her underlying strict rational preference sequentially to remove inferior alternatives from a given choice set. This sequence is menu-independent in the particular sense that if two different pairwise comparisons are both relevant in two distinct choice problems then they are considered in the same order in both. The set of relevant pairwise comparisons, however, is menu dependent. Facing a choice problem the agent may be forced to stop at different points along this sequence of relevant comparisons according to some unobserved and menu-dependent probability distribution. While she is able to make all relevant comparisons with positive probability, it is not certain. Upon stopping she chooses from alternatives that have not been removed from the choice set with equal odds. Choice resulting from this procedure is called a *Gradual Pairwise Comparison Rule* (GPCR).

The random nature of stopping along the sequence makes the choice behaviour stochastic. Such boundedly rational behaviour is consistent with a rich set of (stochastic) choice data, including deterministic rational choice, Luce rules (Luce (1959)) and Additive Perturbed Utility rules (Fudenberg et al. (2015)), but also choice data where the order of choice probabilities across alternatives is menu-dependent.

The assumption of menu-independence of the order of relevant comparisons follows

---

<sup>1</sup>See also Russo and Leclerc (1994) and the review of eye-tracking research in Wedel and Pieters (2008). For more recent work, see Noguchi and Stewart (2014, 2018).

from a more basic assumption that the order reflects the agent's relative ease of making such comparisons, with easier comparisons preceding difficult ones. This relative ease may be subjective and known to the agent alone. Nevertheless, it ensures that the order of comparison is menu-independent (see Section 5.2). This order can be inferred from choice data (Theorem 5).

The random stopping could arise from different sources such as fatigue (from making multiple comparisons) or unobserved time constraints. Section 4 describes how GPCRs can exhibit violations of strong (and moderate) stochastic transitivity, the similarity effect and regularity violations, thereby suggesting one avenue through which fatigue or time constraints could yield such non-standard behaviour.

The primary objective of this study is to carefully analyze a simple decision procedure built upon the empirical finding that choice involves a sequence of pairwise comparisons and the idea that easier comparisons precede difficult ones.<sup>2,3</sup> The exercise is made more compelling by the ability of this procedure to explain disparate behavioural phenomena despite an underlying stable rational preference. The latter facilitates standard welfare analysis.

The ordinal content of choice probabilities in a GPCR is essentially determined by the sequence of pairwise comparisons (Theorem 1). Changing the choice probabilities of a GPCR without changing the ordinal content, leads to a new GPCR where only the random stopping specification needs changing (Theorem 2). As a result, checking whether some choice data is a GPCR amounts to verifying if there exists a sequence of pairwise comparisons that can generate the required choice ranks.

The model is characterized by three simple axioms (Theorem 3) and each GPCR is shown to correspond to a unique underlying strict preference (Theorem 4). The latter is easily identified with the agent strictly preferring  $a$  to  $b$  if and only if  $a$  has

---

<sup>2</sup>Ravid and Steverson (2018) also studies a choice procedure involving pairwise comparisons but with very different structure and implication, as discussed in section 5.1.

<sup>3</sup>Using empirical guidance to derive appropriate properties of a choice procedure, instead of suitable restrictions on choice data, is less common in economics research but has a tradition of its own. See, for instance, Simon (1955), Rubinstein (1988) and Manzini and Mariotti (2007).

the highest choice probability in some set containing  $b$ .

Multiple sequences of pairwise comparisons can be consistent with the same GPCR. It may be, though, that in all such representations, certain pairwise comparisons must be considered before some others. These are fully identified by using a revealed preference approach (Theorem 5).

This study owes a considerable debt to Apesteguia and Ballester (2013) and Manzini and Mariotti (2012). Not only did they introduce the framework of choice resulting from a sequence of pairwise comparisons on which the current model is built, they also made the valuable finding that identifying from choice data the first relevant pairwise comparison for any (collection of) choice set(s), is key to characterizing their sequential procedures. This idea is essential in the current setting too and is captured by one of the three axioms that characterize the model.

The rest of the paper is as follows. Section 2 defines the GPC choice procedure and discusses an example. Section 3 contains all the characterization results. The proof of theorem 1 is retained in the main body of the text to give the reader a better sense of how the model works. All other proofs are collected in the appendix, along with the discussion on the independence of axioms. Section 4 discusses the specific ways in which GPCRs can accommodate violations of stochastic transitivity and other forms of menu-dependent choice. Section 5 discusses the key components of the GPC procedure and how they relate to other models of boundedly rational choice and stochastic choice.

## 2 Stochastic Choice and Procedures

### 2.1 Preliminaries

Consider a nonempty finite set of alternatives  $X$  and let  $\mathcal{X}$  be the set of all non-empty subsets of  $X$ . These are the choice sets the decision-maker faces. The decision-maker is assumed to have a strict rational preference. This is captured by a binary relation,

$P \subseteq X \times X$ , where  $(a, b) \in P$  means that  $a$  is strictly preferred to  $b$ .<sup>4</sup> It will often be convenient to represent this binary relation by  $\succ$  where  $a \succ b \equiv (a, b) \in P$ . Denote the set of all strict rational preferences over  $X$  as  $\mathcal{P}$ .

**Definition 1.** A *stochastic choice rule* is a function  $p : X \times \mathcal{X} \rightarrow [0, 1]$  such that  $\sum_{a \in A} p(a, A) = 1$  for all  $A \in \mathcal{X}$  and  $p(a, A) = 0$  for all  $a \notin A$ .

Here  $p(a, A)$  is the probability with which  $a$  is chosen when the decision-maker faces the choice set  $A$ . Stochastic choice rules are clearly more general than deterministic ones, which in addition require  $p(a, A) \in \{0, 1\}$ . More importantly, they better accommodate observed choice data in that they can represent the relative observed choice frequencies obtained from repeated choices by the decision maker. Let a *stochastic choice rule without ties* be a stochastic choice rule  $p$  such that  $p(a, A) \neq p(b, A)$  for all  $a, b \in A \in \mathcal{X}$ , with  $a \neq b$ . These choice rules turn out to be particularly useful in the characterization results that follow.

## 2.2 Gradual Pairwise Comparison

The choice procedure of gradual pairwise comparison (GPC) is as follows. The decision-maker, endowed with a strict rational preference  $P$ , does not consider all the binary comparisons in  $P$  simultaneously. Instead, she has an *ordered partition*  $\mathbb{P} = \{P_i\}_{i=1}^I$  of  $P$ , in that:  $P_j \cap P_k = \emptyset$  for  $j \neq k$ ,  $\bigcup_i P_i = P$  and  $P_i \neq \emptyset$  for all  $1 \leq i \leq I$ . Let  $P_0 = \emptyset$ . A pairwise comparison  $(a, b)$  is *relevant* given a set of alternatives,  $A$ , if  $\{a, b\} \subseteq A$ . Given a choice set, the agent considers all the relevant pairwise comparisons in  $P_1$  simultaneously, eliminating all alternatives found inferior. With the alternatives that survive she then considers the relevant comparisons (given the set of surviving alternatives) in  $P_2$  and so on. For a given ordered partition  $\mathbb{P}$  of  $P$

---

<sup>4</sup>As a strict rational preference,  $P$  must be *asymmetric*:  $(a, b) \in P \Rightarrow (b, a) \notin P$ , *complete*: for all  $a, b \in X$  if  $a \neq b$  then either  $(a, b) \in P$  or  $(b, a) \in P$  and *transitive*:  $(a, b) \in P$  and  $(b, c) \in P \Rightarrow (a, c) \in P$ .

and a choice set  $A \in \mathcal{X}$ , define the following sets recursively,

$$\begin{aligned} M_0^{\mathbb{P}}(A) &= A \\ M_i^{\mathbb{P}}(A) &= \{x \in M_{i-1}^{\mathbb{P}}(A) \mid \forall y \in M_{i-1}^{\mathbb{P}}(A), (y, x) \notin P_i\}, \quad \forall 1 \leq i \leq I. \end{aligned} \quad (1)$$

$M_i^{\mathbb{P}}(A)$  contains all alternatives that survive after the decision maker has considered the  $i$ th cell of her ordered partition.

For any choice set  $A$ , let  $\tilde{I}^{\mathbb{P}}(A)$  be the cell of the partition that finally reduces the surviving options to a singleton. Formally,  $\tilde{I}^{\mathbb{P}}(A) = i \leq I$  such that  $|M_i^{\mathbb{P}}(A)| = 1$  and either  $|M_{i-1}^{\mathbb{P}}(A)| > 1$  or  $i = 1$ .<sup>5</sup>  $\tilde{I}^{\mathbb{P}}$  is well defined since  $\mathbb{P}$  is a partition of a strict rational preference  $P$ .<sup>6</sup>

If the decision-maker could complete all comparisons, her choice would coincide with deterministic rational choice. Her (possible) inability to do so is captured by a function  $\pi : (\mathbb{P} \cup \{P_0\}) \times \mathcal{X} \rightarrow [0, 1]$ , such that  $\sum_{P_i \in \mathbb{P}} \pi(P_i, A) + \pi(P_0, A) = 1$  and  $\pi(P_{\tilde{I}^{\mathbb{P}}(A)}, A) > 0$  for all  $A \in \mathcal{X}$ , labeled **stopping function**.<sup>7</sup> For any choice set  $A$ ,  $\pi(P_i, A)$  is the probability that the decision maker stops at cell  $P_i$  and is unable to complete the comparisons contained in subsequent cells.  $\pi(P_0, A)$  is the probability with which she is unable to complete any relevant comparison at all. While the premise of this study is that a decision-maker may be unable to make all relevant comparisons, assuming  $\pi(P_{\tilde{I}^{\mathbb{P}}(A)}, A) > 0$  for all  $A \in \mathcal{X}$  requires that her ability to do so cannot be ruled out entirely either. It says that the decision-maker is able to make all relevant comparisons with positive probability. In deterministic rational choice this probability would have to be 1.

Conditional on stopping after considering cell  $P_i$ , the procedure entails the decision-maker choosing with equal odds from among the alternatives that remain,  $M_i^{\mathbb{P}}(A)$ .

<sup>5</sup> $|B|$  denotes the number of elements in the set  $B$ .

<sup>6</sup>For any  $i < I$ , by definition,  $|M_i^{\mathbb{P}}(A)| \geq |M_{i+1}^{\mathbb{P}}(A)|$ .  $\mathbb{P}$  being a partition of a strict rational preference  $P$  implies that  $M_I^{\mathbb{P}}(A)$  is a singleton (containing the most preferred alternative in  $A$ ).

<sup>7</sup>The obvious dependence of the stopping function on the partition  $\mathbb{P}$  is suppressed for notational convenience.

**Definition 2.** A *gradual pairwise comparison rule* (GPCR) is a stochastic choice rule  $p^{\mathbb{P},\pi}$  with an ordered partition  $\mathbb{P}$  of a strict rational preference  $P$  and a stopping function  $\pi$  such that for all  $A \in \mathcal{X}$ ,

$$p^{\mathbb{P},\pi}(a, A) = \begin{cases} \sum_{\{i|a \in M_i^{\mathbb{P}}(A)\}} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} & \text{if } a \in A \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

A *simple pairwise comparison rule* (SPCR) is a GPCR,  $p^{\mathbb{P},\pi}$ , for which each cell of the ordered partition  $\mathbb{P}$  is a singleton. Formally,  $|P_i| = 1$  for all  $P_i \in \mathbb{P}$ . A stochastic choice rule  $p$ , is *rationalizable by gradual pairwise comparison*, if there exists an ordered partition  $\mathbb{P}$  of a preference  $P \in \mathcal{P}$  and a stopping function  $\pi$  such that  $p = p^{\mathbb{P},\pi}$ .

To see how the procedure works consider the following example.

**Example 1.** There are three lotteries,  $a$ ,  $b$  and  $c$ :

$$\begin{array}{l|l} a & (7100, 1005; \frac{3}{4}, \frac{1}{4}) \\ b & (7000, 1000; \frac{3}{4}, \frac{1}{4}) \\ c & (6650, 490; \frac{4}{5}, \frac{1}{5}) \end{array}$$

Read the table above as lottery  $a$  yields \$7100 with probability 3/4 and \$1005 with probability 1/4, and so on. The agent's underlying preference ranks  $a$  over  $b$  over  $c$ . Nevertheless, she is able to make the comparison between  $a$  and  $b$  the earliest, followed by the one between  $a$  and  $c$  and then  $b$  and  $c$ .

Formally,  $X = \{a, b, c\}$ . The decision-maker's preference is  $a \succ b \succ c$ . In other words,  $P = \{(a, b), (a, c), (b, c)\}$ . The ordered partition  $\mathbb{P}$  is

$$\begin{array}{c|c|c|c} P_0 & P_1 & P_2 & P_3 \\ \hline & (a, b) & (a, c) & (b, c) \end{array}$$

Her stopping function is the following.

$\pi(\cdot, A)$	$A = \{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
$P_0$	0	2/3	2/3	0
$P_1$	1	0	0	4/5
$P_2$	0	1/3	0	1/5
$P_3$	0	0	1/3	0

The resulting choice probabilities through gradual pairwise comparison are the following.

$p^{\mathbb{P}, \pi}(\cdot, A)$	$A = \{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
$a$	1	2/3	0	3/5
$b$	0	0	2/3	0
$c$	0	1/3	1/3	2/5

For instance,  $p^{\mathbb{P}, \pi}(b, \{b, c\}) = 2/3$  and  $p^{\mathbb{P}, \pi}(c, \{a, b, c\}) = 2/5$ .

The stopping function captures the agent's ability to complete the comparison  $(a, b)$ , whenever relevant. By contrast, the agent finds the comparison  $(a, c)$  difficult. The probability with which she is unable to complete this comparison when facing the choice set  $\{a, c\}$  is captured by  $\pi(P_0, \{a, c\})$ . With the choice set  $\{a, b, c\}$ , the agent again makes the comparison  $(a, b)$  without fail. However, she is now unable to complete the  $(a, c)$  comparison with higher probability,  $\pi(P_1, \{a, b, c\})$ . Such a stopping function specification therefore captures something akin to fatigue (or time constraints) where successfully carrying out one comparison reduces the chances of completing subsequent comparisons.

**Remark 1.** Note that when choosing from  $\{b, c\}$  the only relevant comparison is  $(b, c)$ ; the agent does not think of any other comparisons. When choosing from  $\{a, b, c\}$ , the relevant comparisons are  $(a, b)$  followed by  $(a, c)$ , while  $(b, c)$  is not considered. Clearly, the set of relevant comparisons is menu-dependent. Nevertheless,



the order of relevant comparisons is menu-independent, in that if there were a larger choice set in which  $(a, b)$  and  $(a, c)$  were relevant comparisons, then  $(a, b)$  would precede  $(a, c)$ . It is this menu-independence of the order of relevant comparisons that allows for a menu-independent specification of the ordered partition, with the understanding that the agent, for a given choice set, only considers the relevant comparisons in each cell of the partition, sequentially.<sup>8</sup>

**Remark 2.** The model allows for a very general class of stopping rules. For instance, with the same ordered partition as in example 1, the model allows an agent to stop with equal probability at every cell of the partition irrespective of the choice set,  $\pi(P_i, A) = 1/4$  for all  $A \in \mathcal{X}$  and all  $i \in \{0, 1, 2, 3\}$ . This specification makes little sense, since for instance, when choosing from the set  $\{b, c\}$  the agent does not think about the comparisons  $(a, b)$  and  $(a, c)$ ; so stopping at  $P_1$  or  $P_2$  does not correspond to any literal description of the choice procedure. Indeed, the following subset of stopping functions captures more accurately the decision procedure;  $\pi : (\mathbb{P} \cup \{P_0\}) \times \mathcal{X} \rightarrow [0, 1]$  is an *exact stopping function* if it is a stopping function such that  $\pi(P_i, A) = 0$  if  $M_i^{\mathbb{P}}(A) = M_{i-1}^{\mathbb{P}}(A)$ , for all  $1 \leq i \leq I$ . It turns out, though, that the set of choice rules rationalizable by gradual pairwise comparison is identical to those rationalizable by GPC using exact stopping functions.

**Observation 1.** *A stochastic choice rule  $p$  is rationalizable by gradual pairwise comparison if and only if  $p = p^{\mathbb{P}, \pi'}$ , where  $\pi'$  is an exact stopping function.*

Relying on this observation, the subsequent analysis continues to use the more general stopping functions, which are easier to describe in proofs. Exact stopping functions are used in all examples. Finally, choosing between these two classes of stopping functions has no impact on any of the results that follow.

---

<sup>8</sup>A similar feature is at work in the representation of strict rational preferences. The menu-independent ranking of any two available alternatives, allows for a menu-independent representation of the entire preference ordering, with the understanding that the agent, for a given choice set, maximizes this preference only from among the available alternatives.

### 3 Characterization

#### 3.1 Choice Probabilities and Choice Rank

The ordinal content of choice probabilities is labeled *choice rank*. This is the ranking of alternatives in a choice set generated by the choice probabilities, with a higher choice probability corresponding to a higher rank. Since the mapping from choice probabilities to choice rank is many-to-one, knowledge of choice rank alone cannot pin down the exact value of choice probabilities.

The two key components of a GPC procedure, the ordered partition  $\mathbb{P}$  and the stopping function  $\pi$ , play very different roles in determining choice rank and choice probabilities. Choice rank is essentially determined by  $\mathbb{P}$ , the sequence in which the constituent pairwise comparisons of the rational preference are considered (and therefore the underlying preference too).

**Theorem 1.** *Fix an ordered partition  $\mathbb{P}$  of some  $P \in \mathcal{P}$ . Let  $\pi$  and  $\pi'$  be stopping functions on  $\mathbb{P}$ . Then for any  $A \in \mathcal{X}$  and  $a, b \in X$ ,*

$$p^{\mathbb{P},\pi}(a, A) > p^{\mathbb{P},\pi}(b, A) \Rightarrow p^{\mathbb{P},\pi'}(a, A) \geq p^{\mathbb{P},\pi'}(b, A).$$

*Proof.* It follows from the definition of these sets in (1) that for any  $\mathbb{P}$ ,  $M_j^{\mathbb{P}}(A) \supseteq M_k^{\mathbb{P}}(A)$  for all  $j, k \leq I$  with  $j < k$ . Also for any  $a, b \in A$ , one of the sets  $\{i|a \in M_i^{\mathbb{P}}(A)\}$  and  $\{i|b \in M_i^{\mathbb{P}}(A)\}$  must be a subset of the other. Therefore

$$\begin{aligned} p^{\mathbb{P},\pi}(a, A) &= \sum_{\{i|a \in M_i^{\mathbb{P}}(A)\}} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} > \sum_{\{i|b \in M_i^{\mathbb{P}}(A)\}} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} = p^{\mathbb{P},\pi}(b, A) \\ &\Rightarrow \{i|a \in M_i^{\mathbb{P}}(A)\} \supset \{i|b \in M_i^{\mathbb{P}}(A)\} \\ \Rightarrow p^{\mathbb{P},\pi'}(a, A) &= \sum_{\{i|a \in M_i^{\mathbb{P}}(A)\}} \frac{\pi'(P_i, A)}{|M_i^{\mathbb{P}}(A)|} \geq \sum_{\{i|b \in M_i^{\mathbb{P}}(A)\}} \frac{\pi'(P_i, A)}{|M_i^{\mathbb{P}}(A)|} = p^{\mathbb{P},\pi'}(b, A) \end{aligned}$$

□

In other words, changing the stopping function cannot reverse strict choice ranks.

Further, any stochastic choice rule that always has a unique most probable alternative and is consistent with the choice ranks of some GPCR can be rationalized by choosing an appropriate stopping function while leaving the ordered partition of the original GPCR unchanged.

**Axiom 1 (Unique Best).** *For all  $A \in \mathcal{X}$  there exists  $a \in A$  such that*

$$p(a, A) > p(b, A), \quad \forall b \in A \setminus \{a\}.$$

**Theorem 2.** *Fix an ordered partition  $\mathbb{P}$  of some  $P \in \mathcal{P}$  and a stopping function  $\pi$ . Let  $p$  be a stochastic choice rule that satisfies unique best and for all  $A \in \mathcal{X}$  and  $a, b \in X$*

$$\begin{aligned} p^{\mathbb{P}, \pi}(a, A) > p^{\mathbb{P}, \pi}(b, A) &\Rightarrow p(a, A) \geq p(b, A) && \text{and} \\ p^{\mathbb{P}, \pi}(a, A) = p^{\mathbb{P}, \pi}(b, A) &\Rightarrow p(a, A) = p(b, A). \end{aligned}$$

*Then there exists  $\pi'$  such that  $p = p^{\mathbb{P}, \pi'}$ .*

Taken together, the two theorems above show that the key step to rationalizing a stochastic choice rule by gradual pairwise comparison is to obtain an ordered partition of the underlying preference that generates the required choice ranks. It is then guaranteed that there exists an appropriate stopping function for the remaining task of matching the exact choice probabilities.

## 3.2 General Characterization

Deterministic rational choice is a particular case of the GPC procedure. Indeed, for any preference  $P \in \mathcal{P}$ , and any ordered partition  $\mathbb{P}$  of it, setting  $\pi(P_{\tilde{I}^{\mathbb{P}}(A)}, A) = 1$  ensures that  $p^{\mathbb{P}, \pi}(a, A) = 1$  if  $a$  is the most preferred element in  $A$  according to  $P$  and

$p^{\mathbb{P},\pi}(a, A) = 0$  otherwise. This is unsurprising, since setting  $\pi(P_{\mathcal{I}^{\mathbb{P}}(A)}, A) = 1$  implies that the agent is able to make all relevant comparisons before making her decision.

More interestingly, the GPC procedure can rationalize choice reversals. For instance, in example 1, adding the alternative  $a$  to the choice set  $\{b, c\}$ , increases the probability of  $c$  being selected from  $1/3$  to  $2/5$  while reducing that of  $b$  from  $2/3$  to  $0$ . Commonly used stochastic choice rules such as the Luce rule cannot allow such choice reversals. In fact, the increased probability of  $c$  violates *regularity*, a property that requires  $p(a, A) \geq p(a, B)$  for all  $a \in A \subseteq B$ . This puts the GPC choice procedure outside the scope of Random Utility Models, which necessarily satisfy regularity.

It is natural, then, to wonder whether the GPC procedure has any empirical content. Indeed, it does and it can be characterized. Start by defining an appropriate notion of revealed preference.

**Definition 3.** *Given a stochastic choice rule  $p$  and  $a, b \in X$ ,  $a$  is **stochastically revealed preferred** to  $b$  if  $\exists A \in \mathcal{X}$  such that  $a, b \in A$  and*

$$p(a, A) \geq p(d, A), \quad \forall d \in A.$$

Note that it is not enough for  $a$  to simply have a higher choice probability than  $b$  to be revealed preferred to it;  $a$  must be the most probable alternative in the presence of  $b$ . This leads to the most immediate testable implication of the GPC procedure, labeled *stochastic weak axiom of revealed preference* or *sWARP*.

**Axiom 2 (sWARP).** *For all  $a, b \in X$ , if  $a$  is stochastically revealed preferred to  $b$  then  $b$  is not stochastically revealed preferred to  $a$ .*

In words, alternatives that were not the most probable in a given set, cannot become the most probable in the presence of the original most probable alternative. Given the assumption of strict preferences, an immediate implication of *sWARP* is that for any choice set there is a unique alternative with the highest choice probability. *sWARP* relates entirely to how the most probable alternative varies across choice

sets. It imposes no restriction on the choice probabilities or even the choice rank of alternatives that are not the most probable. The testable implications of GPC on these are more subtle.

**Definition 4.** *Given a stochastic choice rule  $p$ ,  $B$  is a  **$p$ -truncation** of  $A \in \mathcal{X}$  if  $B \subseteq A$  and*

$$a \in B, b \in A \setminus B \Rightarrow p(a, A) > p(b, A).$$

Let  $A^T(p)$  be the set of all  $p$ -truncations of  $A \in \mathcal{X}$ . In words, any set of the  $m$  highest (choice) ranked alternatives in the set  $A$  under  $p$  is a  $p$ -truncation of  $A$ . Note that  $A \in A^T(p)$ . Then any stochastic choice rule without ties that is rationalizable by GPC must satisfy the following property, labeled ***Invariance to Truncation by Rank (ITR)***.

**Axiom 3 (ITR).** *If  $D \in A^T \cap B^T$  for some  $A, B \in \mathcal{X}$  then,*

$$p(a, A) > p(b, A) \Leftrightarrow p(a, B) > p(b, B), \quad \forall a, b \in D.$$

ITR simply requires that if the same set of alternatives  $D$ , makes up the top  $n$  choice ranks in two different sets then the choice rank of any alternative in  $D$  must be the same in the two sets. Compare this to the far more restrictive Luce's IIA, which requires not only that the choice rank ordering of *any* pair of alternatives is the same across *any* two sets where they both belong but that the ratio of their choice probabilities is menu-independent too.

Consider the universe of choice procedures in which pairwise comparisons are used sequentially to eliminate alternatives, with an earlier elimination corresponding to a worse choice rank. ITR requires that along such a procedure an eliminated alternative should have no bearing on the choice ranks of the alternatives that remain. The GPC procedure satisfies this by virtue of two of its features. First, once an alternative is eliminated, any comparison between the latter and the remaining alternatives is rendered irrelevant. Second, the order of relevant comparisons is menu-independent.

The final axiom is labeled s(tochastic)-reducibility. Let  $\tilde{\mathcal{X}}$  denote the subset of  $\mathcal{X}$  containing all subsets of  $X$  with at least two elements.

**Axiom 4 (s-Reducibility).** *For every nonempty collection of sets  $\mathcal{B} \subseteq \tilde{\mathcal{X}}$ , there exists  $D \in \mathcal{B}$  and  $\{a, b\} \subseteq D$  such that if  $\{a, b\} \subseteq A \in \mathcal{B}$  then*

$$p(c, A) > p(b, A), \quad \forall c \in A \setminus \{b\}.$$

*s-reducibility* says that for any collection of choice sets, there must exist a pair of alternatives  $a$  and  $b$ , such that  $b$  is always (choice) ranked last in a set in this collection whenever  $a$  is also present. Compare this to Luce's IIA, under which any collection of sets must have an alternative that is ranked last whenever available in a set in this collection. *s-reducibility* is a lot weaker, in that  $b$  need not be the lowest ranked whenever available in a set in the collection, but only so in the presence of  $a$ .

Consider again the universe of choice procedures using sequential pairwise comparisons to eliminate alternatives. *s-reducibility* requires that for any collection of choice sets there exists a pairwise comparison that is the first relevant one for the collection. In the GPC procedure this follows from the menu-independence of the order of relevant comparisons.

Not only are *sWARP*, *ITR* and *s-reducibility* necessary, but they are also jointly sufficient for a stochastic choice rule without ties to be rationalizable by GPC.

**Theorem 3.** *A stochastic choice rule without ties  $p$ , is rationalizable by gradual pairwise comparison if and only if  $p$  satisfies *sWARP*, *ITR* and *s-reducibility*.*

The three axioms characterizing the model can be thought of as corresponding to three separate implications of the Order Independence axiom of Tversky (1972a), which is weaker than both Luce's IIA and the Acyclicity axiom in Fudenberg, Iijima and Strzalecki (2015), and requires the choice rank ordering of any pair of alternatives to be the same across all sets that contain them both. The first implication of order independence for a choice rule without ties is exactly *sWARP*, which is retained. The

second implication is the existence in any collection of sets of a worst alternative, which as a result always ranks last whenever available in a set in the collection. This is weakened to *s-reducibility*. The final implication has to do with alternatives that are not the most or least probable, requiring the choice rank ordering of any two such alternatives to be menu-independent. This is weakened to ITR.

A general stochastic choice rule (with possible ties) that is rationalizable by GPC must always have a unique most probable outcome for any choice set. Beyond this, characterizing stochastic choice rules that are rationalizable by GPC turns out to be equivalent to asking whether the ties (if any) in such a choice rule can be broken consistently to arrive at a stochastic choice rule without ties which satisfies sWARP, ITR and *s-reducibility*. The formal result (Theorem 6) is relegated to the appendix.

*s-reducibility* is similar in spirit and purpose to the *reducibility* axiom introduced in Manzini and Mariotti (2012), which fully characterized the deterministic model of acyclic sequentially rationalizable choice. *Reducibility* requires for any collection of sets the existence of pair of alternatives  $a, b$  such that in the presence of  $a$  removing  $b$  from any set in the collection has no effect on the deterministic choice from those sets. Despite the difference in the property required of the pair of alternatives the two axioms seek, their purpose is the same. Both identify the first pairwise comparison in the sequential procedure that is relevant to any set in the collection.<sup>9</sup> This is a critical ingredient in the construction used to prove the sufficiency part of theorem 3.

### 3.3 Revealed Preference and Order of Comparison

Given choice data consistent with a decision-maker using the GPC procedure, it is very easy to infer the unique underlying rational preference relation. Indeed, it is the same as the stochastically revealed preferred relation.

---

<sup>9</sup>The approach taken in Apesteguia and Ballester (2013) is different, but with a similar purpose.

**Theorem 4.** Given a GPC choice rule  $p^{\mathbb{P},\pi}$  where  $\mathbb{P}$  is an ordered partition of  $P \in \mathcal{P}$ ,

$$(a, b) \in P \Leftrightarrow a \text{ is stochastically revealed preferred to } b.$$

Different pairs of  $\mathbb{P}$  and  $\pi$  can rationalize the same stochastic choice rule. For instance, deterministic rational choice with preference  $P$  can be rationalized by any ordered partition  $\mathbb{P}$  of  $P$  and  $\pi$  such that  $\pi(P_{\tilde{I}(A)}, A) = 1$ . So while the choice data always uniquely pins down the underlying preference relation, in this case it offers no clue about the particular order in which the decision-maker considers the pairwise comparisons. Now consider the following choice data.

**Example 1 revisited.** Recall the choice probabilities generated by the GPC choice procedure described in example 1, in which the decision maker considered  $(a, b)$  followed by  $(a, c)$  and finally  $(b, c)$ . Indeed, to rationalize this choice rule by GPC it must be that the comparison  $(a, b)$  is made before either  $(a, c)$  or  $(b, c)$ . There is no other way to have  $b$  ranked strictly below  $c$  in the choice set  $\{a, b, c\}$ . Beyond that, however, there are no further inferences to be made about the order of comparison. The choice data is consistent with  $(a, c)$  being considered both before  $(b, c)$  (as in example 1) and after. For the latter, consider the following  $\mathbb{P}$  and  $\pi$  which also rationalizes the choice rule.

$P_0$	$P_1$	$P_2$	$P_3$
	$(a, b)$	$(b, c)$	$(a, c)$

$\pi(\cdot, A)$	$A = \{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
$P_0$	0	2/3	2/3	0
$P_1$	1	0	0	4/5
$P_2$	0	0	1/3	0
$P_3$	0	1/3	0	1/5

This leads to an obvious question: is it possible to infer from choice data rationalizable by GPC that some pairwise comparison must be considered before some other?



Indeed, it is. Begin by appropriately defining these pairs of comparisons which must always be ordered in a specific way to rationalize some given choice data.

**Definition 5.** For a choice rule  $p$ , rationalizable by GPC,  $(a, b)$  is **revealed compared before**  $(x, y)$  if for all  $\mathbb{P}, \pi$  such that  $p = p^{\mathbb{P}, \pi}$ ,

$$(a, b) \in P_i \text{ and } (x, y) \in P_j \Rightarrow i < j.$$

In choice data rationalizable by GPC, the order of pairwise comparison is revealed by violations of the order independence property. The intuition is simple. Suppose  $b$  is not chosen with lowest probability from some choice set  $A$ . If adding  $a$  to this set makes  $b$  receive the lowest choice probability, then it must be that the comparison  $(a, b)$  precedes any comparison between any other pair of alternatives in  $A \cup \{a\}$ . The following definition captures this binary relation over pairwise comparisons.

**Definition 6.** For a choice rule  $p$ , rationalizable by GPC, with  $P$  the stochastically revealed preference relation and  $(a, b), (x, y) \in P$ , let

$$\begin{aligned} (a, b) \triangleright_p (x, y) & \quad \text{if} & & \quad b \text{ is } \mathbf{not} \text{ choice ranked last in } A \\ & \quad \text{and} & & \quad b \text{ is choice ranked last in } A \cup \{a\} \end{aligned}$$

for some  $A \in \mathcal{X}$  with  $\{x, y\} \subseteq A \cup \{a\}$ ,  $b \in A$  and  $b \neq y$ .

Let  $\bar{\triangleright}_p$  be the transitive closure of  $\triangleright_p$ . The following theorem then fully characterizes what choice data can reveal about the agent's sequence of pairwise comparisons.

**Theorem 5.** Given a choice rule without ties  $p$ , rationalizable by GPC,  $(a, b)$  is revealed compared before  $(x, y)$  if and only if  $(a, b) \bar{\triangleright}_p (x, y)$ .

## 4 Menu Dependence

### 4.1 Stochastic Transitivity

Notions of choice consistency across menus offer a useful way to classify both theories of choice as well as choice data (see Reiskamp, Busemeyer and Mellers (2006)). *Strong Stochastic Transitivity* is one such notion and a strict one, requiring  $\forall a, b, c \in X$ ,  $p(a, \{a, b\}) \geq 1/2, p(b, \{b, c\}) \geq 1/2 \Rightarrow p(a, \{a, c\}) \geq \max\{p(a, \{a, b\}), p(b, \{b, c\})\}$ . This is satisfied by the commonly used Luce's model or the multinomial logit, which meets the stricter requirement, called *Luce's IIA*,  $p(a, A)/p(b, A) = p(a, B)/p(b, B)$ . GPCRs can violate strong stochastic transitivity (and therefore Luce's IIA too), as can be seen in example 1. The reason is simple. The probability of choosing the superior alternative in pairwise choices does not depend on how far apart the alternatives are in the underlying preference order. Instead it depends on how readily the pairwise comparison is made. So even though  $c$  is ranked below  $b$ , the comparison  $(a, b)$  is obvious and  $a$  is selected with certainty from  $\{a, b\}$ , while  $(a, c)$  is hard and yields a lower probability of selecting  $a$  from  $\{a, c\}$ .

GPCRs can violate a weaker consistency notion satisfied by the models in Tversky (1972a) and Natenzon (2018), called *Moderate Stochastic Transitivity*,  $\forall a, b, c \in X$ ,  $p(a, \{a, b\}) \geq 1/2, p(b, \{b, c\}) \geq 1/2 \Rightarrow p(a, \{a, c\}) \geq \min\{p(a, \{a, b\}), p(b, \{b, c\})\}$ . To see this, consider the following example.

**Example 2.**  $a \succ b \succ c$  with

$P_0$	$P_1$	$P_2$	$P_3$
	$(b, c)$	$(a, b)$	$(a, c)$

and

$\pi(\cdot, A)$	$A = \{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
$P_0$	1/3	2/3	0	0
$P_1$	0	0	1	1/3
$P_2$	2/3	0	0	2/3
$P_3$	0	1/3	0	0

The resulting choice probabilities are as follows.

$p^{\mathbb{P}, \pi}(\cdot, A)$	$A = \{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
$a$	5/6	2/3	0	5/6
$b$	1/6	0	1	1/6
$c$	0	1/3	0	0

In particular,  $p^{\mathbb{P}, \pi}(a, \{a, c\}) = 2/3 < \min\{p^{\mathbb{P}, \pi}(a, \{a, b\}), p^{\mathbb{P}, \pi}(b, \{b, c\})\} = 5/6$ .

Recall that the choice probability of the superior alternative in a pairwise comparison depends on the probability with which the decision-maker makes that comparison. Here while considering the larger set  $\{a, b, c\}$  the decision-maker is able to indirectly (by making the comparison  $(b, c)$  followed by  $(a, b)$ ) realize that  $c$  is inferior to  $a$  with a higher probability than if she were assessing the set  $\{a, c\}$  directly.<sup>10</sup> To see this more concretely suppose the choice in example 2 is over bundles of goods of the form  $(x, y, z)$ ;  $x$  units of good  $u$ ,  $y$  of  $v$  and  $z$  of  $w$ . The decision-maker's underlying preference corresponds to a symmetric Cobb-Douglas function, but she finds it easier to consider tradeoffs involving two goods as compared to those with three or more. Suppose  $a = (4, 10, 4)$  and  $c = (5, 13, 2)$ . The decision-maker finds the relevant comparison facing the choice set  $\{a, c\}$  difficult. She needs to consider tradeoffs involving three goods. The menu  $\{a, b, c\}$  with  $b = (3, 13, 4)$ , makes it easier for her. The comparison  $(b, c)$  is immediate, trading off a gain of 2 of  $w$  against a loss

<sup>10</sup>This is similar in spirit to the common occurrence that an otherwise opaque result seems obvious when broken down into an appropriate sequence of obvious lemmas.

of 2 of  $u$ , with a clearly more balanced bundle in  $b$ . The comparison  $(a, b)$  is next, again with a tradeoff involving only two goods. As a result, she is able to indirectly assess  $c$  as being inferior to  $a$  from the set  $\{a, b, c\}$  with a higher probability than if she considered the set  $\{a, c\}$  directly.

GPCRs, however, always satisfy *Weak Stochastic Transitivity*, which requires  $\forall a, b, c \in X, p(a, \{a, b\}) \geq 1/2, p(b, \{b, c\}) \geq 1/2 \Rightarrow p(a, \{a, c\}) \geq 1/2$ . Indeed this is due to the following more general result.

**Observation 2.** *sWARP implies weak stochastic transitivity.*

## 4.2 Similarity Effect

Debreu (1960), through a thought experiment, pointed out that Luce’s IIA is unable to accommodate alternatives that are very similar. The principle that Luce’s IIA runs counter to in the thought experiment and for which there exists considerable empirical support is called the *similarity effect*, described in Tversky (1972b) as follows:

*“The addition of an alternative to an offered set ‘hurts’ alternatives that are similar to the added alternative more than those that are dissimilar to it.”*

The GPC choice procedure accommodates the similarity effect in an obvious way. Alternatives  $b$  and  $c$  being similar has a natural representation within the choice procedure: the decision-maker makes the pairwise comparisons between  $b$  and any other alternative  $x \in X \setminus \{b, c\}$  in an identical way to her comparison of  $c$  and  $x$ , in that both comparisons belong to the same cell of her ordered partition and  $b$  is preferred to  $x$  if and only if  $c$  is preferred to  $x$ . Finally, the comparison between  $b$  and  $c$  itself must happen sufficiently late in the sequence, so that with a high probability the decision maker treats them interchangeably.

Formally,  $b$  and  $c$  being similar in a GPC choice procedure,  $(\mathbb{P}, \pi)$ , is modelled as

follows. Suppose  $\{(b, c), (c, b)\} \cap P_k \neq \emptyset$ . Then,

$$\begin{aligned}
(i) & \forall x \in X \setminus \{b, c\}, \quad (b, x) \in P_i \Leftrightarrow (c, x) \in P_i \quad (x, b) \in P_i \Leftrightarrow (x, c) \in P_i \\
(ii) & \forall x \in X \setminus \{b, c\}, \quad \{(b, x), (x, b)\} \cap P_j \neq \emptyset \quad \Rightarrow j < k \\
(iii) & \forall A \supseteq \{b, c\}, \quad \sum_{i < k} \pi(P_i, A) > 1 - \nu \quad \text{for sufficiently small } \nu.
\end{aligned}$$

While (i) and (ii) capture ordinal features of similar alternatives about how they relate to any third alternative, (iii) is about how they relate to each other cardinally. In particular  $\nu$  is an upper bound to how far apart their choice probabilities can be;  $\nu \geq |p^{\mathbb{P}, \pi}(b, A) - p^{\mathbb{P}, \pi}(c, A)|$  for any  $\{b, c\} \subseteq A \in \mathcal{X}$ . The following proposition then describes how the GPC procedure accommodates the similarity effect.

**Proposition 1.** *Suppose  $b$  is similar to  $c$  and dissimilar to  $a$  with  $\{a, b\} \subseteq A$  and*

*$\sum_{i \leq I} |\pi(P_i, A) - \pi(P_i, A \cup \{c\})| < \epsilon$ . For small enough  $\epsilon$  and  $\nu$ ,*

*(i) if  $p^{\mathbb{P}, \pi}(a, A) < p^{\mathbb{P}, \pi}(b, A)$  then*

$$p^{\mathbb{P}, \pi}(a, A) - p^{\mathbb{P}, \pi}(a, A \cup \{c\}) < p^{\mathbb{P}, \pi}(b, A) - p^{\mathbb{P}, \pi}(b, A \cup \{c\}), \quad (3)$$

*(ii) if  $p^{\mathbb{P}, \pi}(a, A) > p^{\mathbb{P}, \pi}(b, A) > 0$  then*

$$\frac{p^{\mathbb{P}, \pi}(a, A) - p^{\mathbb{P}, \pi}(a, A \cup \{c\})}{p^{\mathbb{P}, \pi}(a, A)} < \frac{p^{\mathbb{P}, \pi}(b, A) - p^{\mathbb{P}, \pi}(b, A \cup \{c\})}{p^{\mathbb{P}, \pi}(b, A)}. \quad (4)$$

Note that adding  $c$  to the choice set  $A$  leaves the set and sequence of relevant comparisons essentially unchanged.<sup>11</sup> It is reasonable then to assume that the stopping rule should not change by much between  $A$  and  $A \cup \{c\}$ . Proposition 1 states that beyond this small (potential) effect of stopping rule change,  $\epsilon$ , adding  $c$ , hurts the similar alternative  $b$  more than the dissimilar one,  $a$ . Specifically, if  $b$  was chosen

---

<sup>11</sup>Any comparison not involving either  $b$  or  $c$  is relevant in  $A$  if and only if it is relevant in  $A \cup \{c\}$ . Further, any relevant comparison in  $A$  involving  $b$  and some other alternative is now accompanied by a duplicate comparison with  $c$  replacing  $b$ , in the same cell of the ordered partition. Finally, even the new comparison of  $(b, c)$  in  $A \cup \{c\}$ , if relevant, is reached with a negligibly small probability  $\nu$ .

with higher probability than  $a$  in  $A$ , then  $b$  loses strictly more probability than  $a$  upon the addition of  $c$ , for sufficiently small change in the stopping rule (inequality 3). If instead,  $a$  was chosen with a higher probability than  $b$  in  $A$ , both positive, then both these alternatives lose essentially the same amount of probability upon the addition of  $c$ . Nevertheless, the similarity effect obtains with  $b$  losing a larger proportion of its probability compared to  $a$  (inequality 4), improving the odds ratio in favour of  $a$ .

The rationale behind the result is simple. The only effect of adding  $c$  to  $A$ , is  $c$  joining all survivor sets  $M_i^{\mathbb{P}}$  that contain  $b$  (except the one following the comparison  $(b, c)$ ), leaving all else unchanged. This makes  $c$  take away more probability from  $b$  than  $a$ , either directly as in part (i) where the set of survivor sets containing  $a$  is a strict subset of those containing  $b$ , or proportionally as in part (ii) where the survivor sets containing  $a$  and those containing  $b$ , that are affected by  $c$  are the same but  $a$  is chosen with the higher probability in  $A$ .

### 4.3 Violation of Regularity Through Fatigue

The regularity property, which requires  $p(a, A) \geq p(a, B)$  for all  $a \in A \subseteq B$  and satisfied by random utility models, is often violated in choice data (see review in Resikamp et al. (2006)). The GPC procedure captures how fatigue may lead to regularity violations.

Consider the choice environment described in figure 1. The decision-maker chooses from among alternatives defined by a pair of attributes. She prefers more to less of each attribute. Suppose her underlying preference is  $a \succ b \succ c$ . The fact that alternative  $c$  is clearly dominated by  $b$  but not by  $a$  makes it natural then that the decision-maker makes the comparison  $(b, c)$  before any other.

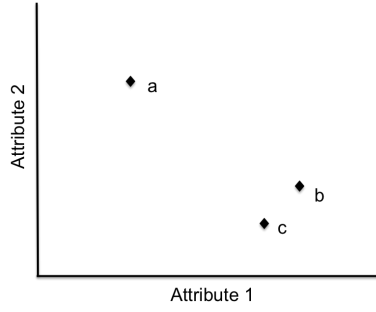


Figure 1: Violation of Regularity

Consider the following order of comparison

$P_0$	$P_1$	$P_2$
	$(b, c)$	$(a, b)$
		$(a, c)$

The resulting choice violates regularity with  $p(b, \{a, b, c\}) > p(b, \{a, b\})$  if  $\pi(P_0, \{a, b, c\})/3 + \pi(P_1, \{a, b, c\})/2 > \pi(P_0, \{a, b\})/2$ . Given the obvious nature of the comparison  $(b, c)$  it is reasonable to expect  $\pi(P_0, \{a, b, c\})$  to be negligibly small, if not 0. Therefore the violation occurs as long as having to make an additional comparison increases the chance of giving up before completing a tougher comparison.

The phenomenon of decision fatigue, in which making multiple decisions negatively affects the quality of subsequent decisions, is now well documented (see for instance Levav et al. (2010) and Hirshleifer et al. (2019)). The example above shows that even in a single decision problem fatigue may worsen choices. Indeed, even adding alternatives that are relatively easily found to be inferior may nevertheless make it less probable that the decision-maker completes the more difficult comparisons.

This phenomenon is distinct from the attraction effect. In the latter the increased probability of choosing  $b$  upon adding  $c$  does not depend on the preference between  $a$  and  $b$ . By contrast, in the argument above it matters that  $a$  is preferred to  $b$ . Here the earlier comparison,  $(b, c)$ , helps the inferior alternative  $b$  in the later one  $(a, b)$ .

## 5 Discussion

### 5.1 Sequence of Pairwise Comparisons

An essential ingredient of the GPC procedure is that choice involves a sequence of pairwise comparisons. Ravid and Steverson (2018) study a model of choice which also contains this feature but is otherwise very different in both structure and implication. In their model labeled Focus, Then Compare (FTC), an agent focuses on an available alternative (with equal odds) and sequentially compares it in a pairwise manner to all other alternatives in a random order. The focal alternative is selected if all comparisons are favourable. Otherwise, a new alternative is selected (with replacement) as focal, again with equal odds, and the procedure repeated.

The most important difference with GPC is that the result of each binary comparison in FTC is random. Requiring these comparisons to reflect a stable rational preference (as in GPC) would reduce the model to deterministic rational choice. So, to the extent that FTC can rationalize context-dependent choice or even non-degenerate stochastic choice, it must depart from rational preferences. Further, in contrast to GPC, the sequentiality of pairwise comparisons (for a given focal option) plays no important procedural role in FTC. Since what matters is whether all the pairwise comparisons are favourable or not, changing the order of these comparisons while holding their outcomes constant would lead to the same decision.

The FTC procedure shows how the random nature of what catches an agent's attention can interact with random (pairwise) choice mistakes to generate context dependent choice. By contrast, in GPC, context dependent choice is the result of the differential ease of making these pairwise comparisons (hence the sequential structure), which interacts with the random number of comparisons the agent is able to make. While not directly addressed in the paper, to the best of my understanding, FTC cannot accommodate the similarity effect.

The simple premise behind choosing an alternative in the GPC procedure is that



it is not found inferior in any relevant pairwise comparison. Based on this alone the agent has no reason to further discriminate among surviving alternatives. This is the logic behind the agent randomizing uniformly across these alternatives. There exist more involved variants of the procedure where the agent uses additional information to randomize non-uniformly across surviving alternatives.<sup>12</sup> It is beyond the scope of this study to consider and compare such variants. Hopefully, subsequent theoretical and empirical work such as Reutskaja et al. (2011) will identify the interesting variants.

## 5.2 Menu (In)Dependence

The menu independence of the order of relevant comparisons follows from the assumption that easier comparisons precede more difficult ones. Consider example 1 again. The agent finds the comparison  $(a, b)$  easier to make than  $(a, c)$ . Since this relative ease depends directly on  $a, b$  and  $c$ , adding another lottery to the menu should have no effect on it. Note, however, that adding a new lottery could make one or both of these comparisons irrelevant.<sup>13</sup> But if  $(a, b)$  and  $(a, c)$  continue to be relevant upon adding another lottery, then the comparison  $(a, b)$  must continue to precede  $(b, c)$ .

While the menu dependent choice described in Section 4 is not an example of framing effects, the GPC procedure can accommodate the latter in a natural way. Framing an alternative differently while leaving its payoff relevant features the same does not change the preference ordering but typically changes the order of relevant pairwise comparisons, thereby affecting choice.

The key driver of context dependent choice in the GPC procedure is the menu-dependent nature of the set of relevant pairwise comparisons. This interacts with the menu-independent order of relevant comparisons, to generate a highly menu-

---

<sup>12</sup>For instance, the agent could keep track of the number of successful comparisons for each surviving alternative. Indeed, different successful comparisons could be weighted differently as well, depending on the identity of the dominated alternative, and so on.

<sup>13</sup>For instance, if lottery  $d$ , which pays \$10,000 with probability 1, were added to the menu, then perhaps the comparisons  $(d, a)$ ,  $(d, b)$  and  $(d, c)$  would precede all others, making the comparisons  $(a, b)$  and  $(a, c)$  irrelevant for the choice set  $\{a, b, c, d\}$ .

dependent sequence of actual comparisons the decision-maker makes.

### 5.3 Random Attention Models

Recent work has focused on a different source of bounded rationality.<sup>14</sup> Only a strict subset of all available alternatives may catch the decision-maker’s attention. Despite a rational preference, her chosen alternative could then be worse than an available option that did not catch her attention. The stochastic nature of what catches the agent’s attention for a given choice set, leads to stochastic choice behaviour.

As for choice data it can rationalize, none of the random attention models nest GPC and vice versa. For instance the model in Manzini and Mariotti (2014) must satisfy regularity (unlike a GPCR) but allows violations of sWARP. The following example is of a GPCR that violates the acyclicity condition that characterizes the most general random attention model, studied in Cattaneo et al. (2018).

**Example 3. (Violation of RAM Acyclicity)**

The ordered partition of the underlying preference  $a \succ b \succ c \succ d$  is

$P_1$	$P_2$	$P_3$	$P_4$	$P_5$
$(c, d)$	$(a, b)$	$(a, d)$	$(a, c)$	$(b, c)$
				$(b, d)$

The stopping function is the following.

$\pi(\cdot, A)$	$A = \{a, c\}$	$\{a, b, c\}$	$\{a, b, d\}$	$\{a, b, c, d\}$	...
$P_1$	0	0	0	1/5	...
$P_2$	0	1/3	1/3	2/5	...
$P_3$	0	0	2/3	0	...
$P_4$	1	2/3	0	2/5	...
$P_5$	0	0	0	0	...

---

<sup>14</sup>See Manzini and Mariotti (2014), Brady and Rehbeck (2016) and Cattaneo et al. (2018).

This results in  $p^{\mathbb{P},\pi}(c, \{a, c\}) = 0$  and  $p^{\mathbb{P},\pi}(c, \{a, b, c\}) = 1/6$ , which in Cattaneo et al. (2018) means that  $c$  is revealed preferred to  $b$ . But with  $p^{\mathbb{P},\pi}(b, \{a, b, d\}) = 0$  and  $p^{\mathbb{P},\pi}(b, \{a, b, c, d\}) = 1/15$ ,  $b$  is revealed preferred to  $c$ , violating acyclicity.

More importantly, the random attention model and GPC constitute fundamentally different choice procedures. This often leads to the same choice behaviour despite vastly different underlying preferences, making inference about such preferences while ignoring the specific source of bounded rationality extremely tenuous.<sup>15</sup> Take for instance example 1, in which the underlying preference in the GPC procedure was  $a \succ b \succ c$ . The choice data in the example is rationalizable by RAM, but the only inferences about preference under RAM would be  $c \succ a$  and  $c \succ b$ .

## 5.4 Luce's Model and Related Work

In Luce's model, if the underlying preference is a strict order, as in this study, then the alternatives can be assigned values,  $u(x) \in \mathbb{R}_{++}, \forall x \in \mathcal{X}$ , such that  $u(x) \neq u(y)$  if  $x \neq y$  and  $p(x, A) = u(x)/(\sum_{y \in A} u(y))$ . All such choice rules are rationalizable by GPC.

Consider the following construction which rationalizes any choice rule where given the function  $u(\cdot)$  above,  $p(x, A) > p(y, A)$  if and only if  $u(x) > u(y)$ .<sup>16</sup> Luce's model is a particular case of this more general class of choice rules. Order the  $n$  alternatives in  $X$  as  $\{x_i\}_{i=1}^n$  where  $i < j \Leftrightarrow u(x_i) > u(x_j)$ . The ordered partition is the following:

$$P_i = \{(x_j, x_{n-i+1}) | j < n - i + 1\} \quad \forall i \leq n.$$

In words, the first element of the ordered partition contains all pairwise comparisons in which the alternative with the lowest value under  $u$  is the inferior alternative. The second element contains all pairwise comparisons in which the alternative with the

<sup>15</sup>Masatlioglu, Nakajima and Ozbay (2012) and Dutta and Horan (2015) make a similar argument against the model-free approach suggested in Bernheim and Rangel (2009) for deterministic choice.

<sup>16</sup>This includes any choice rule without ties that has an Additive Perturbed Utility representation, as in Fudenberg, Iijima and Strzalecki(2015).

second lowest value under  $u$  is the inferior alternative, and so on. Let  $\pi$  be such that  $\pi(P_i, \cdot) > 0$  for all  $i \leq n$ . It then follows that

$$p^{\mathbb{P}, \pi}(x, A) > p^{\mathbb{P}, \pi}(y, A) \Leftrightarrow u(x) > u(y), \quad \forall A \in \mathcal{X}, \{x, y\} \subseteq A.$$

Matching the exact choice probabilities then follows from theorem 2. Also by theorem 2, the  $k(A)$  lowest ranked alternatives in choice set  $A$  can be assigned 0 probability. This formulation, in a natural way, allows for zero probabilities in choice, while staying consistent with the order independence axiom.<sup>17</sup>

Finally, to match the choice probabilities to the Luce rule exactly, the following stopping rule  $\pi$  (along with the ordered partition above) is sufficient. Fix  $A \in \mathcal{X}$ . Let  $x_i$  and  $x_k$  be two alternatives in  $A$  with adjacent choice ranks and  $x_k$  the worse of the two, where  $i$  and  $k$  correspond to the order described in the previous paragraph. Set

$$\pi(P_{n-k+1}, A) = |\{x_m \in A | m \leq i\}| \frac{u(x_i) - u(x_k)}{\sum_{x_j \in A} u(x_j)}$$

and  $\pi(P_{n-q+1}, A) = |A|u(x_q)/(\sum_{x_j \in A} u(x_j))$  where  $x_q$  is the worst choice ranked option in  $A$ .

## A Appendix

**Lemma 1.** *For any GPCR,  $p^{\mathbb{P}, \pi}$ , there exists an SPCR,  $p^{\mathbb{P}', \pi'}$ , such that  $p^{\mathbb{P}, \pi} = p^{\mathbb{P}', \pi'}$ .*

*Proof.* Consider a GPCR,  $p^{\mathbb{P}^n, \pi^n}$  and let  $P_k$  be the first cell in  $\mathbb{P}^n$  that is not a singleton. Since  $P$  is a strict rational preference there must exist some  $(x, y) \in P_k$  such that  $(y, c) \notin P_k$  for any  $c \in X$ . Let  $(a, b) \in P_k$  satisfy this condition. Define a new ordered partition  $\mathbb{P}_{n+1} = \{P'_j\}$  and stopping function  $\pi_{n+1}$  in the following way:  $P'_i = P_i$  for all  $i < k$ ,  $P'_k = (a, b)$ ,  $P'_{k+1} = P_k \setminus \{(a, b)\}$ ,  $P'_{i+1} = P_i$  for all  $i > k$ ,  $\pi_{n+1}(P'_i, \cdot) = \pi_n(P_i, \cdot)$

---

<sup>17</sup>See Echenique and Saito (2018), Ahumada and Ulku (2017) and Horan (2018) for work on extending the Luce model to better incorporate choice with 0 probabilities.

for all  $i < k$ ,  $\pi_{n+1}(P'_k, \cdot) = 0$  and  $\pi_{n+1}(P'_{i+1}, \cdot) = \pi_n(P_i, \cdot)$  for all  $i \geq k$ . It is easy to confirm that  $p^{\mathbb{P}_n, \pi_n} = p^{\mathbb{P}_{n+1}, \pi_{n+1}}$ . So setting  $\mathbb{P}_1 = \mathbb{P}$  and  $\pi_1 = \pi$ , generates a finite sequence  $\{p^{\mathbb{P}_n, \pi_n}\}_{n=1}^m$  using the construction above, such that  $p^{\mathbb{P}_m, \pi_m}$  is an SPCR. Setting  $\mathbb{P}_m = \mathbb{P}'$  and  $\pi_m = \pi'$  concludes the proof.  $\square$

**Lemma 2.** *For any GPCR,  $p^{\mathbb{P}, \pi}$ , there exists a GPCR without ties,  $p^{\mathbb{P}', \pi'}$ , such that for all  $A \in \mathcal{X}$  and  $a, b \in X$ ,*

$$p^{\mathbb{P}', \pi'}(a, A) > p^{\mathbb{P}', \pi'}(b, A) \Rightarrow p^{\mathbb{P}, \pi}(a, A) \geq p^{\mathbb{P}, \pi}(b, A).$$

*Proof.* Fix a GPCR  $p^{\mathbb{P}, \pi}$ . Then by lemma 1 there exists an SPCR, say  $p^{\bar{\mathbb{P}}, \bar{\pi}}$ , such that  $p^{\mathbb{P}, \pi} = p^{\bar{\mathbb{P}}, \bar{\pi}}$ . Set  $\mathbb{P}' = \bar{\mathbb{P}}$ . Pick any stopping function  $\pi'$  on  $\mathbb{P}'$  such that  $\pi'(P_i, \cdot) > 0$  for all  $P_i \in \mathbb{P}'$ . Then  $p^{\mathbb{P}', \pi'}$  is a GPCR without ties. Also,

$$p^{\mathbb{P}', \pi'}(a, A) > p^{\mathbb{P}', \pi'}(b, A) \Rightarrow p^{\bar{\mathbb{P}}, \bar{\pi}}(a, A) \geq p^{\bar{\mathbb{P}}, \bar{\pi}}(b, A) \Leftrightarrow p^{\mathbb{P}, \pi}(a, A) \geq p^{\mathbb{P}, \pi}(b, A).$$

The first implication is because  $\mathbb{P}' = \bar{\mathbb{P}}$  and Theorem 1.  $\square$

**Proof of Observation 1.** It is sufficient to show that given a GPCR  $p^{\mathbb{P}, \pi}$ , it follows that  $p^{\mathbb{P}, \pi} = p^{\mathbb{P}, \pi'}$ , where  $\pi'$  is an exact stopping function. Let  $Z(A) = \{i \leq I \mid M_i^{\mathbb{P}}(A) \neq M_{i-1}^{\mathbb{P}}(A)\} \cup \{0\}$ . For any  $j \in Z(A)$ , let  $k(j)$  be the smallest number in the set  $Z(A)$  that is larger than  $j$ . If  $j$  is the highest number in  $Z(A)$  then let  $k(j) = I + 1$ . Set  $\pi'(P_j, A) = \sum_{j \leq i \leq k(j)-1} \pi(P_i, A)$  if  $j \in Z(A)$ . Otherwise set  $\pi'(P_j, A) = 0$ .  $\pi'$  is

therefore an exact stopping function. Now fix some  $A \in \mathcal{X}$  and  $a \in A$ . Observe that

$$\begin{aligned}
p^{\mathbb{P}, \pi'}(a, A) &= \sum_{\{j \in Z(A) \mid a \in M_j^{\mathbb{P}}(A)\}} \frac{\pi'(P_j, A)}{|M_j^{\mathbb{P}}(A)|} \\
&= \sum_{\{j \in Z(A) \mid a \in M_j^{\mathbb{P}}(A)\}} \frac{\sum_{j \leq i \leq k(j)-1} \pi(P_i, A)}{|M_j^{\mathbb{P}}(A)|} \\
&= \sum_{\{j \in Z(A) \mid a \in M_j^{\mathbb{P}}(A)\}} \sum_{j \leq i \leq k(j)-1} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} \\
&= \sum_{\{i \mid a \in M_i^{\mathbb{P}}(A)\}} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} = p^{\mathbb{P}, \pi}(a, A)
\end{aligned}$$

□

**Proof of Theorem 2.** Given  $\mathbb{P}$  and some  $A \in \mathcal{X}$ , define the sequence of sets,  $\{E_i(A)\}_{i=1}^I$  where  $E_i(A) = \{a \in A \mid a \notin M_i^{\mathbb{P}}(A) \text{ and } a \in M_{i-1}^{\mathbb{P}}(A)\}$ .  $E_i(A)$  contains all alternatives in  $A$  that are eliminated by the GPC choice procedure at the  $i$ 'th cell of  $\mathbb{P}$ . It follows from equation 2 that if  $a$  and  $b$  both belong to  $E_i(A)$  then  $p^{\mathbb{P}, \pi}(a, A) = p^{\mathbb{P}, \pi}(b, A)$ . By the premise of the theorem, it follows that  $p(a, A) = p(b, A)$ . Let  $p(\alpha_j, A)$  denote the choice probability under  $p$  of an element (if any) in  $E_j(A)$ . If  $p^{\mathbb{P}, \pi}(a, A) < p^{\mathbb{P}, \pi}(b, A)$  then again, from equation 2, it follows that  $a \in E_j(A)$  and  $b \in E_k(A)$  with  $j < k$ . Further by the premise of the theorem,  $p(a, A) \leq p(b, B)$ .

In what follows,  $\pi'$  is selected so that  $\pi'(P_i, A) \geq 0$  for all  $0 \leq i \leq I$ . Let  $\underline{j}$  be the smallest number  $i$  for which  $E_i(A)$  is nonempty. Set  $\pi'$  such that  $\sum_{i=0}^{\underline{j}-1} \pi'(P_i, A) = |A|p(\alpha_{\underline{j}}, A)$ . Subsequently, for any  $j$  and  $k$  with  $j < k$  such that  $E_j(A)$  and  $E_k(A)$  are non-empty and  $E_q(A) = \emptyset$  for all  $j < q < k$ , set  $\pi'$  such that

$$\sum_{i=j}^{k-1} \pi'(P_i, A) = |M_{k-1}^{\mathbb{P}}(A)| [p(\alpha_k, A) - p(\alpha_j, A)]. \quad (5)$$

Finally if  $j$  is the largest number  $i$  for which  $E_i(A)$  is non-empty then set  $\pi'(P_j, A) =$

$p(\alpha_j, A) - p(\alpha_h, A)$ , where  $h$  is the second largest number  $i$  for which  $E_i(A)$  is non-empty.

$\pi'$ , so defined, is a stopping function. Indeed, the selections ensure that  $\pi'(P_i, A) \geq 0$  for all  $0 \leq i \leq I$ . Also if  $j$  is the largest number  $i$  for which  $E_i(A)$  is non-empty then it must be that  $j = \tilde{I}^{\mathbb{P}}(A)$ . So  $\pi'(P_{\tilde{I}^{\mathbb{P}}(A)}, A) > 0$ . Let  $Z(A) = \{i | E_i(A) > 0 \text{ and } E_k(A) > 0 \text{ for some } k > i\}$ . For  $j \in Z(A)$  let  $j'$  be the next highest element in  $Z(A)$  (formally,  $j' > j$  and  $\nexists k \in Z(A)$  with  $j < k < j'$ ). Then,

$$\begin{aligned} \sum_{i=0}^I \pi'(P_i, A) &= |A|p(\alpha_{\underline{j}}, A) + \sum_{j \in Z(A)} |M_{j'-1}^{\mathbb{P}}(A)| [p(\alpha_{j'}, A) - p(\alpha_j, A)] + p(\alpha_{\tilde{I}^{\mathbb{P}}(A)}, A) \\ &= \sum_j |E_j(A)|p(\alpha_j, A) + p(\alpha_{\tilde{I}^{\mathbb{P}}(A)}, A) = \sum_{a \in A} p(a, A) = 1. \end{aligned}$$

Since  $p$  satisfies unique best,  $|A| - \sum_j |E_j(A)| = 1$ . This is what ensures that the third equality above holds. It is now straightforward to verify that  $p^{\mathbb{P}, \pi'}(a, A) = p(\alpha_j, A)$  where  $a \in E_j(A)$ .  $\square$

**Lemma 3.**  $M_i^{\mathbb{P}}(A) = B \Rightarrow M_i^{\mathbb{P}}(B) = B$

*Proof.*  $M_i^{\mathbb{P}}(A) = B$  implies that for all  $a, b \in B$ ,  $(a, b) \notin P_j$  for all  $j \leq i$ . This in turn implies that  $M_j^{\mathbb{P}}(B) = B$  for all  $j \leq i$ .  $\square$

**Lemma 4.** If  $a \in M_{\tilde{I}^{\mathbb{P}}(A)}^{\mathbb{P}}(A)$  then  $(a, c) \in P$  for all  $c \in A \setminus \{a\}$ .

*Proof.* Suppose by contradiction, there exists  $b \in A$  such that  $(b, a) \in P$ . Since  $P$  is a strict rational preference there must be a  $P$ -maximal element in  $A$ , say  $d \in A$ . If  $(b, a) \in P$  then  $d \neq a$ . Since by definition  $M_{\tilde{I}^{\mathbb{P}}(A)}^{\mathbb{P}}(A)$  is a singleton,  $d \notin M_{\tilde{I}^{\mathbb{P}}(A)}^{\mathbb{P}}(A) \Rightarrow (c, d) \in P$  for some  $c \in A$ . This contradicts  $d$  being  $P$ -maximal in  $A$ .  $\square$

**Proof of Theorem 3. Necessity:** It is sufficient to show that if  $p^{\mathbb{P}, \pi}$  is an SCR without ties then it satisfies *sWARP*, *s-reducibility* and *ITR*. It turns out that any  $p^{\mathbb{P}, \pi}$  (not just those without ties) satisfies *sWARP*.

**sWARP:** Suppose under the SCR  $p^{\mathbb{P},\pi}$ ,  $a$  is stochastically revealed preferred to  $b$ . So for some  $A \in \mathcal{X}$  with  $\{a, b\} \subseteq A$ ,

$$\begin{aligned}
& p^{\mathbb{P},\pi}(a, A) \geq p^{\mathbb{P},\pi}(c, A), & \forall c \in A \\
\Rightarrow & \sum_{\{i|a \in M_i^{\mathbb{P}}(A)\}} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} \geq \sum_{\{i|c \in M_i^{\mathbb{P}}(A)\}} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|}, & \forall c \in A \\
\Rightarrow & \{i|a \in M_i^{\mathbb{P}}(A)\} \supseteq \{i|c \in M_i^{\mathbb{P}}(A)\}, & \forall c \in A \\
\Rightarrow & \{i|a \in M_i^{\mathbb{P}}(A)\} \supset \{i|c \in M_i^{\mathbb{P}}(A)\}, & \forall c \in A \setminus \{a\}
\end{aligned}$$

The final implication obtains because  $M_{\mathbb{P}(A)}^{\mathbb{P}}(A)$  is a singleton by definition, and so it must contain  $a$ . This means that  $(a, c) \in P, \forall c \in A \setminus \{a\}$  by lemma 4.

Now, suppose by contradiction there exists  $B \in \mathcal{X}$  such that  $a, b \in B$  and  $p^{\mathbb{P},\pi}(b, B) \geq p^{\mathbb{P},\pi}(c, B)$  for all  $c \in B$ . Then exactly by the argument above it must be that  $M_{\mathbb{P}(B)}^{\mathbb{P}}(B) = \{b\}$ . Again by lemma 4 this implies that  $(b, a) \in P$ , a contradiction.

**s-reducibility:** By lemma 1,  $p^{\mathbb{P},\pi}$  can be assumed to be an SPCR without loss of generality. Fix a collection of sets  $\mathcal{B} \subseteq \tilde{\mathcal{X}}$ . Given  $\mathbb{P}$ , let  $P_k$  be the first cell to contain a pairwise comparison  $(a, b)$  such that  $\{a, b\} \subseteq B$  for some  $B \in \mathcal{B}$ . Formally, for all  $P_i$  with  $i < k$ ,  $(x, y) \in P_i \Rightarrow \{x, y\} \not\subseteq A, \forall A \in \mathcal{B}$  and  $(a, b) \in P_k$  is such that  $\{a, b\} \subseteq B$  for some  $B \in \mathcal{B}$ . Since  $p^{\mathbb{P},\pi}$  is an SPCR,  $(a, b)$  is the only element in  $P_k$ .

So if  $\{a, b\} \subseteq A \in \mathcal{B}$  then  $M_i^{\mathbb{P}}(A) = A$  for all  $i < k$ . Since  $(a, b) \in P_k$  it must be that  $M_k^{\mathbb{P}}(A) = M_{k-1}^{\mathbb{P}}(A) \setminus \{b\}$ . This implies  $\{i|b \in M_i^{\mathbb{P}}(A)\} \subset \{i|c \in M_i^{\mathbb{P}}(A)\}$  for all  $c \in A$ . Therefore since  $p^{\mathbb{P},\pi}$  is an SPCR without ties,  $p^{\mathbb{P},\pi}(c, A) > p^{\mathbb{P},\pi}(b, A)$ ,  $\forall c \in A \setminus \{b\}$ .

**ITR:** Suppose  $D \in A^T(p^{\mathbb{P},\pi})$  for some  $A \in \tilde{\mathcal{X}}$ . Then it must be that  $M_k^{\mathbb{P}}(A) = D$  for some  $k$ . Lemma 3 then ensures that  $M_k^{\mathbb{P}}(D) = D$ . This implies that  $M_j^{\mathbb{P}}(A) = M_j^{\mathbb{P}}(D)$  for all  $j \geq k$ . So  $\{i|c \in M_i^{\mathbb{P}}(A)\} = \{i|c \in M_i^{\mathbb{P}}(D)\}$  for all  $c \in D$ . Therefore  $p^{\mathbb{P},\pi}(a, A) > p^{\mathbb{P},\pi}(b, A) \Leftrightarrow p^{\mathbb{P},\pi}(a, D) > p^{\mathbb{P},\pi}(b, D)$  for all  $a, b \in D$ . Observing that this is true for all  $A \in \tilde{\mathcal{X}}$  with  $D \in A^T(p^{\mathbb{P},\pi})$ , concludes the argument.



**Sufficiency:** Let  $p$  be a stochastic choice rule without ties that satisfies sWARP, ITR and s-reducibility. The proof is by construction and relies on defining a sequence of sets in a recursive manner. For any non-empty  $\mathcal{C}_i \subseteq \tilde{\mathcal{X}}$ , the subsequent set  $\mathcal{C}_{i+1}$  will be constructed along with  $P_i$ . Fix some non-empty  $\mathcal{C}_i \subseteq \tilde{\mathcal{X}}$ . Since  $p$  satisfies s-reducibility, there exists  $D \in \mathcal{C}_i$  with  $\{x, y\} \subset D$  such that if  $\{x, y\} \subset A \in \mathcal{C}_i$  then  $p(c, A) > p(y, A)$  for all  $c \in A \setminus \{y\}$ . There may be multiple pairs  $\{x, y\}$  that satisfy this condition. Simply pick one, say  $\{a_i, b_i\}$ . Set  $P_i = \{(a_i, b_i)\}$  and let  $\mathcal{C}_{i+1} = \mathcal{C}_i \setminus \mathcal{E}_i$  where  $\mathcal{E}_i = \{A \in \mathcal{C}_i \mid \{a_i, b_i\} \subseteq A\}$ .

Setting  $\mathcal{C}_1 = \tilde{\mathcal{X}}$  yields a sequence  $\{P_i\}_{i=1}^I$ . Let  $\mathbb{P} = \{P_i\}_{i=1}^I$ . Also let  $\pi$  be a stopping function on  $\mathbb{P}$  such that  $\pi(P_i, \cdot) > 0$  for all  $1 \leq i \leq I$ . It will now be shown that  $\mathbb{P}$  so defined is a partition of a strict rational preference. By construction, if  $(a, b) \in P_i$  then  $\{a, b\} \not\subseteq A$  for any  $A \in \mathcal{C}_j$  with  $j > i$ . Therefore  $(b, a) \notin P_j$  for  $j > i$ . This proves asymmetry. Next, by *s-reducibility*, as long as  $\mathcal{C}_i$  is non-empty,  $\mathcal{C}_{i+1} \subset \mathcal{C}_i$ . So for a given pair  $a, b$ , either  $\{a, b\} \in \mathcal{C}_j$  for some  $j$  such that  $A \notin \mathcal{C}_j$  for any  $A \supset \{a, b\}$  or  $(a, b) \in P_i$  or  $(b, a) \in P_i$  for some  $i < j$ . In the first case it must be that either  $(a, b) \in P_i$  or  $(b, a) \in P_i$  for some  $j \leq i \leq I$ . This proves completeness. Finally, to prove transitivity, note that by construction,  $(x_i, y_i) \in P_i \Rightarrow p(x_i, \{x_i, y_i\}) > p(y_i, \{x_i, y_i\})$ . Suppose by contradiction there exists a sequence  $\{x_i, y_i\}_{i=1}^n$  such that for each  $1 \leq i \leq n$ ,  $(x_i, y_i) \in P_{j(i)}$ ,  $y_i = x_{i+1}$ ,  $\forall i < n$  and  $x_1 = y_n$ . Since  $p$  satisfies sWARP there must be a unique most probable alternative under  $p$  in the set  $A = \cup_{i=1}^n \{x_i, y_i\}$ , say  $a$ . But then there must exist some  $1 \leq i \leq n$  for which  $a = y_i$ . This contradicts the assumption that  $p$  satisfies sWARP.

Therefore  $p^{\mathbb{P}, \pi}$  is a well defined SPCR. Further by construction  $p^{\mathbb{P}, \pi}$  is an SCR without ties. To see why, note first that since  $p^{\mathbb{P}, \pi}$  is an SPCR, for any  $A \in \tilde{\mathcal{X}}$  and  $a, b \in A$ , one of the two sets  $\{i \mid a \in M_i^{\mathbb{P}}(A)\}$  and  $\{i \mid b \in M_i^{\mathbb{P}}(A)\}$ , must be a strict subset of the other. Now since  $\pi(P_i, \cdot) > 0$  for all  $1 \leq i \leq I$  it must be that either  $p^{\mathbb{P}, \pi}(a, A) > p^{\mathbb{P}, \pi}(b, A)$  or  $p^{\mathbb{P}, \pi}(a, A) < p^{\mathbb{P}, \pi}(b, A)$ .

To complete the proof it is sufficient to show that for all  $A \in \mathcal{X}$  and  $a, b \in X$ ,

$p^{\mathbb{P},\pi}(a, A) > p^{\mathbb{P},\pi}(b, A) \Rightarrow p(a, A) > p(b, A)$ . Theorem 2 then guarantees the existence of a  $\pi'$  such that  $p^{\mathbb{P},\pi'} = p$ , since *sWARP* implies unique best.

Suppose by contradiction there exists an  $A \in \tilde{\mathcal{X}}$  such that the condition above does not hold. Consider the choice ranks defined by  $p^{\mathbb{P},\pi}$  and  $p$  on the alternatives in  $A$ . In particular, start with the lowest ranked alternative (smallest choice probability) according to each and if they are the same then move one rank up. If there exists  $x, y \in A$  such that  $p^{\mathbb{P},\pi}(x, A) > p^{\mathbb{P},\pi}(y, A)$  but  $p(y, A) > p(x, A)$ , then eventually this process must end with the  $m$ 'th ranked alternative according to  $p^{\mathbb{P},\pi}$  being different from that according to  $p$ , with  $m > 1$ , while all lower ranked alternatives are identical. Let  $B$  be the set of all alternatives in  $A$  ranked  $m$  or better by  $p^{\mathbb{P},\pi}$ . By construction  $B$  is also the set of alternatives in  $A$  ranked  $m$  or better by  $p$ . So  $B$  is both a  $p$ -truncation and a  $p^{\mathbb{P},\pi}$ -truncation of  $A$ .

Suppose  $a$  is the  $m$ 'th ranked alternative in  $A$  (and therefore also  $B$ ) under  $p^{\mathbb{P},\pi}(\cdot, A)$  while it is  $b$  under  $p(\cdot, A)$  with  $b \neq a$ . Since  $a$  is the lowest ranked alternative in  $B$  under  $p^{\mathbb{P},\pi}(\cdot, A)$  it must be that the first pairwise comparison in  $\mathbb{P}$  relevant to  $B$  is of the form  $(x, a)$  for some  $x \in B \setminus \{a\}$ . Suppose  $P_k$  is the cell of  $\mathbb{P}$  that contains this  $(x, a)$ . So for any  $y, z \in B$ ,  $(y, z) \notin P_j$  for all  $j < k$ . Then by construction of  $\mathbb{P}$  it must be that  $B \in \mathcal{C}_k$ . Then (again by construction)  $(x, a) \in P_k$  implies that since  $\{x, a\} \subseteq B$ , we get  $p(c, B) > p(a, B)$  for all  $c \in B \setminus \{a\}$ . Finally by ITR it must be that if  $B \in A^T(p)$  then  $p(c, A) > p(a, A)$  for all  $c \in B \setminus \{a\}$ . This contradicts the assertion that  $b$  is choice ranked last in  $B$  under  $p(\cdot, A)$ .  $\square$

**Theorem 6.** *A stochastic choice rule  $p$ , is rationalizable by gradual pairwise comparison if and only if  $p$  satisfies unique best and there exists a stochastic choice rule without ties  $p'$  that satisfies *sWARP*, ITR and *s-reducibility* such that for any  $A \in \mathcal{X}$  and  $a, b \in X$ ,*

$$p'(a, A) > p'(b, A) \Rightarrow p(a, A) \geq p(b, A).$$

**Proof of Theorem 6. Necessity:** It is sufficient to show that for any GPCR  $p^{\mathbb{P},\pi}$ , unique best is satisfied and that there exists a GPCR without ties  $p^{\mathbb{P}',\pi'}$  such that

for all  $A \in \mathcal{X}$  and  $a, b \in X$ ,  $p^{\mathbb{P}', \pi'}(a, A) > p^{\mathbb{P}', \pi'}(b, A) \Rightarrow p^{\mathbb{P}, \pi}(a, A) \geq p^{\mathbb{P}, \pi}(b, A)$ . The latter follows directly from lemma 2. As for the former, it has already been shown in the proof of theorem 3 that a *GPCR* necessarily satisfies sWARP. It is easy to see that sWARP implies unique best.

**Sufficiency:** Suppose  $p$  satisfies unique best and there exists an SCR without ties  $p'$  that satisfies sWARP, ITR and  $s$ -reducibility such that for all  $A \in \mathcal{X}$  and  $a, b \in X$

$$p'(a, A) > p'(b, A) \Rightarrow p(a, A) \geq p(b, A). \quad (6)$$

Since  $p'$  is an SCR without ties and satisfies sWARP, ITR and  $s$ -reducibility, by theorem 3 there exists a GPCR  $p^{\mathbb{P}, \pi}$  such that  $p^{\mathbb{P}, \pi} = p'$ . Moreover, since the (possible) ties in  $p$  are consistent with  $p' = p^{\mathbb{P}, \pi}$  as in 6, by theorem 2, there exists  $\pi'$  such that  $p = p^{\mathbb{P}, \pi'}$ .  $\square$

**Proof of Theorem 4.** “ $\Rightarrow$ ” : If  $(a, b) \in P$  then  $p^{\mathbb{P}, \pi}(a, \{a, b\}) > p^{\mathbb{P}, \pi}(b, \{a, b\})$  since by assumption  $\pi(P_{\mathbb{P}(A)}, A) > 0$  for all  $A \in \mathcal{X}$ .

“ $\Leftarrow$ ” : The fact that  $a$  is stochastically revealed preferred to  $b$  implies that there exists some  $A \in \mathcal{X}$  with  $b \in A$  such that  $M_{\mathbb{P}(A)}^{\mathbb{P}}(A) = \{a\}$ . The result then follows directly from lemma 4.  $\square$

**Lemma 5.** Suppose  $p^{\mathbb{P}, \pi}$  is an SPCR with  $P_j = (a, b)$ ,  $P_{j+1} = (x, y)$  and  $\pi(P_i, \cdot) > 0$  for all  $P_i \in \mathbb{P}$ . If  $Z \not\subseteq M_{j-1}^{\mathbb{P}}(A)$  for all  $A \in \mathcal{X}$  where  $Z = \{a, b\} \cup \{x, y\}$  then

$$p^{\mathbb{P}, \pi}(v, B) > p^{\mathbb{P}, \pi}(w, B) \Leftrightarrow p^{\mathbb{P}', \pi'}(v, B) > p^{\mathbb{P}', \pi'}(w, B), \quad \forall v, w \in B, B \in \tilde{\mathcal{X}}$$

where  $P'_i = P_i$  for all  $i < j$  and  $i > j + 1$ ,  $P'_j = (x, y)$  and  $P'_{j+1} = (a, b)$  and  $\pi'(P'_i, \cdot) > 0$  for all  $P'_i \in \mathbb{P}'$ .

*Proof.* Fix  $B \in \tilde{\mathcal{X}}$ . Suppose  $v \in B$  and  $v \notin Z$ . Then clearly  $v \in M_i^{\mathbb{P}}(B) \Leftrightarrow v \in M_i^{\mathbb{P}'}(B)$  for all  $1 \leq i \leq I$ . Further, if  $v \in M_{j-1}^{\mathbb{P}}(B)$  then  $v \in M_{j+1}^{\mathbb{P}}(B)$ . If instead,  $v \in B \cap Z$ , then  $v \in M_i^{\mathbb{P}}(B) \Leftrightarrow v \in M_i^{\mathbb{P}'}(B)$  for all  $i < j$ .

Now since  $Z \not\subseteq M_{j-1}^{\mathbb{P}}(B)$ , it must be that either  $M_j^{\mathbb{P}}(B) = M_{j-1}^{\mathbb{P}}(B)$  or  $M_{j+1}^{\mathbb{P}}(B) = M_j^{\mathbb{P}}(B)$  or both. If  $M_j^{\mathbb{P}}(B) = M_{j-1}^{\mathbb{P}}(B)$  then it must be that  $M_{j+1}^{\mathbb{P}'}(B) = M_j^{\mathbb{P}'}(B)$ . Further, since  $P'_j = P_{j+1}$ , it follows that  $M_j^{\mathbb{P}'}(B) = M_{j+1}^{\mathbb{P}}(B)$  and therefore  $M_{j+1}^{\mathbb{P}'}(B) = M_{j+1}^{\mathbb{P}}(B)$ . Suppose now that  $M_{j+1}^{\mathbb{P}}(B) = M_j^{\mathbb{P}}(B)$ , while  $M_j^{\mathbb{P}}(B)$  may or may not be the same as  $M_{j-1}^{\mathbb{P}}(B)$ . Either way, it must be that  $M_j^{\mathbb{P}'}(B) = M_{j-1}^{\mathbb{P}'}(B)$ . This in turn means that  $M_{j+1}^{\mathbb{P}'}(B) = M_j^{\mathbb{P}}(B)$ . Therefore again  $M_{j+1}^{\mathbb{P}'}(B) = M_{j+1}^{\mathbb{P}}(B)$ . As a result,  $Z \not\subseteq M_{j-1}^{\mathbb{P}}(B)$  implies that  $M_{j+1}^{\mathbb{P}'}(B) = M_{j+1}^{\mathbb{P}}(B)$  and that no more than one element in  $M_{j-1}^{\mathbb{P}}(B)$  could be missing from  $M_{j+1}^{\mathbb{P}}(B)$ , and this element must belong in  $Z$ . Note that  $M_{j+1}^{\mathbb{P}'}(B) = M_{j+1}^{\mathbb{P}}(B)$  implies that  $M_i^{\mathbb{P}'}(B) = M_i^{\mathbb{P}}(B)$  for all  $i \geq j+1$ .

In summary,  $M_i^{\mathbb{P}'}(B) = M_i^{\mathbb{P}}(B)$  for all  $i \neq j$  and  $|M_{j+1}^{\mathbb{P}}(B) \setminus M_{j-1}^{\mathbb{P}}(B)| \leq 1$ . So

$$\begin{aligned} p^{\mathbb{P},\pi}(v, B) > p^{\mathbb{P},\pi}(w, B) &\Leftrightarrow v \in M_i^{\mathbb{P}}(B), w \notin M_i^{\mathbb{P}}(B) && \text{for some } 1 \leq i \leq I \\ &\Leftrightarrow v \in M_k^{\mathbb{P}'}(B), w \notin M_k^{\mathbb{P}'}(B) && \text{for some } 1 \leq k \leq I \\ &\Leftrightarrow p^{\mathbb{P}',\pi'}(v, B) > p^{\mathbb{P}',\pi'}(w, B). \end{aligned}$$

The only non-obvious component of the second implication above involves  $v, w \in M_i^{\mathbb{P}}(B)$  for all  $i < j$  and  $v \in M_j^{\mathbb{P}}(B)$  but  $w \notin M_j^{\mathbb{P}}(B)$ . Then notice that  $|M_{j+1}^{\mathbb{P}}(B) \setminus M_{j-1}^{\mathbb{P}}(B)| \leq 1$  ensures that  $v \in M_{j+1}^{\mathbb{P}}(B)$  while still  $w \notin M_{j+1}^{\mathbb{P}}(B)$ . The implication then follows from  $M_{j+1}^{\mathbb{P}'}(B) = M_{j+1}^{\mathbb{P}}(B)$ .  $\square$

**Proof of Theorem 5.** “ $\Leftarrow$ ” : Suppose  $(a, b) \triangleright_p (x, y)$  and let  $p = p^{\mathbb{P},\pi}$ . So, there exists  $A \in \mathcal{X}$  with  $\{x, y\} \subseteq A \cup \{a\}$  such that  $p(b, A) > p(z, A)$  for some  $z \in A$  and  $p(b, A \cup \{a\}) < p(w, A \cup \{a\})$  for all  $w \in A \cup \{a\}$ ,  $w \neq b$ . Without loss of generality suppose  $z$  is choice ranked last in  $A$ . Further let  $1 \leq i \leq I$  be the smallest number for which  $z \notin M_i^{\mathbb{P}}(A)$  (by assumption,  $b \in M_i^{\mathbb{P}}(A)$ ). Therefore  $(w, v) \notin P_j$  for all  $j < i$  and  $w, v \in A$ . Now,  $b$  being choice ranked last in  $A \cup \{a\}$  implies that  $b \notin M_j^{\mathbb{P}}(A \cup \{a\})$  for some  $j < i$ . This means  $(a, b) \in P_j$  for some  $j < i$ . So  $(a, b)$  precedes  $(w, v)$  in the ordered partition for all  $w, v \in A$ . Further, since  $b$  is choice ranked last,  $(a, b)$  precedes  $(a, v)$  in the ordered partition for all  $v \in A$ . Therefore

$(a, b)$  is revealed compared before  $(x, y)$ .

By the transitivity of the revealed compared relation, it follows that if  $(a, b) \overline{\mathcal{E}}_p(x, y)$  then also  $(a, b)$  is revealed compared before  $(x, y)$ .

“ $\Rightarrow$ ” : Suppose  $p = p^{\mathbb{P}, \pi}$ . By lemma 1,  $p^{\mathbb{P}, \pi}$  is assumed to be an SPCR without loss of generality. Further, since  $p$  is without ties, we can set  $\pi(P_i, \cdot) > 0$  for all  $P_i \in \mathbb{P}$ . Now suppose  $P_j = (a, b)$  and  $P_{j+1} = (x, y)$ . It will be shown that  $(a, b) \not\mathcal{F}_p(x, y)$  implies that  $Z \not\subseteq M_{j-1}^{\mathbb{P}}(A)$  for all  $A \in \mathcal{X}$  where  $Z = \{a, b\} \cup \{x, y\}$ . Then by lemma 5 a new ordered partition  $\mathbb{P}'$  which is exactly the same as  $\mathbb{P}$  except for the interchanged positions of  $(a, b)$  and  $(x, y)$  would generate the same choice ranks as  $\mathbb{P}$ . By theorem 2, there then exists an appropriate  $\pi'$  such that  $p = p^{\mathbb{P}', \pi'}$ . This would establish that  $(a, b)$  is not revealed compared before  $(x, y)$ .

Consider the sets  $Z \setminus \{a\}$  and  $Z$ . If  $M_{j-1}^{\mathbb{P}}(Z) = Z$  and  $M_{j-1}^{\mathbb{P}}(Z \setminus \{a\}) = Z \setminus \{a\}$  then it must be that  $y$  is choice ranked last in  $Z \setminus \{a\}$  while  $b$  is choice ranked last in  $Z$  under  $p^{\mathbb{P}, \pi}$ . So, if  $(a, b) \mathcal{F}_p(x, y)$  then it must be that  $M_{j-1}^{\mathbb{P}}(Z) \neq Z$ . Therefore there must be  $v, w \in Z$  such that  $P_i = (v, w)$  for some  $i < j$ . This in turn means that for any  $A \in \mathcal{X}$ ,  $Z \not\subseteq M_{j-1}^{\mathbb{P}}(A)$ , since either  $w \notin M_{j-1}^{\mathbb{P}}(A)$  or  $v \notin M_{j-1}^{\mathbb{P}}(A)$ .

It has therefore been shown that if  $(a, b) \not\mathcal{F}_p(x, y)$  and if  $(a, b)$  and  $(x, y)$  hold adjacent positions in  $\mathbb{P}$ , with  $(a, b)$  preceding  $(x, y)$ , where  $p^{\mathbb{P}, \pi} = p$  then switching the order of  $(a, b)$  and  $(x, y)$  and leaving all else unchanged generates the same choice ranks. Suppose now that  $(a, b)$  still precedes  $(x, y)$  but with other comparisons in between in  $\mathbb{P}$ . Then take the comparison  $(p, q)$  nearest to  $(x, y)$  along the order in  $\mathbb{P}$ , which lies between  $(a, b)$  and  $(x, y)$  such that  $(a, b) \overline{\mathcal{E}}_p(p, q)$ . Since  $(a, b) \not\mathcal{F}_p(x, y)$  it must be that  $(p, q) \not\mathcal{F}_p(x, y)$ . The procedure outlined above then shows that  $(p, q)$  can be shifted one adjacent switch at a time to eventually take a position after  $(x, y)$  in the ordered partition of  $P$ , without affecting any choice ranks. This same procedure can be carried out for the comparisons from  $(p, q)$  all the way to  $(a, b)$  without changing the order of the comparisons between and including  $(a, b)$  and  $(p, q)$ . This would eventually lead to  $(a, b)$  following  $(x, y)$  in the ordered partition and yet generating

the same choice ranks as before. This concludes the proof.  $\square$

**Proof of Observation 2.** Suppose  $p$  is an SCR that violates weak stochastic transitivity. Then there exists  $\{a, b, c\} \subseteq X$  such that  $p(a, \{a, b\}) \geq 1/2$ ,  $p(b, \{b, c\}) \geq 1/2$  but  $p(c, \{a, c\}) > 1/2$ . Then for some  $x \in \{a, b, c\}$ ,  $p(x, \{a, b, c\}) \geq p(y, \{a, b, c\})$  for all  $y \in \{a, b, c\}$ . This violates sWARP.  $\square$

**Proof of Proposition 1.** (i) Suppose  $\{(b, c), (c, b)\} \cap P_k \neq \emptyset$ . Let  $W = \{i | \{a, b\} \subseteq M_i^{\mathbb{P}}(A)\}$  and  $Z = \{i | \{a, b\} \cap M_i^{\mathbb{P}}(A) = \{b\}\}$ . Then the relevant expressions are:

$$\begin{aligned} p^{\mathbb{P}, \pi}(b, A) &= \sum_{i \in W} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} + \sum_{i \in Z} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} \\ p^{\mathbb{P}, \pi}(b, A \cup \{c\}) &= \sum_{i \in W} \frac{\pi(P_i, A \cup \{c\})}{|M_i^{\mathbb{P}}(A) + 1|} + \sum_{i \in Z, i < k} \frac{\pi(P_i, A \cup \{c\})}{|M_i^{\mathbb{P}}(A) + 1|} + \mathbb{1}(b \succ c) \sum_{i \in Z, i \geq k} \pi(P_i, A \cup \{c\}) \\ p^{\mathbb{P}, \pi}(a, A) &= \sum_{i \in W} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|}, \quad p^{\mathbb{P}, \pi}(a, A \cup \{c\}) = \sum_{i \in W} \frac{\pi(P_i, A \cup \{c\})}{|M_i^{\mathbb{P}}(A) + 1|}. \end{aligned}$$

So,

$$\begin{aligned} p^{\mathbb{P}, \pi}(b, A) - p^{\mathbb{P}, \pi}(b, A \cup \{c\}) &= p^{\mathbb{P}, \pi}(a, A) - p^{\mathbb{P}, \pi}(a, A \cup \{c\}) \\ &+ \sum_{i \in Z, i < k} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|} - \sum_{i \in Z, i < k} \frac{\pi(P_i, A \cup \{c\})}{|M_i^{\mathbb{P}}(A) + 1|} \\ &+ \sum_{i \in Z, i \geq k} \pi(P_i, A) - \mathbb{1}(b \succ c) \sum_{i \in Z, i \geq k} \pi(P_i, A \cup \{c\}) \end{aligned}$$

This is equal to

$$\begin{aligned} &p^{\mathbb{P}, \pi}(a, A) - p^{\mathbb{P}, \pi}(a, A \cup \{c\}) \\ &+ \sum_{i \in Z, i < k} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|(|M_i^{\mathbb{P}}(A)| + 1)} + \sum_{i \in Z, i \geq k} \pi(P_i, A) \\ &+ \sum_{i \in Z, i < k} \frac{\pi(P_i, A) - \pi(P_i, A \cup \{c\})}{|M_i^{\mathbb{P}}(A)| + 1} - \mathbb{1}(b \succ c) \sum_{i \in Z, i \geq k} \pi(P_i, A \cup \{c\}) \end{aligned}$$

The second line in that expression is strictly positive since  $p^{\mathbb{P},\pi}(b, A) > p^{\mathbb{P},\pi}(a, A)$ . The third line in that expression can be no smaller than  $-\epsilon - \nu$ . Therefore for sufficiently small  $\epsilon$  and  $\nu$  we get  $p^{\mathbb{P},\pi}(b, A) - p^{\mathbb{P},\pi}(b, A \cup \{c\}) > p^{\mathbb{P},\pi}(a, A) - p^{\mathbb{P},\pi}(a, A \cup \{c\})$ .

(ii) Let  $E = \{i \mid \{a, b\} \cap M_i^{\mathbb{P}}(A) = \{a\}\}$ . Then similarly to part (i) we get

$$p(a, A) - p(a, A \cup \{c\}) = p(b, A) - p(b, A \cup \{c\}) + \sum_{i \in E} \frac{\pi(P_i, A) - \pi(P_i, A \cup \{c\})}{|M_i^{\mathbb{P}}(A)|}.$$

So  $p(a, A) - p(a, A \cup \{c\}) < p(b, A) - p(b, A \cup \{c\}) + \epsilon$ . This in turn means

$$\frac{p(a, A) - p(a, A \cup \{c\})}{p(a, A)} < \frac{p(b, A) - p(b, A \cup \{c\}) + \epsilon}{p(a, A)}.$$

Then as long as  $p(b, A) - p(b, A \cup \{c\}) > 0$ , for small enough  $\epsilon$  we must have

$$(p(a, A) - p(a, A \cup \{c\}))/p(a, A) < (p(b, A) - p(b, A \cup \{c\}))/p(b, A).$$

$$p(b, A) - p(b, A \cup \{c\}) = \sum_{b \in M_i^{\mathbb{P}}(A)} \frac{\pi(P_i, A) - \pi(P_i, A \cup \{c\})}{|M_i^{\mathbb{P}}(A)| + 1} + \sum_{b \in M_i^{\mathbb{P}}(A)} \frac{\pi(P_i, A)}{|M_i^{\mathbb{P}}(A)|(|M_i^{\mathbb{P}}(A)| + 1)}$$

This is strictly greater than  $\frac{p(b, A)}{|A|} - \epsilon$ , which is positive for small enough  $\epsilon$ .  $\square$

## Independence of Axioms

To see the independence of the three axioms used to characterize the GPC procedure, consider the following examples.

**Example A.1.**[Violation of *sWARP*]

$p(\cdot, A)$	$A = \{a, b, c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$
$a$	1/2	0	1/4	0
$b$	1/3	3/4	0	3/4
$c$	1/6	1/4	3/4	1/4

The only non-trivial truncation here is that of  $\{a, b\}$  from  $\{a, b, c\}$  and the choice

rank ordering is the same in both sets. The pair  $\{b, c\}$  satisfies the requirement of s-reducibility in any collection with a set containing  $\{b, c\}$ . For any other collection, both  $a$  is always ranked last in the presence of  $c$  and  $c$  is always ranked last in the presence of  $b$ . The violation of sWARP follows from  $a$  being stochastically revealed preferred to  $c$  and vice versa.

**Example A.2.**[Violation of *ITR*]

$p(\cdot, A)$	$A = \{a, b, c, d\}$	$\{a, b, c\}$	$\{a, b, d\}$	$\{a, c, d\}$	$\{b, c, d\}$	...
$a$	1/2	1/2	1/2	1/2	0	...
$b$	1/3	1/6	1/3	0	1/2	...
$c$	1/6	1/3	0	1/3	1/3	...
$d$	0	0	1/6	1/6	1/6	...

Here the change in the rank ordering of  $b$  and  $c$  in the sets  $\{a, b, c, d\}$  and  $\{a, b, c\}$  despite the latter being a truncation of the former violates ITR. Any completion of the example where the choice rank ordering in any two-alternative choice set aligns with the strict preference ordering  $a \succ b \succ c \succ d$ , satisfies both sWARP and s-reducibility. For the latter, notice that  $d$  is ranked last in the presence of  $c$  in  $\tilde{\mathcal{X}}$ . For any collection without a set containing  $\{c, d\}$ , but with one containing  $\{a, d\}$ ,  $d$  is ranked last in the presence of  $a$ . In any other collection with at least one set with more than 2 elements,  $b$  is ranked last in the presence of  $a$ . Any collection of sets that contain no more than two alternatives satisfies the requirement of s-reducibility trivially.

**Example A.3.**[Violation of *s-reducibility*]

$p(\cdot, A)$	$A = \{a, b, c, d\}$	$\{b, c, d\}$	$\{a, c, d\}$	...
$a$	1/2	0	1/2	...
$b$	1/3	1/2	0	...
$c$	1/6	1/6	1/6	...
$d$	0	1/3	1/3	...



It is easy to fill in the remaining data points to satisfy both sWARP and ITR, but no such choice rule can satisfy s-reducibility. The choice ranks in the set  $\{a, b, c, d\}$  require that  $d$  is ranked last in the presence of some alternative  $y$  in the collection of the three sets mentioned in the example, with  $y \in \{a, b, c\}$ . It cannot be  $a$  or  $c$  since  $d$  is not ranked last in  $\{a, c, d\}$  and not  $b$  since  $d$  is not ranked last in  $\{b, c, d\}$ .

## References

- [1] Agranov, Marina and Pietro Ortoleva. 2017. “Stochastic Choice and Preferences for Randomization.” *Journal of Political Economy* 125 (1) 40-68.
- [2] Ahumada, Alonso and Levent Ulku. 2018. “Luce Rule with Limited Consideration.” *Mathematical Social Sciences* 93 52-56.
- [3] Apesteguia, Jose and Miguel A. Ballester. 2013. “Choice by Sequential Procedures.” *Games and Economic Behavior* 77 90–99.
- [4] Brady, Richard L. and John Rehbeck. 2016. “Menu-Dependent Stochastic Feasibility.” *Econometrica* 84(3) 1203-1223.
- [5] Bernheim, Douglas, and Antonio Rangel. 2009. “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics.” *Quarterly Journal of Economics* 124(1) 51–104.
- [6] Cattaneo, Matias D., Xinwei Ma, Yusufcan Masatlioglu and Elchin Suleymanov. 2018. “A Random Attention Model.” *working paper*.
- [7] Debreu, Gerard. 1960. “Review of ‘Individual Choice Behaviour’ by R. D. Luce.” *American Economic Review* 50 186-188.
- [8] Dutta, Rohan and Sean Horan. 2014. “Inferring Rationales from Choice: Identification for Rational Shortlist Methods.” *AEJ: Microeconomics* 7(4) 179-201.

- [9] Echenique, Federico and Kota Saito. 2018. “General Luce Model.” *Economic Theory* forthcoming.
- [10] Fudenberg, Drew, Ryota Iijima and Tomasz Strzalecki. 2015. “Stochastic Choice and Revealed Perturbed Utility.” *Econometrica* 83(6) 2371-2409.
- [11] Hirshleifer, David, Yaron Levi, Ben Lourie and Siew Hong Teoh. 2019 “Decision Fatigue and Heuristic Analyst Forecasts.” *Journal of Financial Economics* 133 83-98.
- [12] Horan, Sean. 2018. “Threshold Luce Rules.” working paper.
- [13] Levav, Jonathan, Mark Heitmann, Andreas Herrmann and Sheena S. Iyengar. 2010. “Order in Product Customization Decisions: Evidence from Field Experiments.” *Journal of Political Economy* 118(2) 274-299.
- [14] Luce, Robert D. 1959. “Individual Choice Behavior: A Theoretical Analysis.” *New York: Wiley*.
- [15] Manzini, Paola and Marco Mariotti. 2007. “Sequentially Rationalizable Choice.” *American Economic Review* 97(5) 1824–1839.
- [16] Manzini, Paola and Marco Mariotti. 2012. “Choice by Lexicographic Semi-orders.” *Theoretical Economics* 7 1–23.
- [17] Manzini, Paola and Marco Mariotti. 2014. “Stochastic Choice and Consideration Sets.” *Econometrica* 82(3) 1153-1176.
- [18] Masatlioglu, Yusufcan, Daisuke Nakajima and Erkut Ozbay. 2012. “Revealed Attention.” *American Economic Review* 102(5) 2183–2205.
- [19] Natenzon, Paulo. 2018. “Random Choice and Learning.” *Journal of Political Economy* forthcoming.

- [20] Noguchi, Takao and Neil Stewart. 2014. "In the Attraction, Compromise and Similarity Effects, Alternatives are Repeatedly Compared in Pairs on Single Dimensions." *Cognition* 132 44-56.
- [21] Noguchi, Takao and Neil Stewart. 2018. "Multialternative Decision by Sampling: A Model of Decision Making Constrained by Process Data." *Psychological Review* 4 512-544.
- [22] Ravid, Doron and Kai Steverson. 2018. "Focus, then Compare." *working paper*.
- [23] Reiskamp, Jorg, Jerome R. Busemeyer and Barbara A. Mellers. 2006. "Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice." *Journal of Economic Literature* 44 631-661.
- [24] Reutskaja, Elena, Rosemarie Nagel, Colin F. Camerer and Antonio Rangel. 2011. "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study." *American Economic Review* 101 900-926.
- [25] Rubinstein, Ariel. 1988. "Similarity and Decision-Making under Risk." *Journal of Economic Theory* 46(1) 145-153.
- [26] Russo, J. Edward and France Leclerc. 1994. "An Eye-Fixation Analysis of Choice Processes for Consumer Nondurables." *Journal of Consumer Research* 21(2) 274-290.
- [27] Russo, J. Edward and Larry D. Rosen. 1975. "An Eye Fixation Analysis of Multialternative Choice." *Memory and Cognition* 3(3) 267-276.
- [28] Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *The Quarterly Journal of Economics* 69(1) 99-118.
- [29] Tversky, Amos. 1972a. "Choice by Elimination." *Journal of Mathematical Psychology* 9 341-367.

- [30] Tversky, Amos. 1972b. "Elimination by Aspects: A Theory of Choice." *Psychological Review* 79(4) 281-299.
- [31] Wedel, Michel and Rik Pieters. 2008. "A Review of Eye-Tracking Research in Marketing." in Naresh K. Malhotra (ed.) *Review of Marketing Research* Volume 4, Emerald Group Publishing Limited, 123-147.