CIREQ

Centre interuniversitaire de
recherche en économie quantitative

Cahier 23-2018

Gradual Pairwise Comparison and Stochastic Choice

Rohan Dutta

CIREQ

Cahier 23-2018

Gradual Pairwise Comparison and Stochastic Choice

Rohan Dutta

CIREQ, Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal (Québec) H3C 3J7 Canada
☎ (514) 343-6557 / 🖷 (514) 343-7221
cireq@umontreal.ca
http://www.cireqmontreal.com
@CIREQMTL
https://www.facebook.com/CIREQ/

Dépôt légal - Bibliothèque nationale du Canada, 2018, ISSN 0821-4441
Dépôt légal - Bibliothèque et Archives nationales du Québec, 2018
ISBN-13 : 978-2-89382-736-0

Gradual Pairwise Comparison and Stochastic Choice*

Rohan Dutta†

November 2018

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. 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, attraction and compromise effects.

Keywords: Revealed preference, bounded rationality, stochastic choice.

*I thank David Levine, Salvatore Modica, Yusufcan Masatlioglu and Anna Winterbottom for helpful comments and encouragement. All errors are mine.

†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), finds compelling evidence of the actual choice procedure consisting primarily of a sequence of pairwise comparisons. For a rational agent capable of considering all relevant comparisons 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 pairwise comparisons of her underlying strict rational preference sequentially in some fixed menu-independent order to remove inferior alternatives from a given choice set. Facing a choice problem she may be forced to stop at different points along this order according to some unobserved and possibly 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 and Luce rules (Luce (1959)), but also choice data where the order of choice probabilities across alternatives is menu-dependent. No specific stand is taken on what causes the random stopping or what determines the sequence of comparisons. However, the distribution function modelling random stopping could capture fatigue (from making multiple comparisons). Similarly, the sequence in which the agent considers the pairwise comparisons could reflect the ease of making such comparisons. Modelling these as such leads in a natural way to choice exhibiting behavioural phenomena such as the similarity, attraction and compromise effects (Section 4). While accommodating the similarity effect, the analysis highlights the incomplete (and therefore partly misleading) nature of the famous Debussy-Beethoven thought experiment in Debreu (1960).

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 5) and each GPCR is shown to correspond to a unique underlying strict preference (Theorem 7). The latter is easily identified with the agent strictly preferring a to b if and only if a has 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 with some additional structure on the choice procedure (Section 3.3). Using a revealed preference approach it is then possible to uncover the entire sequence of pairwise comparisons. Section 5.2 provides a construction that can be used to rationalize any choice data in which the order of choice probabilities for any two available alternatives is menu-independent, which includes any Luce rule.

This study owes a considerable debt to Apesteguia and Ballester (2013) (AB) and Manzini and Mariotti (2012) (MM), who study a decision-maker who makes pairwise comparisons sequentially (without stopping prematurely) to make a choice. The resulting deterministic choice behaviour can violate WARP if the pairwise comparisons violate transitivity. By contrast, in the current study the pairwise comparisons are constituents of a strict rational preference. As a result, the only overlap is deterministic rational choice. Despite the difference in scope, AB and MM make 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.[1] This idea proves very useful 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 a couple of examples. Section 3 contains all the characterization results. The proofs of results in section 3.1 are retained in the main body of the text to help the reader get a better sense of how the model works. All other proofs are collected in the appendix. Section 4 discusses the specific ways in which GPCRs can accommodate violations of stochastic transitivity and other forms of menu-dependent

[1] Perhaps more importantly they showed that such models admit a simple characterization through restrictions on choice data, despite the complexity. As evidence of the latter, identification of the sequence of pairwise comparisons in these models remains an open problem. The present study has more success on this count (see Section 3.3).

choice behaviour. Section 5 concludes by comparing the GPC choice procedure 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. Being a strict rational preference requires P to 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 $(b, c) \in P \Rightarrow (a, c) \in P$. 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* $p : X \times \mathcal{X} \to [0, 1]$ *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* $1 \leq$

$i \leq I$. Given a choice set she 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 in P_2 and so on. For a given ordered partition $\mathbb{P}$ of P and a choice set $A \in \mathcal{X}$, define the following sets recursively,

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

$M_i^{\mathbb{P}}(A)$ contains all alternatives that survive after the decision maker has considered the ith 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* $i = 1$.[2] $\tilde{I}^{\mathbb{P}}$ is well defined since $\mathbb{P}$ is a partition of a strict rational preference P.[3]

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 (possibly menu-dependent) 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***.[4] For any choice set A, $\pi(P_i, A)$ is the probability that the decision maker stops at cell P_i and does not make the comparisons contained in subsequent cells. $\pi(P_0, A)$ is the probability with which she considers no cell 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 choice procedure entails the decision-maker choosing with equal odds from among the alternatives that remain, $M_i^{\mathbb{P}}(A)$.

Definition 2. *A* ***gradual pairwise comparison rule*** *(GPCR) is a stochastic*

[2] $|B|$ denotes the number of elements in the set B.

[3] 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).

[4] The obvious dependence of the stopping function on the partition $\mathbb{P}$ is suppressed for notational convenience.

choice rule $p^{\mathbb{P},\pi}$ with an ordered partition $\mathbb{P}$ of a strict rational preference P and a stopping function π such that for all $A \in \mathcal{X}$,

$$p^{\mathbb{P},\pi}(a,A) = \begin{cases} \sum\limits_{\{i|a\in M_i^{\mathbb{P}}(A)\}} \dfrac{\pi(P_i,A)}{|M_i^{\mathbb{P}}(A)|} & \textit{if } a \in A \\ 0 \quad \textit{otherwise.} \end{cases} \tag{2}$$

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 π such that

$$p = p^{\mathbb{P},\pi}.$$

To see how the procedure works consider the following two examples.

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

a	$\left(7100, 1005;\ \frac{3}{4}, \frac{1}{4}\right)$
b	$\left(7000, 1000;\ \frac{3}{4}, \frac{1}{4}\right)$
c	$\left(6650, 490;\ \frac{4}{5}, \frac{1}{5}\right)$

Read the table above as lottery a yields \$7100 with probability 3/4 and \$1005 with probability 1/4, and so on.

In both examples, the agent's underlying preference ranks a over b over c. Nevertheless, the agent 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.

Example 1a. $X = \{a,b,c\}$. 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

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

The decision maker is equally likely to stop at any element of her ordered partition irrespective of the choice set; $\pi(P_i,A) = 1/3$ for all $A \in \mathcal{X}$ and all $i \in \{1,2,3\}$.

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	5/6	0	5/6
b	0	0	2/3	0
c	0	1/6	1/3	1/6

For instance, $p^{\mathbb{P},\pi}(b, \{b, c\}) = 2/3$ and $p^{\mathbb{P},\pi}(c, \{a, b, c\}) = 1/6$.

Example 1b. Everything is as in example 1, except the stopping function. In particular, $X = \{a, b, c\}$, $P = \{(a, b), (a, c), (b, c)\}$ and $\mathbb{P}$ is

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

The decision maker is now equally likely to stop only at those elements of her ordered partition that are relevant to her choice problem.

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

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	1	0	5/6
b	0	0	1	0
c	0	0	0	1/6

The two different stopping function specifications in examples 1a and 1b capture different sources of the bounded rationality constraining the agent from making all relevant comparisons. 1b captures something akin to fatigue. While perfectly able to carry out one round of pairwise comparison, facing the prospect of another, the agent may "give up". By contrast, in 1a the agent finds the comparison between b and c intrinsically difficult enough that even if she were only choosing between b and c, with positive probability she would "give up" and choose with equal odds from the two.

Menu-independent stopping functions seem appropriate for 1a while menu-dependent stopping functions better suit 1b. Both these sources of bounded rationality, however, seem reasonable and therefore the subsequent analysis, instead of making an arbitrary restriction, allows for both.

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 π, play very different roles in determining choice rank and choice probabilities. Choice rank is essentially determined by the sequence in which the constituent pairwise comparisons of the rational preference are considered, the ordered partition $\mathbb{P}$.

Theorem 1. *Fix an ordered partition $\mathbb{P}$ of some $P \in \mathcal{P}$. Let π and π' 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

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

□

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*

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

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

$$p^{\mathbb{P},\pi}(a, A) > p^{\mathbb{P},\pi}(b, A) \Rightarrow p(a, A) \geq p(b, A) \qquad \textit{and}$$
$$p^{\mathbb{P},\pi}(a, A) = p^{\mathbb{P},\pi}(b, A) \Rightarrow p(a, A) = p(b, A).$$

Then there exists π' such that

$$p = p^{\mathbb{P},\pi'}.$$

Proof. Given $\mathbb{P}$ and some $A \in \mathcal{X}$, define the sequence of sets, $\{E_i(A)\}_{i=1}^{I}$ where $E_i(A) = \{a \in A | a \notin M_i^{\mathbb{P}}(A) \textit{ and } a \in M_{i-1}^{\mathbb{P}}(A)\}$. $E_i(A)$ 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, π' 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 π' 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 π' such that

$$\sum_{i=j}^{k-1} \pi'(P_i, A) = |M_{k-1}^{\mathbb{P}}(A)| \left[p(\alpha_k, A) - p(\alpha_j, A) \right]. \tag{3}$$

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.

π', 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 $Z(A) = \{i | E_i(A) > 0 \textit{ and } E_k(A) > 0 \textit{ for some } k > i\}$. 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,

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

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 $a \in E_j(A)$. $\square$

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.

The next simple and useful result shows that every GPCR can also be represented as an SPCR.

Theorem 3. *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 π_{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 $i > k$, $\pi_{n+1}(P'_i, \cdot) = \pi_n(P_i, \cdot)$ 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. □

Theorem 2 implies that starting with a GPCR without ties, it is possible to construct a GPCR with ties, as long as these ties are consistent with the choice ranks of the original GPCR. Something of a converse to this statement is true too. Starting with a GPCR with ties, it is possible to construct a GPCR without ties such that the ties in the original GPCR are consistent with the choice ranks of the constructed GPCR.

Theorem 4. *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 Theorem 3 there exists an SPCR, say $p^{\overline{\mathbb{P}},\overline{\pi}}$, such that $p^{\mathbb{P},\pi} = p^{\overline{\mathbb{P}},\overline{\pi}}$. Set $\mathbb{P}' = \overline{\mathbb{P}}$. Pick any stopping function π' 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,

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

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

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 $\pi(P_{\tilde{I}(A)}, A) = 1$ 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 hardly surprising, since setting $\pi(P_{\tilde{I}(A)}, A) = 1$ implies that the decision-maker is able to make all relevant comparisons before making her decision.

More interestingly, the GPC procedure can rationalize choice reversals. For instance, in example 1b, adding the alternative a to the choice set $\{b, c\}$, increases the probability of c being selected from 0 to 1/6 while reducing that of b from 1 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. The task will be split into two parts. First the set of all stochastic choice rules without ties that are rationalizable by GPC will be characterized. This result will then be used to characterize all stochastic choice rules that are rationalizable by GPC, including those with ties.

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*

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

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.*

A simple implication of *sWARP* is that for any choice set there is a unique alternative with the highest choice probability. Further, 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. *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

$A \in A^T(p)$. Let $\tilde{\mathcal{X}}$ denote the subset of $\mathcal{X}$ containing all subsets of X with at least two elements. 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,*

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

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 ordering of any pair of alternatives 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.

The final axiom is labeled s(tochastic)- reducibility.

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*

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

$s-reducibility$ 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.

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 5. *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.

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 $a, b \in X$,*

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

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.[5] This is a critical ingredient in the construction used to prove the sufficiency part of theorem 5.

To see the independence of the three axioms, consider the following examples.

[5]The approach taken in Apesteguia and Ballester (2013) is a little different, but with a similar purpose.

Example 2.[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

In example 2 the only non-trivial truncation 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 3.[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	...

In example 3 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 4.[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 of example 4 in a way that satisfies both sWARP and ITR. Nevertheless, no such choice rule can satisfy s-reducibility. Given the choice ranks in the set $\{a, b, c, d\}$, it must be that d is ranked last in the presence of some alternative y in the collection of the three sets mentioned in the example, with $y \in \{a, b, c\}$. But y cannot be a or c since d is not ranked last in $\{a, c, d\}$. y cannot be b either since d is not ranked last in $\{b, c, d\}$.

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 preference relation.

Theorem 7. *Given a GPC choice rule $p^{\mathbb{P},\pi}$ where $\mathbb{P}$ is an ordered partition of a strict rational preference P,*

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

It is possible for different pairs of $\mathbb{P}$ and π to rationalize the same stochastic choice rule. Consider, for instance, the case of deterministic rational choice rationalizable by a strict rational preference P. In this case, any ordered partition $\mathbb{P}$ of P and π such that $\pi(P_{\tilde{I}(A)}, A) = 1$ can rationalize the choice rule. 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 5.

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

These are the choice probabilities generated by the GPC choice procedure described in example 1a, 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 1a) and after. For the latter, consider the following $\mathbb{P}$ and π which also rationalizes the choice rule.

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_1	1	1/3	2/3	1/3
P_2	0	0	1/3	0
P_3	0	2/3	0	2/3

This leads to an obvious question: when can we infer from choice data rationalizable by GPC that some pairwise comparison must have been considered before others. In other words, can choice data systematically reveal if a pairwise comparison is considered before another? Unfortunately, the GPC procedure turns out to be too flexible to obtain a clean inference. With a little more structure, however, it is possible to uncover the entire sequence of pairwise comparisons, as is studied in the next subsection.

3.3 Revealed Consideration and Revealed Preference

Consider now a decision-maker who in addition to making binary comparisons sequentially, does so in a certain order, labeled *salience order*.[6] In particular, she considers the alternative most salient in her choice set and eliminates all alternatives it dominates, and only then moves to the next most salient alternative. In other words, if she considers the pairwise comparison (a, b) before (c, d), then she does not consider any comparison (c, x) before (a, y). This leads to the following subclass of gradual pairwise comparison choice rules.

Definition 5 (Ordered Gradual Pairwise Comparison)**.** *An ordered gradual pairwise*

[6]I thank Yusufcan Masatlioglu for suggesting this interpretation.

comparison (OGPC) rule is a GPCR $p^{\mathbb{P},\pi}$ such that if $(a,b) \in P_i$ and $(c,d) \in P_j$ with $i < j$ and $(a,x) \in P_k$, $(c,y) \in P_l$ then $k < l$.

Characterizing the set of choice rules rationalizable by OGPC requires strengthening the s-reducibility axiom alone. Given any binary relation P on the set X and some alternative a, let $L^P(a)$ be the set of all alternatives ranked below a under P. Formally, $L^P(a) = \{x \in X | (a,x) \in P\}$.

Definition 6. *Given a collection of sets $\mathcal{B} \subseteq \mathcal{X}$, c is* ***outranked in the presence of a*** *by b, if there exists a sequence of sets $\{X_i\}_{i=1}^n$ and alternatives $\{x_i\}_{i=1}^{n+1} \subset L^P(a)$ such that for all $1 \leq i \leq n$,*

$$\{a, x_i, x_{i+1}\} \subseteq X_i \in \mathcal{B} \qquad \textit{and} \qquad p(x_i, X_i) > p(x_{i+1}, X_i)$$

with $x_1 = b$ and $x_{n+1} = c$, where P is the stochastically revealed preferred relation under p.

In words, c is outranked in the presence of a by b, if b either directly or indirectly has a higher choice rank than c with a available and stochastically revealed preferred to both. The direct route would involve a set containing a, b and c with b having a higher choice rank than c. The indirect route would involve a sequence of x_i, x_{i+1} with each x_{i+1} outranked in the presence of a by x_i by the direct route, and b, c belonging in the transitive closure of this binary relation.

Axiom 5 (**Strong s-Reducibility**). *For every nonempty collection of sets $\mathcal{B} \subseteq \tilde{\mathcal{X}}$, there exists $D \in \mathcal{B}$ and $a \in D$ with $L^P(a) \cap D \neq \emptyset$ such that*
(i) if $a \in A \in \mathcal{B}$ then

$$p(c, A) > p(d, A) \qquad \forall c \in A \setminus L^P(a), d \in L^P(a) \cap A, \qquad \textit{and}$$

(ii) given $\mathcal{B}$ and for all $b, c \in L^P(a)$, if c is outranked in the presence of a by b then

$$p(b, A) > p(c, A) \qquad \forall A \in \mathcal{B} \textit{ with } \{a, b, c\} \subseteq A$$

where P is the stochastically revealed preferred relation under p.

Recall that the order independence axiom requires, in any collection of sets, an option that is always ranked last when available. s-reducibility weakens this to the requirement of an option that is always ranked last in the presence of another specific (dominant) option. In other words, in the presence of the dominant option, the dominated option must have a worse choice rank than all other options. Strong s-reducibility strengthens this by requiring, for any collection of sets, the existence of a dominant action, such that any available option it is stochastically revealed preferred to (dominated alternative) must have a worse choice rank than those that are not. Further, the choice rank ordering of any pair of these dominated alternatives in any set in this collection in the presence of the dominant alternative must be consistent with some fixed order over all the dominated options.

Strong s-reducibility is still a weakening of the order independence axiom. Under the latter, for any given alternative in a choice set, the options it is stochastically revealed preferred to would have worse choice ranks than the others. Moreover the choice rank ordering of any pair of these dominated alternatives would always agree with the order defined on them by their choice ranks in the choice set X.

Theorem 8. *A stochastic choice rule p without ties can be rationalized by ordered gradual pairwise comparison if and only if p satisfies sWARP, ITR and strong s-reducibility.*

Theorem 9. *A stochastic choice rule p, is rationalizable by ordered 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 strong s-reducibility such that for any $A \in \mathcal{X}$ and $a, b \in X$,*

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

Recall that different pairs of $\mathbb{P}$ and π can be used to rationalize the same choice rule by GPC. This is true for rationalization by OGPC too. So to infer that a decision maker considers a particular pairwise comparison before another, it must be true for every pair of $\mathbb{P}$ and π that can rationalize the choice data.[7]

Definition 7. *For a choice rule p, rationalizable by OGPC, a is* ***revealed considered before*** *b if for any OGPC rule $p^{\mathbb{P},\pi}$ such that $p = p^{\mathbb{P},\pi}$,*

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

[7] The case for this conservative approach to extending revealed preference techniques to models of bounded rationality in choice is made compellingly in Masatlioglu, Nakajima and Ozbay (2012).

For choice probabilities generated by an OGPC choice procedure, the salience order is revealed by violations of the order independence property. The intuition is simple. Consider two alternatives, (at least) one of which is stochastically revealed preferred to by a dominant third alternative, in a menu. The relative choice ranks of these two alternatives in the menu depends on the order in which they are eliminated in the procedure. Suppose the dominated option is the lower ranked of the two. Then at the very least the dominated option must be getting eliminated either by the dominant one or some alternative more salient than the dominant one. If adding a new alternative changes the relative choice ranks of these original two alternatives then it must be that the order in which they are eliminated has changed. But this is only possible if the new alternative is considered before (more salient than) the dominant alternative.

Lemma 1. *Let p be rationalizable by OGPC and P be the stochastically revealed preference relation. If $\exists A \ni b$ and $d \in L^P(b)$ such that*

$$p(c, A) > p(d, A) \text{ and } p(c, A \cup \{a\}) < p(d, A \cup \{a\})$$

then a is revealed considered before b.

In words, if adding a to a set containing b reverses the choice rank ordering of two alternatives, with b stochastically revealed preferred to the originally lower ranked one, then a must be more salient than b.

For a given choice rule, p and rational preference P, lemma 1 defines the following binary relation.

Definition 8. *For any $a, b \in X$,*

$$a \vartriangleright_{p,P} b \qquad \text{if} \qquad p(c, A) > p(d, A) \text{ and } p(c, A \cup \{a\}) < p(d, A \cup \{a\})$$

for some $A \ni b$ and $d \in L^P(b)$.

Let $\overline{\vartriangleright}_{p,P}$ be the transitive closure of $\vartriangleright_{p,P}$. With a slight abuse of notation, $\overline{\vartriangleright}_{p,P}$ will also be used to denote the set of all pairs of alternatives in X that satisfy the binary relation $\overline{\vartriangleright}_{p,P}$. So $(a, b) \in \overline{\vartriangleright}_{p,P} \Leftrightarrow a \overline{\vartriangleright}_{p,P} b$. The following theorem then fully describes what choice data can reveal about the decision-maker's salience order.

Theorem 10. *Let p be a choice rule without ties that is rationalizable by OGPC and let P be the stochastically revealed preferred relation. Then a is revealed considered before b if and only if $a\overline{\rhd}_{p,P}b$.*

Theorems 7 and 10 uncover the underlying preference and salience order from choice data. All that remains is the order in which a particular alternative is compared to those it is stochastically revealed preferred to. Suppose a is stochastically revealed preferred to b and c. Consider a menu containing a, b and c with a worse choice rank for b compared to c. So b is eliminated in the choice procedure before c. Suppose further that such a menu contains no alternative that is strictly preferred to either b or c *and* is also more salient than a. This means that a must be doing the eliminating (in the choice procedure) of b and c in this menu. Such choice data, if rationalizable by OGPC, would require the comparison (a, b) to be made before (a, c).

Definition 9. *For a choice rule p, rationalizable by OGPC, (a, b) is* ***revealed compared before*** *(a, c) if for all $\mathbb{P}, \pi$ such that $p = p^{\mathbb{P},\pi}$,*

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

For a given preference $P \in \mathcal{P}$, let $U^P(a)$ be the set of all alternatives strictly preferred to a. For a choice rule without ties p, rationalizable by OGPC, let $\rhd_p$ be the revealed considered before relation and let $U^{\rhd_p}(a)$ be the set of all alternatives revealed considered before a.

Lemma 2. *Suppose p is a choice rule without ties, rationalizable by OGPC, with revealed preference relation P and $\{(a, b), (a, c)\} \subset P$. Then (a, b) is revealed compared before (a, c) if $\exists A \in \mathcal{X}$ such that*

$$\{a, b, c\} \subseteq A, \qquad p(b, A) < p(c, A) \qquad \text{and } \left[U^P(b) \cup U^P(c)\right] \cap U^{\rhd_p}(a) \cap A = \emptyset.$$

For a choice rule p rationalizable by OGPC, with revealed preference relation P, for any alternative $a \in X$, Lemma 2 defines the following binary relation.

Definition 10. *For any $b, c \in L^P(a)$,*

$$b \rhd^a_{p,P} c \qquad \text{if} \qquad \exists A \in \mathcal{X} \text{ such that}$$

$$\{a, b, c\} \subseteq A, \qquad p(b, A) < p(c, A) \qquad \text{and } \left[U^P(b) \cup U^P(c)\right] \cap U^{\rhd_p}(a) \cap A = \emptyset.$$

Let $\overline{\rhd}^a_{p,P}$ be the transitive closure of $\rhd^a_{p,P}$. The following theorem then fully characterizes what else choice data can reveal about the decision-maker's sequence of pairwise comparisons.

Theorem 11. *Let p be a choice rule without ties that is rationalizable by OGPC, with revealed preference relation P and $\{(a,b),(a,c)\} \subset P$. Then (a,b) is revealed compared before (a,c) if and only if $b\overline{\rhd}^a_{p,P}c$.*

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 even stricter requirement, called *Luce's IIA*,

$$\frac{p(a,A)}{p(b,A)} = \frac{p(a,B)}{p(b,B)}.$$

(O)GPC choice rules can violate strong stochastic transitivity (and therefore Luce's IIA too). Indeed they can violate the weaker consistency notion discussed in Tversky (1972a) and Natenzon (2017), 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 6. $a \succ b \succ c$ with

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

and $\pi(P_i, \cdot) = 1/3$ for $i \in \{1, 2, 3\}$.
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, $p^{\mathbb{P},\pi}(a, \{a, c\}) = 2/3 < \min\{p^{\mathbb{P},\pi}(a, \{a, b\}), p^{\mathbb{P},\pi}(b, \{b, c\})\} = 5/6$.

The rationale here is that for a GPCR the choice probability of the superior alternative in a pairwise comparison depends on the probability with which the decision-maker makes that comparison. The GPC procedure imposes very little structure on how these probabilities vary across menus. Interestingly, this allows for phenomenon like in the example above where 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.[8]

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 1. *sWARP implies weak stochastic transitivity.*

4.2 Similarity Effect

Debreu (1960) offers the following thought experiment as a critique of Luce's IIA. The universe of alternatives consists of three elements; D, a recording of the Debussy quartet, and B_F and B_K, both recordings of Beethoven's eighth symphony by the same orchestra but led by different conductors, F and K. The decision-maker chooses D with the same probability 3/5 when choosing from $\{D, B_F\}$ and $\{D, B_K\}$, revealing a clear preference for the Debussy recording. Furthermore she chooses B_F

[8]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.

and B_K with equal probability, 1/2, from the set $\{B_F, B_K\}$, revealing her indifference across the two Beethoven records. Such preferences suggest that the decision-maker would choose Debussy with the higher probability compared to Beethoven when all three records are available. Luce's IIA, however, would force the decision-maker to choose D with probability 3/7, requiring Beethoven to be selected with the higher probability. The principle that Luce's IIA runs counter to here 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.*"

Notice that Debreu argues for something stronger than the similarity effect, in that he expects the Debussy record to be chosen from the three element set with more than 1/2 probability. In this regard, observe that a minor variant of Debreu's example where the decision-maker picks D with probability 11/16 when choosing from $\{D, B_F\}$ and $\{D, B_K\}$, delivers the same critique, since Luce's IIA would force D to be chosen with probability 11/21 (less than half) with all three options available.

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

Example 7. (Debreu, Debussy and Beethoven)
$D \succ B_F \succ B_K$ with

P_0	P_1	P_2
	(D, B_F)	(B_F, B_K)
	(D, B_K)	

and $\pi(P_0, \cdot) = \alpha > 0$, $\pi(P_1, \cdot) = \beta > 0$ and $\pi(P_2, \cdot) = \epsilon > 0$ with $\alpha + \beta + \epsilon = 1$.

The resulting choice probabilities are as follows.

$p^{\mathbb{P},\pi}(\cdot, A)$	$A = \{D, B_F\}$	$\{D, B_K\}$	$\{B_F, B_K\}$	$\{D, B_F, B_K\}$
D	$\alpha/2 + \beta + \epsilon$	$\alpha/2 + \beta + \epsilon$	0	$\alpha/3 + \beta + \epsilon$
B_F	$\alpha/2$	0	$1/2 + \epsilon$	$\alpha/3$
B_K	0	$\alpha/2$	$1/2 - \epsilon$	$\alpha/3$

The first observation is that the odds ratio of choosing the Debussy record to a given Beethoven record increases with the addition of the second Beethoven record, since

$$\frac{\alpha/2 + \beta + \epsilon}{\alpha/2} < \frac{\alpha/3 + \beta + \epsilon}{\alpha/3}.$$

This is precisely the similarity effect and it holds irrespective of the particular values taken by the parameters α, β and ϵ.

The second observation is more subtle. Notice that Debussy is chosen with higher probability than Beethoven from the set of three (the stronger effect that Debreu expects) only if $\alpha/3 < \beta + \epsilon$. This does not hold if the parameters α, β and ϵ are selected to match $p^{\mathbb{P},\pi}(D, \{D, B_F\}) = p^{\mathbb{P},\pi}(D, \{D, B_K\}) = 3/5$ exactly as in Debreu's example. However, if instead α, β and ϵ are selected to match $p^{\mathbb{P},\pi}(D, \{D, B_F\}) = p^{\mathbb{P},\pi}(D, \{D, B_K\}) = 11/16$ as in the variant of Debreu's example that delivered the same critique, then indeed $\alpha/3 < \beta + \epsilon$. So while the similarity effect obtains in all cases, the even stronger effect that Debreu expects, is consistent with Debussy being selected in the pairwise choices with the higher probability of 11/16 but not the lower 3/5.

Debreu suggests that all that matters in his example is that the Debussy recording is selected with the higher probability against either of the Beethoven records and that the Beethoven recordings are almost identical. What is missing, however, is an explanation for why the Beethoven record is being chosen with probability 2/5. Why is it not an even lower probability, say 5/16 or even 0? With choice via GPC the probability of choosing the Beethoven record in pairwise choice matters, since it corresponds to the probability with which the decision-maker is unable to compare the two alternatives before making her choice. If this probability is high, then adding the second Beethoven record will lead to the two Beethoven records gaining two-thirds of this larger share of probability. This in turn may reduce the probability of Debussy being selected over Beethoven to less than half.

The final observation is that in example 7, the probability of stopping at a particular cell was held constant across menus. Making these menu dependent can easily rationalize the original Debreu example, as the following stopping function does.

$\pi'(\cdot, A)$	$A = \{D, B_F\}$	$\{D, B_K\}$	$\{B_F, B_K\}$	$\{D, B_F, B_K\}$
P_0	$4/5$	$4/5$	$4/5$	$3/5$
P_1	$1/5 - \epsilon$	$1/5 - \epsilon$	$1/5 - \epsilon$	$2/5 - \epsilon$
P_2	ϵ	ϵ	ϵ	ϵ

With ϵ taking a small enough value the probability of choosing either alternative from the set $\{B_F, B_K\}$ can be made arbitrarily close to 1/2. Further, now $p^{\mathbb{P},\pi'}(D, \{D, B_F\}) = p^{\mathbb{P},\pi'}(D, \{D, B_K\}) = p^{\mathbb{P},\pi'}(D, \{D, B_F, B_K\}) = 3/5$.

The next example shows how the similarity effect persists if Debreu's decision-maker has the opposite preference, in that she prefers either Beethoven recording to the Debussy.

Example 8. (Debreu, Beethoven and Debussy)
$B_F \succ B_K \succ D$ with

P_0	P_1	P_2
	(B_F, D)	(B_F, B_K)
	(B_K, D)	

and $\pi(P_0, \cdot) = \alpha > 0$, $\pi(P_1, \cdot) = \beta > 0$ and $\pi(P_2, \cdot) = \epsilon > 0$ with $\alpha + \beta + \epsilon = 1$. The resulting choice probabilities are as follows.

$p^{\mathbb{P},\pi}(\cdot, A)$	$A = \{D, B_F\}$	$\{D, B_K\}$	$\{B_F, B_K\}$	$\{D, B_F, B_K\}$
D	$\alpha/2$	$\alpha/2$	0	$\alpha/3$
B_F	$\alpha/2 + \beta + \epsilon$	0	$1/2 + \epsilon$	$\alpha/3 + \beta/2 + \epsilon$
B_K	0	$\alpha/2 + \beta + \epsilon$	$1/2 - \epsilon$	$\alpha/3 + \beta/2 - \epsilon$

For the similarity effect to hold the odds ratio of choosing a given Beethoven record to the Debussy record should decrease upon adding the other Beethoven record to the choice set. So the required inequality is

$$\frac{\alpha/2 + \beta + \epsilon}{\alpha/2} > \frac{\alpha/3 + \beta/2 + \epsilon}{\alpha/3},$$

which simplifies to $\beta/2 > \epsilon$. This inequality would hold for any given positive value of

β and all small enough ϵ, as is required to make the two Beethoven records (almost) identical.

4.3 Attraction and Compromise Effects

The regularity property, which requires

$$p(a, A) \geq p(a, B) \qquad \forall a \in A \subseteq B,$$

is neither implied by nor implies any of the notions of stochastic transitivity discussed earlier. While always satisfied in the large class of random utility models, it is often violated in choice data. The attraction and compromise effects (Huber et al. (1982) and Simonson (1989)) are two kinds of such violation that have been extensively documented in the marketing literature. In both, the addition of a new alternative (often labeled decoy) leads to an already available alternative (target) being selected with a higher probability. The GPC choice procedure can generate both these effects in a natural way and are discussed in turn.

Attraction Effect. The effect is illustrated in figure 1. The decision-maker chooses from among alternatives defined by a pair of attributes. She prefers more to less of each attribute. The probability with which she selects alternative b in the presence of a is higher if c, which is strictly dominated by b but not by a, is also available; $p(b, \{a, b, c\}) > p(b, \{a, b\})$.

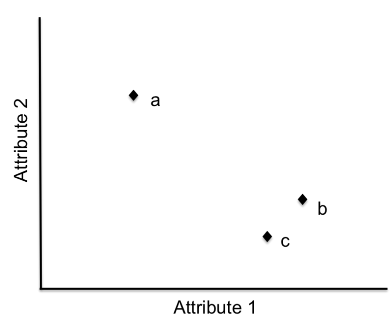

Figure 1: Attraction Effect

The literature clarifies that the key property of the decoy (c) is the asymmetry in its pairwise relation with the two other options, clearly dominated by b but not by a. Notice that there is another closely related asymmetry here. Precisely because of the dominance relation, b and c are a lot easier to compare than any other pair. So if the true preference of the decision-maker is $a \succ b \succ c$ and if she makes the pairwise comparisons gradually, it is natural that she makes the comparison (b, c) before either (a, b) or (a, c). Consider the following order of comparison

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

If the decision-maker were to stop with equal odds after each element of her partition relevant to the decision problem, an acute attraction effect appears. In particular, with $\pi(P_1, \{a, b, c\}) = \pi(P_2, \{a, b, c\}) = 1/2$ and $\pi(P_2, \{a, b\}) = 1$, the resulting choice data has $p^{\mathbb{P},\pi}(b, \{a, b\}) = 0$ while $p^{\mathbb{P},\pi}(b, \{a, b, c\}) = 1/4$.

Note that the use of $\pi(P_2, \{a, b\}) = 1$ is not necessary to obtain the attraction effect. Indeed, given the non-obvious nature of the comparison (a, b) the decision-maker may "give up" before making this comparison even when choosing from $\{a, b\}$. This can be modelled by a stopping function π' with $\pi'(P_0, \{a, b\}) > 0$. Then, as long as $\pi'(P_0, \{a, b, c\})/3 + \pi'(P_1, \{a, b, c\})/2 > \pi'(P_0, \{a, b\})/2$, it is true that $p^{\mathbb{P},\pi'}(b, \{a, b\}) < p^{\mathbb{P},\pi'}(b, \{a, b, c\})$. So in a sense, as long as having to make an additional comparison (even if it is an easy one) only increases the chance of giving up before making a tougher comparison later, the attraction effect persists.

Compromise Effect. The effect is illustrated in figure 2. The alternatives are again defined by a pair of attributes and the decision-maker prefers more to less of each attribute. She selects alternative c with a higher probability in the presence of a or b if the other (of a and b) is also available; $p(c, \{a, b, c\}) > p(c, \{a, c\}), p(c, \{b, c\})$.

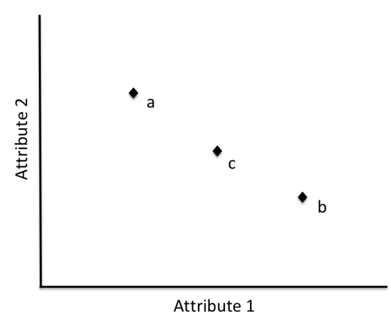

Figure 2: Compromise Effect

Notice that none of the alternatives dominate any other. In this case, the decision-maker's preference depends on how much of a tradeoff she is willing to make across the two attributes. It seems natural then that this comparison is easier to make if the gains and losses involved are substantial, as is the case when comparing a and b. Consider then the following sequence of comparison of the rational preference $a \succ b \succ c$,

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

Further, like in example 1b, she is equally likely to stop at only relevant elements of her partition, meaning $\pi(P_1, \{a, b\}) = \pi(P_2, \{a, c\}) = \pi(P_3, \{b, c\}) = 1$ and $\pi(P_i, \{a, b, c\}) = 1/3$ for $i \in \{1, 2, 3\}$. It then follows that $p^{\mathbb{P},\pi}(c, \{a, c\}) = p^{\mathbb{P},\pi}(c, \{b, c\}) = 0 < 1/6 = p^{\mathbb{P},\pi}(c, \{a, b, c\})$, thereby generating the compromise effect.

In the explanation above, the decision-maker does not have any intrinsic preference for compromise. Indeed, the compromise alternative is the least preferred of the three. Nevertheless, it is more obvious that one of the extreme alternatives is not the most preferred of the three. As long as with positive probability the decision-maker gives up after making this comparison, the compromise effect obtains.

5 Discussion

5.1 Random Attention Models

Some recent work has focused on a different source of bounded rationality. A strict subset of all available alternatives may catch the decision-maker's attention when she makes her choice. 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.

In Manzini and Mariotti (2014) an alternative in a given choice set catches the agent's attention with a probability that is independent of the choice set itself. Brady and Rehbeck (2016) allow for correlation of this random attention across alternatives by considering more general distributions over consideration sets (sets that catch the agent's attention) for a given choice set. Cattaneo et al. (2018) consider an even more general framework, requiring only that the probability with which an agent focuses on a given consideration set (conditional on it being feasible) cannot increase if the underlying choice set becomes larger.

In terms of choice data it can rationalize, none of the random attention models nest GPC and vice versa. It is easy to see this with the model in Manzini and Mariotti (2014) which must satisfy regularity (unlike a GPCR) but allows the violation of sWARP. The latter of course remains true for the more general models of Brady and Rehbeck (2016) and Cattaneo et al. (2018). The following example is of a GPCR (with a simple SPCR counterpart) that violates the acyclicity condition that characterizes the random attention model in Cattaneo et al. (2018).

Example 9. (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)

Like in example 1b, the decision-maker is equally likely to stop only at those elements of the ordered partition that are relevant to her choice problem. In particular,

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

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 then $p^{\mathbb{P},\pi}(b, \{a,b,d\}) = 0$ and $p^{\mathbb{P},\pi}(b, \{a,b,c,d\}) = 1/15$, which would mean that b is revealed preferred to c, violating acyclicity.

Perhaps more importantly, the random attention model and gradual pairwise comparison constitute fundamentally different choice procedures. This difference 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.[9] Take for instance example 1b, 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 $c \succ b$.

Finally, intuition suggests that the relative importance of GPC and random attention should depend on the size of the choice set. It is hard to imagine random attention not being of first order importance in large choice sets. It is less convincing with choice sets of, for instance, 5 or fewer alternatives. In the latter, it seems more plausible that while the agent is fully aware of all alternatives in her choice set, the differential ease of pairwise comparability among these leads to menu-dependent choice ranks. A simple model that combines both these sources of bounded rationality remains to be worked out.

5.2 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 $p(x, A) = \frac{u(x)}{\sum_{y \in A} u(y)}$. 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 $u(x) > u(y)$.[10] 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:

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

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 second lowest value under u is the inferior alternative, and so on. Let π be such that

[9] Masatlioglu, Nakajima and Ozbay (2012) and Dutta and Horan (2015) make a similar argument against model-free approaches to welfare analysis suggested in Bernheim and Rangel (2009) for deterministic choice settings.

[10] This property is implied by the Acyclicity condition in Fudenberg, Iijima and Strzalecki(2015). Therefore the construction can be used to rationalize any choice rule without ties that has an Additive Perturbed Utility representation.

$\pi(P_i, \cdot) > 0$ for all $i \leq n$. It then follows that

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

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.[11]

Finally, if the aim is to match the choice probabilities to the Luce rule exactly, the following stopping rule π (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 set $\pi(P_{n-q+1}, A) = |A| \frac{u(x_q)}{\sum_{x_j \in A} u(x_j)}$ where x_q is the alternative with the worst choice rank in A.

Gul, Natenzon and Pesendorfer (2014) study a generalization of the Luce model in which the attributes of an alternative, often shared across alternatives, play a key role in determining choice probabilities. The reliance on the richer environment of attributes makes the model difficult to compare with the GPC procedure. Nevertheless, such attribute rules do not contain all GPCRs since they must satisfy regularity. The same is true for the model of Additive Perturbed Utility studied in Fudenberg, Iijima and Strzalecki (2015).

The sequential nature of the GPC procedure may suggest similarity with the Perception Adjusted Luce Model (PALM) in Echenique, Saito and Tserenjigmid (2018). The two procedures, however, are very different both in structure and choice data it can rationalize. For the latter note that without an outside option, PALM is the same as a Luce rule. Pairwise comparisons play no procedural role in PALM. The procedure in PALM involves the decision-maker considering available alternatives in a sequence (perception priority), with options later in the sequence being chosen (with probabil-

[11] 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.

ity satisfying a Luce formula) only conditional on those before not being chosen. As an example of the implications of the procedural difference, in PALM the event of not choosing the first option in the perception priority contains the event of not choosing the second option in the perception priority. By contrast, in the GPC procedure, for any pair of alternatives a, b whether the event of not choosing a contains the event of not choosing b or the other way around can be menu dependent.

The satisficing procedure studied in Aquiar, Boccardi and Dean (2016) is another example of a choice procedure with a sequential aspect (search order). Beyond the obvious differences in the procedures themselves, the empirical content of the Full Support Satisficing Model (FSSM) does not nest that of GPC. Example 1b, for instance, violates the axiom SARP (necessary for FSSM), which rules out a getting chosen with positive probability from a set in which c not chosen at all, if there exists another set in which c is selected with positive probability in the presence of a.

Interestingly, the model of Bayesian probit studied in Natenzon (2018) despite a fundamentally different approach, shares a number of features with GPC. In addition to rationalizing decoy effects, it captures the idea that some pairwise comparisons are more difficult than others. Further, the difference in the true value of the options in the pairwise comparison may only have a partial influence on the ease of making the comparison. In Bayesian probit, for each alternative in a choice set the agent observes a noisy signal of its utility. Allowing the signals to be correlated lets an additional alternative in a choice set bring new information about the relative ranking of existing alternatives. This new information allows for menu-dependent choice ranks. By contrast, in a GPC rule it is the different orders of pairwise comparison with different alternatives as the superior option that generate menu-dependent choice ranks.

A Appendix

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 $j \leq i$. □

Lemma 4. *If* $a \in M^{\mathbb{P}}_{\bar{I}^{\mathbb{P}}(A)}(A)$ *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 $M^{\mathbb{P}}_{\tilde{I}^{\mathbb{P}}(A)}(A)$ is a singleton, $d \notin M^{\mathbb{P}}_{\tilde{I}^{\mathbb{P}}(A)}(A) \Rightarrow (c,d) \in P$ for some $c \in A$. This contradicts d being P-maximal in A. $\square$

Proof of Theorem 5*. *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 $M^{\mathbb{P}}_{\tilde{I}^{\mathbb{P}}(A)}(A)$ 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 $M^{\mathbb{P}}_{\tilde{I}^{\mathbb{P}}(B)}(B) = \{b\}$. Again by lemma 4 this implies that $(b,a) \in P$, a contradiction.
s-reducibility: By Theorem 3, $p^{\mathbb{P},\pi}$ 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

$$p(c,A) > p(y,A), \qquad \forall c \in A \setminus \{y\}.$$

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 $\mathcal{E}_i = \{A \in \mathcal{C}_i | \{a_i,b_i\} \subseteq A\}$.

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 π be a stopping function on $\mathbb{P}$ such that $\pi(P_i,\cdot) > 0$ for all $1 \le i \le I$. 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 $j \le i \le I$. 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 $1 \le i \le n$, $(x_i,y_i) \in P_{j(i)}$, $y_i = x_{i+1}, \forall i < n$ 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 $1 \le i \le n$ 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 $\{i | a \in M_i^{\mathbb{P}}(A)\}$ and $\{i | b \in M_i^{\mathbb{P}}(A)\}$, must be a strict subset of the other. Now since $\pi(P_i,\cdot) > 0$ for all $1 \le i \le I$ 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 π' 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 $\{x, a\} \subseteq B$,

$$p(c, B) > p(a, B), \qquad \forall c \in B \setminus \{a\}.$$

Then by ITR it must be that if $B \in A^T(p)$

$$p(c, A) > p(a, A), \qquad \forall c \in B \setminus \{a\}.$$

This contradicts the assertion that b is the lowest ranked among all alternatives in B under $p(\cdot, A)$. □

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 theorem 4. As for the former, it has already been shown in the proof of theorem 5 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 $a, b \in X$

$$p'(a, A) > p'(b, A) \Rightarrow p(a, A) \geq p(b, A). \tag{4}$$

Since p' is an SCR without ties and satisfies sWARP, ITR and s-reducibility, by theorem 5 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 4, by theorem 2, there exists π' such that $p = p^{\mathbb{P},\pi'}$. □

Proof of Theorem 7*.* " $\Rightarrow$ " : If $(a, b) \in P$ then $p^{\mathbb{P},\pi}(a, \{a, b\}) > p^{\mathbb{P},\pi}(b, \{a, b\})$ since by assumption $\pi(P_{\tilde{I}^{\mathbb{P}}(A)}, A) > 0$ for all $A \in \mathcal{X}$. Therefore a is stochastically revealed preferred to b.
" $\Leftarrow$ " : 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 $M^{\mathbb{P}}_{\tilde{I}^{\mathbb{P}}(A)}(A) = \{a\}$. The result then follows directly from lemma 4. □

Lemma 5. *Let p be a stochastic choice rule (SCR) without ties that satisfies strong s-reducibility. Then given a collection of sets $\mathcal{B} \subseteq \tilde{\mathcal{X}}$ if $a \in D \in \mathcal{B}$ satisfies requirements (i) and (ii) of strong s-reducibility then there exists $E \in \mathcal{B}$ with $\{a, b\} \subseteq E$ such that if $\{a, b\} \subseteq A \in \mathcal{B}$ then*

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

Proof. Fix a collection of sets $\mathcal{B} \subseteq \tilde{\mathcal{X}}$ and let $a \in D \in \mathcal{B}$ satisfy requirements (i) and (ii) of strong s-reducibility. Suppose there exists $X_i \in \mathcal{B}$ such that $a \in X_i$ and $Ł^P(a) \cap X_i \neq \emptyset$. Then by requirement (i) of strong s-reducibility, the lowest ranked alternative in X_i, say x_i, must be an element of $Ł^P(a)$. Now either, $p(c, A) >$

$p(x_i, A)$ for all $c \in A \setminus \{x_i\}$ whenever $\{a, x_i\} \subseteq A \in \mathcal{B}$ or there exists $X_{i+1} \in \mathcal{B}$ and $x_{i+1} \in L^P(a)$ such that $\{a, x_i, x_{i+1}\} \subseteq X_{i+1}$ and $p(c, X_{i+1}) > p(x_{i+1}, X_{i+1})$ for all $c \in X_{i+1} \setminus \{x_{i+1}\}$. Notice that x_{i+1} is outranked in the presence of a by x_i.

Since p satisfies strong s-reducibility and $L^P(a) \cap D \neq \emptyset$, there must exist $x \in L^P(a) \cap D$ such that $p(c, D) > p(x, D)$ for all $c \in D \setminus \{x\}$. Set $x_1 = x$. Then the argument above generates a sequence, $\{x_i\}_{i=1}^n$ that must be finite. This is because $L^P(a)$ is finite and $x_j \neq x_k$ where $j < k$. Indeed, $x_j = x_k$ with $j < k$ would violate requirement (ii) of strong s-reducibility. $j < k$ implies that for all $i > j$, x_i is outranked in the presence of a by x_j. In particular x_{k-1} is outranked in the presence of a by x_j. But if $x_k = x_j$ then it must be that $p(x_j, X_k) < p(x_{k-1}, X_k)$ with $\{a, x_j, x_{k-1}\} \subseteq X_k$, thereby violating requirement (ii) of strong s-reducibility.

The last element of this finite sequence, x_n, by construction satisfies $p(c, A) > p(x_n, A)$ for all $c \in A \setminus \{x_n\}$ whenever $\{a, x_n\} \subseteq A \in \mathcal{B}$.

□

Lemma 6. *If p is an SCR without ties that satisfies strong s-reducibility then it satisfies s-reducibility.*

Proof. Follows directly from lemma 5. □

Lemma 7. *Given a collection of sets $\mathcal{B} \subseteq \tilde{\mathcal{X}}$, if c is outranked in the presence of a by b, then for a collection of sets $\mathcal{B}'$ such that $\mathcal{B} \subseteq \mathcal{B}' \subseteq \tilde{\mathcal{X}}$ it is still true that c is outranked in the presence of a by b.*

Proof. Follows directly from definition 6. □

Proof of Theorem 8. Necessity:
It is sufficient to show that if $p^{\mathbb{P},\pi}$ is an OGPCR without ties then it satisfies sWARP, ITR and strong s-reducibility. Since an OGPCR is also a GPCR, it has already been shown in theorem 5 that it satisfies sWARP and ITR.
strong s-reducibility: By theorem 3, $p^{\mathbb{P},\pi}$ 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 $B \in \mathcal{B}$. Let m be the highest number no smaller than k such that $(a, \cdot) \in P_m$. Since $p^{\mathbb{P},\pi}$ is an OGPCR and an SPCR, for all $k \leq i \leq m$, $P_i = \{(a, y_i)\}$ for some $y_i \in X$. By theorem 7, $(a, y_i) \in P \Leftrightarrow y_i \in L^P(a)$.

Now let $A \in \mathcal{B}$ such that $a \in A$. By the selection of P_k it must be that $M^{\mathbb{P}}_{k-1}(A) = A$. Further if $y \in L^P(a) \cap A$ then $(a, y) \in P_i$ for some $k \leq i \leq m$. Therefore $M^{\mathbb{P}}_m(A) = A \setminus L^P(a)$. This implies $\{i | d \in M^{\mathbb{P}}_i(A)\} \subset \{i | c \in M^{\mathbb{P}}_i(A)\}$ for all $c \in A \setminus L^P(a)$ and $d \in L^P(a) \cap A$. Therefore, since $p^{\mathbb{P},\pi}$ is an OGPCR without ties, $p^{\mathbb{P},\pi}(c, A) > p^{\mathbb{P},\pi}(d, A)$, for all $c \in A \setminus L^P(a)$ and $d \in L^P(a) \cap A$. This proves part (i).

Suppose that given $\mathcal{B}$, c is outranked in the presence of a by b. So there exists a sequence of sets $\{X_i\}_{i=1}^n$ and alternatives $\{x_i\}_{i=1}^{n+1} \subset L^P(a)$ such that for any $1 \leq i \leq n$,

$$\{a, x_i, x_{i+1}\} \subseteq X_i \in \mathcal{B} \qquad \textit{and} \qquad p^{\mathbb{P},\pi}(x_i, X_i) > p^{\mathbb{P},\pi}(x_{i+1}, X_i)$$

with $x_1 = b$ and $x_{n+1} = c$. As shown earlier, $X_i \in \mathcal{B} \Rightarrow M^{\mathbb{P}}_{k-1}(X_i) = X_i$. Further $\{x_i, x_{i+1}\} \subseteq L^P(a)$ implies that $(a, x_i) \in P_{j(i)}$ and $(a, x_{i+1}) \in P_{j(i+1)}$ for some $k \leq j(i), j(i+1) \leq m$. Finally, $p^{\mathbb{P},\pi}(x_i, X_i) > p^{\mathbb{P},\pi}(x_{i+1}, X_i)$ implies that $j(i+1) < j(i)$. Therefore it must be that $j(1) < j(n+1)$, implying $\{i | b \in M^{\mathbb{P}}_i(A)\} \subset \{i | c \in M^{\mathbb{P}}_i(A)\}$ for any $A \in \mathcal{B}$ which in turn implies $p^{\mathbb{P},\pi}(b, A) > p^{\mathbb{P},\pi}(c, A)$ for any $A \in \mathcal{B}$. This proves part (ii).

Sufficiency: Let p be a stochastic choice rule without ties that satisfies sWARP, ITR and strong 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 strong s-reducibility, there exists $a \in D \in \mathcal{B}$ with $L^P(a) \cap D \neq \emptyset$ that satsfies requirements (i) and (ii) of strong s-reducibility. There may exist multiple a's that satisfy this condition. If $P_{i-1} = (w, z)$ and w satisfies the condition above then set $a = w$. Otherwise, simply pick one of the a's.

By lemmas 5 and 6, not only does p also satisfy $s - \textit{reducibility}$ but given the alternative a identified above there exists an alternative y such that $\{a, y\} \subseteq E$ for some $E \in \mathcal{B}$ and if $\{a, y\} \subseteq A \in \mathcal{B}$ then

$$p(c, A) > p(y, A), \qquad \forall c \in A \setminus \{y\}.$$

$\{a, y\}$ could satisfy this condition for multiple values of y. Simply pick one, say y_i and set $P_i = (a, y_i)$. Next, let $\mathcal{C}_{i+1} = \mathcal{C}_i \setminus \mathcal{E}_i$ where $\mathcal{E}_i = \{A \in \mathcal{C}_i | \{a, y_i\} \subseteq A\}$.

Setting $\mathcal{C}_1 = \tilde{\mathcal{X}}$ and using the construction above yields a sequence $\{P_i\}_{i=1}^I$. Let $\mathbb{P} = \{P_i\}_{i=1}^I$. Also let π be a stopping function on $\mathbb{P}$ such that $\pi(P_i, \cdot) > 0$ for all $1 \leq i \leq I$.

The construction above is a special case of the construction used in the proof of theorem 5. As a result, the fact that $\mathbb{P}$, so defined, is a partition of a strict rational preference and that by construction $p^{\mathbb{P},\pi}$ is an SCR without ties, follow from the same arguments as in the proof of theorem 5 (sufficiency part), since p satisfies sWARP and s-reducibility (implied by strong s- reducibility). Also, exactly the same arguments as in the proof of theorem 5 (sufficiency part) shows that for all $A \in \mathcal{X}$ and $a, b \in A$,

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

It does need to be shown that $p^{\mathbb{P},\pi}$ by the construction above is an OGPCR, in that $(w, x) \in P_m$ and $(y, z) \in P_n$ with $m < n, w \neq y$ and $(w, r) \in P_k$, $(y, s) \in P_l$ requires $k < l$. Suppose this is true for all P_j with $j \leq i$. Then it is sufficient to show that if $P_i = \{(a, b)\}$ and a is stochastically revealed preferred to x but $P_j \neq \{(a, x)\}$ for all $j \leq i$ then $P_{i+1} \neq \{(g, h)\}$ where $g \neq a$. Given the construction above, it would then be sufficient to show that given $\mathcal{C}_{i+1}$, the alternative a satisfies requirements (i) and (ii) of strong reducibility. This would ensure that the construction sets $P_{i+1} = (a, y)$ for some $y \in X$.

Now, $P_i = \{(a, b)\}$ implies that given $\mathcal{C}_i$, there exists $D \in \mathcal{C}_i$ with $\{a, b\} \subseteq D$ such that (i) if $a \in A \in \mathcal{C}_i$ then

$$p(c, A) > p(d, A) \qquad \forall c \in A \setminus L^P(a), d \in L^P(a) \cap A, \qquad \textit{and}$$

(ii) given $\mathcal{C}_i$ and for all $b, c \in L^P(a)$, if c is outranked in the presence of a by b then

$$p(b, A) > p(c, A) \qquad \forall A \in \mathcal{B} \textit{ with } \{a, b, c\} \subseteq A$$

By the premise above there exists $x \in X$ such that a is stochastically revealed preferred to x but $P_j \neq \{(a, x)\}$ for any $j \leq i$. By the construction of $\mathcal{C}_{i+1}$ this means that there must exist $E \in \mathcal{C}_{i+1}$ with $\{a, x\} \subseteq E$. Next, since $\mathcal{C}_{i+1} \subset \mathcal{C}_i$, by (i) above it follows that if $a \in A \in \mathcal{C}_{i+1}$ then

$$p(c, A) > p(d, A) \qquad \forall c \in A \setminus L^P(a), d \in L^P(a) \cap A.$$

Again since $\mathcal{C}_{i+1} \subset \mathcal{C}_i$, if for some $b, c \in L^P(a)$

$$p(b, A) > p(c, A) \qquad \forall A \in \mathcal{C}_i \textit{ with } \{a, b, c\} \subseteq A$$

then if $\{a,b,c\} \subseteq A \in \mathcal{C}_{i+1}$ it must still be that $p(b,A) > p(c,A)$. Finally by lemma 7, if c is *not* outranked in the presence of a by b in $\mathcal{C}_i$ then c is still *not* outranked in the presence of a by b in $\mathcal{C}_{i+1}$, as $\mathcal{C}_{i+1} \subset \mathcal{C}_i$. This proves that a does indeed satisfy requirements (i) and (ii) of strong s-reducibility, given $\mathcal{C}_{i+1}$. □

Proof of Theorem 9. The proof is almost identical to that of theorem 6. □

Proof of Lemma 1. Let $p^{\mathbb{P},\pi}$ be an OGPC rule such that $p = p^{\mathbb{P},\pi}$, and $A \ni b$ be such that $p(c,A) > p(d,A)$ for some $d \in L^P(b)$. Now suppose that $(b,x) \in P_i, (a,y) \in P_j \Rightarrow i < j$. This means that a is not revealed considered before b. Let k be the smallest number such that $(a,\cdot) \in P_k$. Then $d \notin M^{\mathbb{P}}_{k-1}(A)$. This is because $(b,d) \in P_i$ for some $i < k$. So if, $b \in M^{\mathbb{P}}_{i-1}(A)$ then $d \notin M^{\mathbb{P}}_i(A)$. If $b \notin M^{\mathbb{P}}_{i-1}(A)$ then there exists some y such that $(y,b) \in P_j$ for some $j < i$. Let x be the first such y in the OGPC procedure. Then it must be that $(x,d) \in P_q$ for some $q < i$, in which case again $d \notin M^{\mathbb{P}}_i(A)$ and therefore $d \notin M^{\mathbb{P}}_{k-1}(A)$.

Now, $(a,y) \notin P_j$ for $j < k$ and any $y \in X$ implies that if $c \neq a$ then

$$c \in M^{\mathbb{P}}_j(B) \Leftrightarrow c \in M^{\mathbb{P}}_j(B \cup \{a\}) \qquad \forall j < k. \tag{5}$$

Then,

$$\begin{aligned}
& p(c,A) > p(d,A) \\
& \Rightarrow p^{\mathbb{P},\pi}(c,A) > p^{\mathbb{P},\pi}(d,A) \\
& \Rightarrow \{i | c \in M^{\mathbb{P}}_i(A)\} \supset \{i | d \in M^{\mathbb{P}}_i(A)\} && \textit{by } (2) \\
& \Rightarrow \{i \leq k-1 | c \in M^{\mathbb{P}}_i(A)\} \supset \{i \leq k-1 | d \in M^{\mathbb{P}}_i(A)\} && (\textit{ since } d \notin M^{\mathbb{P}}_{k-1}(A)) \\
& \Rightarrow \{i \leq k-1 | c \in M^{\mathbb{P}}_i(A \cup \{a\})\} \supset \{i \leq k-1 | d \in M^{\mathbb{P}}_i(A \cup \{a\})\} && \textit{by } (5) \\
& \Rightarrow \{i | c \in M^{\mathbb{P}}_i(A \cup \{a\})\} \supset \{i | d \in M^{\mathbb{P}}_i(A \cup \{a\})\} \\
& \Rightarrow p^{\mathbb{P},\pi}(c, A \cup \{a\}) \geq p^{\mathbb{P},\pi}(d, A \cup \{a\}) \\
& \Rightarrow p(c, A \cup \{a\}) \geq p(d, A \cup \{a\}).
\end{aligned}$$

□

Proof of Theorem 10. " $\Leftarrow$ " follows immediately from lemma 1 and the fact that the *revealed considered before* relation is transitive.

To prove " $\Rightarrow$ " the following claim will be proven first. Suppose $p^{\mathbb{P},\pi}$ is an OGPC rule such that $p^{\mathbb{P},\pi} = p$. Further let a be the alternative immediately preceding b in the salience order under $\mathbb{P}$ but $(a, b) \notin \overline{\rhd}_{p,P}$. Suppose $p^{\mathbb{P}',\pi'}$ is an OGPC rule with the only difference between $\mathbb{P}'$ and $\mathbb{P}$ being that the salience order of a and b is reversed. π' defined on $\mathbb{P}'$ is such that $\pi'(P_i, \cdot) > 0$ for all $P_i \in \mathbb{P}'$. The claim is that

$$p^{\mathbb{P},\pi}(c, A) > p^{\mathbb{P},\pi}(d, A) \Leftrightarrow p^{\mathbb{P}',\pi'}(c, A) > p^{\mathbb{P}',\pi'}(d, A), \qquad \forall A \in \tilde{\mathcal{X}}, \forall c, d \in A.$$

If the claim is true then by theorem 2 there exists π'' such that $p = p^{\mathbb{P}',\pi''}$. This would prove that a is not revealed considered before b. Without loss of generality, (by theorem 3), both $p^{\mathbb{P}',\pi'}$ and $p^{\mathbb{P},\pi}$ are assumed to be SPCRs.

Suppose P_j (P_m) and $P_k(P_n)$ are the first and last cells in $\mathbb{P}$ with a (b) as the superior outcome in the pairwise comparison they contain. By assumption $j < k = m - 1 < m < n$. Similarly let $P'_{m'}$ ($P'_{j'}$) and $P'_{n'}(P'_{k'})$ be the first and last cells in $\mathbb{P}$ with b (a) as the superior outcome in the pairwise comparison they contain. Of course, now $m' < n' = j' - 1 < j' < k'$. Further by construction, $j = m'$ and $n = k'$. So,

$$\begin{aligned} P'_i &= P_i && \textit{for all } i < j \textit{ and } i > n \\ P'_i &= P_{i+m-j} && \textit{if } j \le i \le j + n - m \\ P'_i &= P_{i-(n-m)-1} && \textit{if } j + n - m < i \le n. \end{aligned}$$

Fix $A \in \tilde{\mathcal{X}}$. Let $c, d \in A$. If $\{c, d\} \cap L^P(b) = \emptyset$ or if $\{a, b\} \not\subset A$ then

$$\{i | c \in M_i^{\mathbb{P}}(A)\} \supset \{i | d \in M_i^{\mathbb{P}}(A)\} \Leftrightarrow \{i | c \in M_i^{\mathbb{P}'}(A)\} \supset \{i | d \in M_i^{\mathbb{P}'}(A)\}$$

implying $p^{\mathbb{P},\pi}(c, A) > p^{\mathbb{P},\pi}(d, A) \Leftrightarrow p^{\mathbb{P}',\pi'}(c, A) > p^{\mathbb{P}',\pi'}(d, A)$.

Now suppose $\{a, b, c, d\} \subset A \in \tilde{\mathcal{X}}$ with $\{c, d\} \cap L^P(b) = \{d\}$. Then it cannot be that $p^{\mathbb{P}',\pi'}(c, A) < p^{\mathbb{P}',\pi'}(d, A)$ and $p^{\mathbb{P},\pi}(c, A) > p^{\mathbb{P},\pi}(d, A)$. To see why, note that since $d \in L^P(b)$, $d \notin M_{n'}^{\mathbb{P}'}(A)$. Since $c \notin L^P(b)$ if $c \in M_{j-1}^{\mathbb{P}'}(A)$ then $c \in M_{n'}^{\mathbb{P}'}(A)$. Therefore $p^{\mathbb{P}',\pi'}(c, A) < p^{\mathbb{P}',\pi'}(d, A)$ requires that $c \notin M_{j-1}^{\mathbb{P}'}(A)$. Since $P'_i = P_i$ *for all* $i < j$, it must be that $\{i < j | c \in M_i^{\mathbb{P}'}(A)\} \subset \{i < j | d \in M_i^{\mathbb{P}'}(A)\} \Leftrightarrow \{i < j | c \in M_i^{\mathbb{P}}(A)\} \subset \{i < j | d \in M_i^{\mathbb{P}}(A)\}$, which in turn means $p^{\mathbb{P},\pi}(c, A) < p^{\mathbb{P},\pi}(d, A)$.

The only remaining possibility is that $\{a, b, c, d\} \subset A \in \tilde{\mathcal{X}}$, $d \in L^P(b)$ (c may or may not be in $L^P(b)$) with $p^{\mathbb{P},\pi}(c, A) < p^{\mathbb{P},\pi}(d, A)$ and $p^{\mathbb{P}',\pi'}(c, A) > p^{\mathbb{P}',\pi'}(d, A)$. First note that if $\{c, d\} \cap M_{j-1}^{\mathbb{P}}(A) \neq \{c, d\}$ then such a reversal cannot happen since

$P'_i = P_i$ *for all* $i < j$. So with $\{c, d\} \subseteq M^{\mathbb{P}}_{j-1}(A) = M^{\mathbb{P}'}_{j-1}(A)$ and $d \in L^P(b)$, the inequality $p^{\mathbb{P}',\pi'}(c, A) > p^{\mathbb{P}',\pi'}(d, A)$ implies that there exists $j \leq q \leq j + n - m$ such that $\{c, d\} \cap M^{\mathbb{P}'}_q(A) = \{c\}$. Then there exists $m \leq r \leq n$ such that $\{c, d\} \cap M^{\mathbb{P}}_r(A \setminus \{a\}) = \{c\}$ since $\{c, d\} \subseteq M^{\mathbb{P}}_{j-1}(A) \Rightarrow \{c, d\} \subseteq M^{\mathbb{P}}_{j-1}(A \setminus \{a\})$. This implies $p^{\mathbb{P},\pi}(c, A \setminus \{a\}) > p^{\mathbb{P},\pi}(d, A \setminus \{a\})$. But then since $(a, b) \notin \rhd_{p,P}$ it cannot be that $p^{\mathbb{P},\pi}(c, A) < p^{\mathbb{P},\pi}(d, A)$. This concludes the proof for the initial claim.

It has therefore been shown that if $(a, b) \notin \overline{\rhd}_{p,P}$ and if a and b hold adjacent positions in the salience order of $\mathbb{P}$, with a preceding b, where $p^{\mathbb{P},\pi} = p$ then switching the order of a and b and leaving all else unchanged generates the same choice ranks. Suppose now that a still precedes b but with other alternatives in between in the salience order. Then take the alternative, c nearest to b in the salience order, which lies in between a and b such that $(a, c) \in \overline{\rhd}_{p,P}$. Since $(a, b) \notin \overline{\rhd}_{p,P}$ it must be that $(c, d) \notin \overline{\rhd}_{p,P}$ for all alternatives d in the salience order between c and b. The procedure outlined above then shows that c can be shifted one adjacent switch at a time to eventually take a position after b in the salience order, without affecting any choice ranks. This same procedure can be carried out for the alternatives from c all the way to a without changing the salience order of the alternatives between and including a and c. This would eventually lead to a following b in the salience order and yet generating the same choice ranks as before. This concludes the proof.

□

Proof of Lemma 2. Suppose $p = p^{\mathbb{P},\pi}$ where $p^{\mathbb{P},\pi}$ is an OGPC rule without ties that is also an SPCR (without loss of generality by theorem 3). Let $P_k(P_l)$ be the first (last) cell of the ordered partition $\mathbb{P}$ to contain (a, y) for some $y \in X$. So for all $i < k$ and $i > l$ $P_i \neq \{(a, w)\}$ for any $w \in X$. Since $\left[U^P(b) \cup U^P(c)\right] \cap U^{\rhd_p}(a) \cap A = \emptyset$, theorem 10 implies that $\{a, b, c\} \subseteq M^{\mathbb{P}}_{k-1}(A)$. Further, $\{(a, b), (a, c)\} \subset P$ implies that $\{b, c\} \cap M^{\mathbb{P}}_l(A) = \emptyset$. Therefore to satisfy $p^{\mathbb{P},\pi}(b, A) < p^{\mathbb{P},\pi}(c, A)$ it must be that if $P_i = (a, b)$ and $P_j = (a, c)$ then $i < j$. □

Proof of Theorem 11. " $\Leftarrow$ " follows immediately from lemma 2 and the fact that the *revealed compared before* relation is transitive.
To prove " $\Rightarrow$ " the following claim will be proven first. Suppose $p^{\mathbb{P},\pi}$ is an OGPC rule without ties that is also an SPCR such that $p^{\mathbb{P},\pi} = p$. Further let (a, b) be the pairwise comparison immediately preceding (a, c) in $\mathbb{P}$ but $(b, c) \notin \overline{\rhd}^a_{p,P}$. Suppose $p^{\mathbb{P}',\pi'}$ is an OGPC rule with the only difference between $\mathbb{P}'$ and $\mathbb{P}$ being that the order of (a, b)

and (a,c) is reversed (therefore also an SPCR). So if $P_k = (a,b)$ and $P_{k+1} = (a,c)$ then $P'_i = P_i$ for all $i \notin \{k, k+1\}$, $P'_k = (a,c)$ and $P'_{k+1} = (a,b)$. π' defined on $\mathbb{P}'$ is such that $\pi'(P_i, \cdot) > 0$ for all $P_i \in \mathbb{P}'$. The claim is that

$$p^{\mathbb{P},\pi}(x,A) > p^{\mathbb{P},\pi}(y,A) \Leftrightarrow p^{\mathbb{P}',\pi'}(x,A) > p^{\mathbb{P}',\pi'}(y,A), \qquad \forall A \in \tilde{\mathcal{X}}, \forall x,y \in A.$$

If the claim is true then by theorem 2 there exists π'' such that $p = p^{\mathbb{P}',\pi''}$. This would prove that (a,b) is not revealed compared before (a,c).

Fix $A \in \tilde{\mathcal{X}}$. Let $x \in A$ such that $x \notin \{b,c\}$. Then notice that $x \in M_i^{\mathbb{P}}(A) \Leftrightarrow x \in M_i^{\mathbb{P}'}(A)$ for all $1 \leq i \leq I$. Moreover $x \in M_{k+1}^{\mathbb{P}}(A) \Leftrightarrow x \in M_{k-1}^{\mathbb{P}}(A)$. If $x \in \{b,c\}$ then $x \in M_i^{\mathbb{P}}(A) \Leftrightarrow x \in M_i^{\mathbb{P}'}(A)$ for all $1 \leq i \leq I$ such that $i \neq k$. Therefore for any $x,y \in X$ such that $\{x,y\} \neq \{b,c\}$ it must be that $p^{\mathbb{P},\pi}(x,A) > p^{\mathbb{P},\pi}(y,A) \Leftrightarrow p^{\mathbb{P}',\pi'}(x,A) > p^{\mathbb{P}',\pi'}(y,A)$.

Now suppose by contradiction, there exists some $A \in \tilde{\mathcal{X}}$ such that $p^{\mathbb{P},\pi}(b,A) < p^{\mathbb{P},\pi}(c,A)$ but $p^{\mathbb{P}',\pi'}(b,A) > p^{\mathbb{P}',\pi'}(c,A)$. Then clearly $\{a,b,c\} \subseteq M_{k-1}^{\mathbb{P}}(A)$, which implies that $\left[U^P(b) \cup U^P(c)\right] \cap U^{\rhd_p}(a) \cap A = \emptyset$. But that would imply $b \rhd_{p,P}^a c$, a contradiction.

It has therefore been shown that if $(b,c) \notin \overline{\rhd}_{p,P}^a$ and if (a,b) and (a,c) hold adjacent positions in $\mathbb{P}$, with (a,b) preceding (a,c), where $p^{\mathbb{P},\pi} = p$ then switching the order of (a,b) and (a,c) and leaving all else unchanged generates the same choice ranks. Suppose now that (a,b) still precedes (a,c) but with other comparisons in between in $\mathbb{P}$. Then take the comparison, (a,d) nearest to (a,c) along the order in $\mathbb{P}$, which lies in between (a,b) and (a,c) such that $(b,d) \in \overline{\rhd}_{p,P}^a$. Since $(b,c) \notin \overline{\rhd}_{p,P}^a$ it must be that $(d,e) \notin \overline{\rhd}_{p,P}^a$ for all alternatives e where (a,e) lies in between (a,d) and (a,c) in $\mathbb{P}$. The procedure outlined above then shows that (a,d) can be shifted one adjacent switch at a time to eventually take a position after (a,c) in the ordered partition of P, without affecting any choice ranks. This same procedure can be carried out for the comparisons from (a,d) all the way to (a,b) without changing the order of the comparisons between and including (a,b) and (a,d). This would eventually lead to (a,b) following (a,c) in the ordered partition and yet generating the same choice ranks as before. This concludes the proof. □

Proof of Observation 1. 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. □

References

[1] Aguiar, Victor H., Maria Jose Boccardi and Mark Dean. 2016. "Satisficing and Stochastic Choice." *Journal of Economic Theory* 166 445-482.

[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] Echenique, Federico, Kota Saito and Gerelt Tserenjigmid. 2018. "The Perception Adjusted Luce Model." *Mathematical Social Sciences* 93 67-76.

[11] Fudenberg, Drew, Ryota Iijima and Tomasz Strzalecki. 2015. "Stochastic Choice and Revealed Perturbed Utility." *Econometrica* 83(6) 2371-2409.

[12] Gul, Faruk, Paulo Natenzon and Wolfgang Pesendorfer. 2014. "Random Choice as Behavioral Optimization." *Econometrica* 82(5) 1873-1912.

[13] Horan, Sean. 2018. "Threshold Luce Rules." working paper.

[14] Huber, Joel, J. W. Payne and C. Puto. 1982. "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis." *Journal of Consumer Research* 9(1) 90-98.

[15] Luce, Robert D. 1959. "Individual Choice Behavior: A Theoretical Analysis." *New York: Wiley.*

[16] Manzini, Paola, and Marco Mariotti. 2012. "Choice by Lexicographic Semiorders." *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] 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.

[21] Russo, J. Edward and Larry D. Rosen. 1975. "An Eye Fixation Analysis of Multialternative Choice." *Memory and Cognition* 3(3) 267-276.

[22] Simonson, Itamar. 1989. "Choice Based on Reasons: The Case of Attraction and Compromise Effects." *The Journal of Consumer Research* 16(2) 158-174.

[23] Tversky, Amos. 1972a. "Choice by Elimination." *Journal of Mathematical Psychology* 9 341-367.

[24] Tversky, Amos. 1972b. "Elimination by Aspects: A Theory of Choice." *Psychological Review* 79(4) 281-299.

Récents cahiers de recherche du CIREQ
Recent Working Papers of CIREQ

La liste complète des cahiers de recherche du CIREQ est disponible à l'adresse suivante : http://www.cireqmontreal.com/cahiers-de-recherche .

The complete list of CIREQ's working papers is available at the following address : http://www.cireqmontreal.com/cahiers-de-recherche .

07-2018 Bengui, J., T. Phan, "Asset Pledgeability and Endogenously Leveraged Bubbles", juin 2018, 48 pages

08-2018 Andersson, T., L. Ehlers, L.-G. Svensson, R. Tierney, "Gale's Fixed Tax for Exchanging Houses", juin 2018, 22 pages

09-2018 Ramachandran, R., C. Rauh, "Discrimination Without Taste - How Discrimination Can Spillover and Persist", juin 2018, 38 pages

10-2018 Gonçalves, S., B. Perron, "Bootstrapping Factor Models With Cross Sectional Dependence", juillet 2018, 30 pages

11-2018 Andersson, T., Á. Cseh, L, Ehlers, A. Erlanson, "Organizing Time Banks : Lessons from Matching Markets", juillet 2018, 24 pages

12-2018 Hoelzemann, J., N. Klein, "Bandits in the Lab", août 2018, 52 pages

13-2018 Klein, N., P. Wagner, "Strategic Investment and Learning With Private Information", août 2018, 24 pages

14-2018 Godefroy, R., "The Birth of the Congressional Clinic", août 2018, 47 pages

15-2018 Godefroy, R., J. Lewis, "The Impact of Uncertainty in Agriculture", août 2018, 40 pages

16-2018 Horan, S., P. Manzini, M. Mariotti, "Precision May Harm : The Comparative Statics of Imprecise Judgement", août 2018, 32 pages

17-2018 Bouakez, H., O. Rachedi, E. Santoro, "Sectoral Heterogeneity, Production Networks, and the Effects of Government Spending", août 2018, 44 pages

18-2018 Benchekroun, H., M. Breton, A. Ray Chaudhuri, "Mergers in Nonrenewable Resource Oligopolies and Environmental Policies", septembre 2018, 40 pages

19-2018 Dudley, L., C. Rauh, "Innovation Growth Clusters : Lessons from the Industrial Revolution", août 2018, 32 pages

20-2018 Amodio, F., L. Baccini, M. DiMaio, "Security, Trade, and Political Violence", septembre 2018, 58 pages

21-2018 Sprumont, Y., "Belief-weighted Nash Aggregation of Savage Preferences", septembre 2018, 27 pages

22-2018 Andersson, T., L. Ehlers, A. Martinello, "Dynamic Refugee Matching", octobre 2018, 36 pages