Essays in Decision Theory

A dissertation submitted

by

Bhavook Bhardwaj

to

The Indian Statistical Institute

in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

in

Quantitative Economics

Advisor: Prof. Debasis Mishra

May 2024

Abstract

This thesis contains three chapters on individual decision-making and choice. The first chapter introduces a general model of decision-making where alternatives are sequentially examined by a decision maker. Our main object of study is a *decision rule* that maps infinite sequences of alternatives to a decision space. Within the class of decision rules, we focus on two natural subclasses: *stopping* and *uniform stopping* rules. Our main result establishes an equivalence between these two subclasses. Next, we introduce the notion of computability of decision rules using Turing machines and show that computable rules can be implemented using a simpler computational device: a finite automaton. We further show that computability of *choice rules* —a subclass of decision rules—is implied by their continuity with respect to a natural topology. Finally, we provide a revealed preference "toolkit" and characterize some natural choice procedures in our framework.

The second chapter introduces a model of decision-making that formalizes the idea of rejection behavior using binary relations. We propose a procedure where a decision maker *rejects* the minimal alternatives from any decision problem. We provide an axiomatic foundation of this procedure and introduce a shortlisting model of choice where this procedure leads to a new type of a consideration set mapping: the *rejection filter*. We study the testable implications of this shortlisting model using observed reversals in choice. Next, we relate our findings to the existing literature and show that our model provides a novel explanation of some empirically observed behavior. Finally, we introduce and characterize a simple two-stage model of stochastic choice using rejection filters.

The third chapter studies studies the *Copeland set*, a popular tournament solution, from a revealed preference perspective. Two choice procedures where a decision maker has a fixed underlying tournament are introduced and behaviorally characterized: (i) a deterministic choice rule that selects for every menu, the Copeland set of the tournament restricted to that menu; and (ii) a stochastic choice rule that assigns to every menu, a probability distribution over it in a "Luce" manner, where the Luce "weight" of each alternative is generated using its the Copeland score in that menu.

Acknowledgments

From sheer confusion, to getting comfortable with and to finally being able to appreciate and enjoy mathematical arguments in particular and economic theory in general, it has been a long journey. During this journey, I have had the good fortune to come across bright minds and honest individuals dedicated to their craft and that has been a constant source of inspiration for me. Three people stand out. The first one is my advisor, Debasis Mishra. His clarity of thought and insights have shaped my way of thinking about research problems. He has been always welcoming and has given me the academic freedom to explore my areas of interest. The second one is Arunava Sen. His breadth of knowledge, depth of understanding of the subject and most importantly, passion for research is infectious and a source of constant motivation. His friendly, welcoming nature and a sharp mind —a rare combination —makes one want to repeatedly engage and learn from him.

The third one is my colleague, friend and co-author, Siddharth Chatterjee. Some time around two years ago, by a fortunate accident, we struck a conversation and since then there has been no looking back. I have learnt a lot from him in terms of thinking about research, importance of methods in economic theory as well as developing a certain academic discipline. His immersion into the field never fails to inspire me and keeps me charged up. I have had some of the most profound moments of my PhD during our endless discussions that often resulted in “aha!” moments of discovery.

In my academic life, I have had the privilege to interact with some inspiring personalities that include Ariel Rubinstein, Hervé Moulin, Jean Drèze, Pradeep Dubey, Badal Mukherji and Kalyan Chatterjee. Albeit of short duration, some of these interactions have had a lasting impact on me in terms of striving for excellence in research. In particular, my interactions with Ariel Rubinstein have had a profound impact on the way I think about economic theory (and our shared love for working out of cafes!). I would also like to thank Sean Horan for providing invaluable guidance during the initial phase of my PhD journey and Josué Ortega for an enriching collaborative experience while I was a visiting researcher at Queen’s University in Belfast.

While at ISI, I made some good friends. Three of them deserve a special mention. First, Kriti Manocha, with whom I had my initial forays into choice theory and our long discussions made the research process enjoyable. We ended up becoming coauthors. Second, Kolagani Paramahamsa, from whom I learnt cooking, Yoga as well as some mechanism design. His way of conducting life inspired me on many levels. Third, Moyuree Mukherjee, who provided much needed emotional support during difficult times in the PhD. Further, we ended up becoming partners in working out of cafes. Other people I am thankful to from ISI are Vilok Taori, Prerna Dewan, Abhay Gupta, Ankit Singh, Rishabh Gupta, Ojasvi Khare, Nikita Sangwan, Ravinder, Hemant Mishra, Partha Pratim Ghosh and Ahana Bashista.

Beyond ISI, different sets of people have contributed in their unique ways to my development both as an academician as well as an individual. From school, Tejas Tilak, Sagar Gambhir and Shreyak Chauhan....from undergraduate college, Disha Bajaj, Utkarsh Goyal, Sagar Bhatia, Himanshu Khurana and Vaibhav Bhardwaj...from masters college, Aastha Malhotra, Arshiya Sibia, Gargi Pal, Roshni Tara, Tanmeet Thukral and Vedanta Dhamija...from my working days as a research assistant, Arpita Mukherjee, Sanjay Pulipaka, Suvi Agrawal, Stuti Manchanda, Tanu Goyal and Sahil Gandhi...and from my football group, Manish Payal, Naveen Gusain, Vedant Rana, Ishwar Singh (Podo), Kaushal Jora, Nikhil Gahlot and Arjun Kalra.

Lastly, I am grateful for my family. My parents, whom I thank for my unconventional upbringing, instilling emotional values in me and for understanding the value of giving freedom in order to develop a child's individuality. My brother, whom I thank for being a protective and a loving figure all along. My sister-in-law who has always been caring and supportive. Finally, my two bundles of joy, my niece and nephew who never fail to bring a smile to my face.

CONTENTS

Introduction

Individual decision-making forms the basis of much of economic theory. While the classical models of decision-making such as the rational choice model have served as a foundation to economic analysis, there has been mounting empirical evidence against these models, both in a laboratory setting and outside of it. To address this problem, the literature has seen emergence of new models of decision-making that fall under the umbrella term of "bounded rationality" and use insights from related fields such as psychology. This thesis presents and analyzes three such models of decision-making.

The first chapter (co-authored with Siddharth Chatterjee) studies decision-making over *infinite* sequences of alternatives. A general model is developed where a decision maker (DM) is represented by a *decision rule*, a function that maps infinite sequences of alternatives to an arbitrary decision space. This model enriches the classical setup of choice over *sets* of alternatives to capture order effects in decision-making. Further, it also enriches the setup of choice over finite *lists* of Rubinstein and Salant (2006) by enabling the analysis of fully endogenous stopping behavior, a study of which is precluded by the finite nature of lists. The notion of endogenous stopping is embodied in a subclass of decision rules that we term *stopping rules*. For every sequence, these require a stopping point $k \in \mathbb{N}$, beyond which the contents of the sequence are "irrelevant" for the decision-making process. Since the domain of stopping rules is infinite, the set of all stopping points is not guaranteed to have a finite upper bound. The stopping rules for which the set of all stopping points is bounded are called *uniform stopping rules*. While it is clear that by definition uniform stopping rules are stopping rules, the main result of this chapter, the *reduction lemma*, shows that the converse is also true. A proof via diagonalization and a proof via topological methods is given for this result.

Two implications of the reduction lemma are explored in this chapter. First, the perspective of computability of decision-making is adopted. Computable and simply computable rules are defined using the models of *Turing machines* and *finite automatons* respectively. While in the theory of computation Turing machines and finite automatons represent distinct models of computation in terms of their computational capabilities, in the setup of infinite sequences, the reduction lemma establishes an equivalence between these two models. That is, any decision rule that can be implemented using a Turing machine can be implemented using a finite automaton and vice-versa.

Second, a revealed preference approach is considered in this model. Many natural choice procedures are subsumed within the class of stopping rules. Due to the reduction lemma, they are subsumed within the class of uniform stopping rules as well. With a finite set of alternatives, this enables the *testability* of specific choice procedures within the class of stopping rules. A revealed preference "toolkit" is developed to formulate testable *axioms* in this setup. To demonstrate its usefulness, two choice procedures capturing the idea of *satisficing* (see Simon (1955)) are formulated and axiomatically characterized.

The final part of this chapter drops the assumption of endogenous stopping behavior. It focuses on settings where decision-making involves selecting from a collection of infinite streams of alternatives. This is captured in a class of decision rules called *configuration dependent rules* where the DM views every infinite sequence as a collection of configurations: one for each alternative. While configuration dependent rules subsume a wide range of behaviors, a specific way of choosing is by maximizing a preference relation over the set of all configurations. These rules, called rational configuration dependent rules, are stated formally and characterized using a neutrality and an acyclicity axiom.

The second chapter (co-authored with Kriti Manocha) introduces a model of decision-making that formalizes the idea of *rejection* behavior using binary relations. The standard models of decision-making in economic theory do not distinguish between *choice* and *rejection*. However, evidence from the literature on experimental psychology suggests otherwise. When a decision problem is framed as one of "rejection" than that of "choice", individuals tend to behave differently. In particular, selections are larger in case of rejection problems vis-à-vis choice problems. This chapter proposes a procedure that explains this behavioral difference.

While operating in the standard setup where a DM is represented by a binary relation, the act of rejection is captured by the elimination of *minimal* alternatives as against selection of the *maximal* ones for every decision problem. This generates a consideration set mapping that is termed as a *rejection filter*. An axiomatic foundation of rejection filters is provided and contrasted with other filters in the literature such as the attention filter of Masatlioglu et al. (2012) and the competition filter of Lleras et al. (2017).

A two-stage shortlisting method à la Manzini and Mariotti (2007) is introduced where the first stage involves shortlisting using a rejection filter and the second stage involves maximization of a complete binary relation on the shortlisted set to generate a unique choice. The choice function thus generated is termed *Choice by Rejection* (CBR) and a behavioral characterization of it is provided. The analysis of CBR follows that of Horan (2016) by identifying certain "reversals" in observed choice that form the basis of a behavioral characterization that uses four conditions on the choice function. Two conditions are with respect to a revealed relation that is defined using these reversals: (i) a congruence condition à la Tyson (2013) and Richter (1966) and; (ii) an acyclicity condition on the revealed relation. The other two are weakenings of standard consistency conditions used in choice theory.

The issue of identification of the underlying binary relations is addressed next. It is observed that the underlying relations are not identified uniquely. Again, the reversals defined using observed choices are useful in providing insight into the "common" parts to every representation of a CBR choice function. A "small menu" property is shown to hold which states that if two CBR choice functions coincide on menus of size 2 and 3, then they must be identical.

In the final part of this chapter, a connection of CBR with the rational shortlist method (RSM) of Manzini and Mariotti (2007) is studied. Within the class of CBR choice functions, the RSM choice functions are shown to be the ones that do not display a specific reversal. Finally, a simple model of *stochastic* choice is studied that uses a rejection filter to shortlist alternatives before the final probability distribution is assigned. That is, the minimal alternatives get zero probabilities in every menu. This is a special case of the general Luce model of Echenique and Saito (2019). A characterization of this model follows by suitably adapting the axioms characterizing the rejection filter to a stochastic setup and the cyclical independence axiom of Echenique and Saito (2019).

The third chapter studies decision-making using tournaments, binary relations that are asymmetric and complete. Tournaments are well studied objects in economic theory and arise in both in the contexts of individual and group decision-making. While the rational choice model assumes transitivity of the underlying binary relation, there is ample empirical evidence of cyclical choices. Choices via tournaments allow for cyclical choices and this chapter studies two such procedures.

In the first procedure, the DM chooses for every menu of alternatives, the Copeland set of an underlying *fixed* tournament restricted to that menu. Copeland set is a popular tournament solution that selects the set of alternatives that have the highest number of pairwise "wins", also known as Copeland scores. The choice correspondences thus generated are termed *Copeland choice rules.* An axiomatic characterization of Copeland choice rules is provided using five axioms. Two new axioms are used in the characterization: (i) *Responsiveness* which is a choice theoretic adaptation of the *positive responsiveness* axiom of Rubinstein (1980); (ii) *Symmetry* that requires the chosen alternatives for any menu to be treated symmetrically upon addition of a "dominated" or a "dominating" alternative. The third axiom is a standard axiom called *Binary Dominance Consistency* (BDC). The fourth axiom is a weakening of the contraction consistency (condition α) axiom of Sen (1971) called *Independence of Dominating Alternatives*(IDA) and the fifth axiom is a weakening of the Weakened WARP(WWARP) axiom of Ehlers and Sprumont (2008) called *Weak* WWARP.

In the second procedure, the DM chooses randomly in every menu using the *Copeland scores* of alternatives in the tournament restricted to that menu. The primitives are a tournament and a scoring function that assigns a positive real number —the "score" —to every alternative using its Copeland score. Then, the probability assigned to an alternative in a menu is equal to its relative score in that menu. Two variants of this procedure, one with the option of abstaining in every menu and the other without, are axiomatically characterized.

In the last part of this chapter, the relation of Copeland rules with top cycle rules (Ehlers and Sprumont (2008)) and uncovered set rules (Lombardi (2008)) is examined. As a tournament solution, the Copeland set is a subset of the uncovered set which is a subset of the top cycle set. Top cycle rules are characterized by WWARP, BDC and Weak Contraction Consistency (WCC). It is shown the Copeland rules satisfy the latter two but may violate WWARP. Uncovered set rules are characterized by Weak Expansion (WE), Non-Discrimination (ND), Weakened Chernoff (WC) and BDC. It is shown that Copeland rules satisfy the latter three but may violate WE.

Chapter 1

Decisions over Sequences

1.1 Introduction

1.1.1 *An overview*

Imagine a decision maker (DM) who faces alternatives sequentially before *she* decides to "stop" and make a decision. The alternatives keep on being presented and the decision to stop may depend on the history of alternatives examined thus far. A wide range of situations where decision-making involves seeking recommendations, receiving bitstreams of information, meeting people, viewing alternatives on websites etc. fall under such a description. There is substantial evidence from the literature on marketing and psychology that the order in which alternatives are presented to the DM affects the final decision. We propose a general model to study such situations where the ordering of alternatives can affect decision-making. Our objects of interest are *decision rules* that map *infinite* sequences of alternatives to a decision space.

The classical model of choice studies choice functions over *sets* of alternatives. To incorporate the well-documented order effects in decision-making, Rubinstein and Salant (2006) initiated the literature on choice functions over finite *lists*. Our model is a natural generalization of theirs since (i) it allows the lists to be infinite (and allow for duplication) and; (ii) it allows the choice to land in an arbitrary decision space making choice from within the set of alternatives a special case. However, there is an important conceptual difference. In our model, we allow the decision to stop to be completely endogenous to the DM. That is, the DM can go on examining the alternatives as long as *she* "wishes" to in the above mentioned situations of decision-making. For instance, in seeking recommendations regarding which restaurant to go to, the *DM* decides when to stop receiving recommendations. This is in contrast with model of choice over lists. Despite its ability to study endogenous stopping behavior, the finite nature of lists has an exogeneity built into it i.e. the last term in a list is the exogenous stopping point. Given this conceptual difference, we investigate whether the behavioral predictions of our model are different from that of the model on lists. We find that from the perspective of *computability* of decision-making, the two are significantly different.[1]

The idea that a DM endogenously stops and makes a decision is captured in a broad subclass of decision rules that we term *stopping* rules. These decision rules require, for every infinite sequence, a "point" in the sequence beyond which the terms in the sequence are irrelevant in the decision-making process. However, this "stopping point" may possibly be dependent on the sequence as the following example illustrates:

Example 1 (Cardinal satisficing). Let X be a finite set of movies and $\{X_i\}_{i=1}^N$ denote a partition of the set of movies into N "genres" for some $N \in \mathbb{N}$. A DM wishes to watch a movie and relies on recommendations. She attaches a "weight" to each genre which indicates the value she attaches to each genre i.e. there exists a function $w : N \to \mathbb{R}_+$ such that every movie in the i^{th} genre is given the same weight. Her decision procedure involves seeking recommendations *sequentially* from different sources such as peer groups, websites etc. She has a "threshold" weight in her mind and for every sequence of recommendations, she selects the first movie whose cumulative weight (due to repetitions) crosses the threshold weight.

Despite the set of alternatives being finite, the set of inputs to a stopping rule is uncountably infinite and hence, the stopping points are not guaranteed to have a finite upper bound. A subclass of stopping rules with a finite bound on the set of all stopping points are called *uniform stopping* rules. Such rules have an attractive feature that when a decision rule restricted to be from the class of uniform stopping rules it, loosely put, turns the decision problem to a finite one. That is, only a finite set of "segments" are relevant for decision-making. This finite nature of the uniform stopping rules enables one to formulate and study *testable* implications of various "choice heuristics".

[1] Another instance of a conceptual difference leading to differing predictions is in game theory. In the case of repeated games, as one switches from the assumption of finite repetition to infinite repetition of the stage game, the predictions can differ significantly. In particular, cooperation can be sustained as an equilibrium outcome in the repeated prisoner's dilemma.

The first question we ask is what restrictions on the class of stopping rules pin down the subclass of uniform stopping rules? Our main result —the *reduction lemma* —shows that there are none. That is, these two subclasses of decision rules are in fact equivalent. We establish this equivalence using a diagonalization argument and a critical ingredient is the assumption that the set of alternatives generating the sequences is finite. We provide a generalized version as well as a topological approach to our main result in the Appendix. We observe that the reduction lemma is essentially a result concerning stopping times as defined in probability theory. It shows that any stopping time that is inputwise bounded has a finite bound. Further, the generalization of the reduction lemma extends it to a sequence of stopping times.

Next, we define the notion of *computability* of decision rules via Turing machines. We call a decision rule computable if it can be implemented via a Turing machine. According to the Church-Turing thesis, any physically realizable computer can be expressed as a Turing machine. On the other hand, we define the notion of *simple computability* of decision rules via finite automatons. In the theory of computation, finite automatons are simpler models of computation than Turing machines. That is, a computation that can be performed by a finite automaton can be performed by a Turing machine as well. But the converse is not true. However, in our setup, using the reduction lemma, we show that any computable decision rule is also simply computable—an equivalence that fails to hold in the list setup.

In the second part of the paper, we focus our attention on another subclass of decision rules that we term *choice* rules. These are natural analogues of choice functions over menus and choice functions over lists in our setting. That is, choice rules are those decision rules that output for every sequence, an element of the sequence. Our first result shows that within the subclass of choice rules, computable rules are characterized by their continuity when the domain (set of all sequences) and the co-domain (set of alternatives) are endowed with product and discrete topology respectively.

With the assumption of computability in the background and our main result, any choice rule can be fully specified by its output on a *finite* set of intitial "segments" of sequences. These segments are what we call *minimal sufficient* segments. Using a mathematically equivalent but conceptually different object than a stopping rule—which we term a *decision procedure*—we are able to fully identify these segments. This provides us with a "toolkit" for conducting revealed preference analysis of choice procedures. We then introduce and behaviorally characterize two procedures: *cardinal* satisficing and *ordinal* satisficing. These are natural adaptations of the idea of *satisficing* introduced by Simon (1955) in our setup.

While much of our analysis concerns stopping rules, our framework can be used to model non-stopping behavior as well. With a different viewpoint to decision rules, behavior such as one involving choosing from infinite streams can be studied. To that end, we introduce a subclass of choice rules that we term *configuration-dependent* rules. Such rules are useful to study situations where the positions as well as the frequency of alternatives in a sequence can affect decision-making. As the name suggests, the only relevant information for these rules is the "configuration" of an alternative in a sequence that is represented by a 0-1 sequence. We characterize a subclass of such rules which we term *rational* configuration dependent rules, where the DM chooses by maximizing a ranking over these configurations.

1.1.2 *Some examples and outline*

To illustrate the applicability of our model in varied settings, we provide some examples of decision rules.

Example 2 (Ordinal Satisficing). A DM has to choose a partner based on repeated interactions with a finite set of potential partners. She has a fixed attention span $k \in \mathbb{N}$ i.e. she engages in only first k interactions for any sequence. She has a preference order[2] over the set of potential partners and has some "threshold" partner in her mind. She chooses the first potential partner within the first-k interactions that is ranked above the threshold. Otherwise she chooses the highest ranked partner according to her preference from the first k interactions.

Example 3 (Two-stage status quo choice). Let X be a finite set of alternatives. A DM relies on two objects to make a decision: a shortlisting function Γ and an preference order $\succ$ on X. For any sequence, the *status quo* is its first element and it dictates the shortlist generated by the Γ function —the consideration *set*—from the elements of the sequence. Having generated a consideration set, the DM chooses its $\succ$-maximal set.

Example 4 (Investment strategy). A company wishes to invest in a fund based on inputs provided by a fund rating agency. Let X denote the finite set of possible performances of the fund (often denoted as A_+, B_{++} etc.). The inputs are in the form of a sequence of performances of the fund i.e. at every time period $t \in \mathbb{N}$, the agency

[2] A reflexive, antisymmetric, complete and transitive binary relation.

announces the performance of the fund. Let $Y = \mathbb{R} \times \mathbb{N}$. The company uses an "algorithm" that —based on the input of a sequence performances—decides the amount of money and the time period for which to invest i.e. $y \in Y$.

Example 5 (Stochastic attention). Let X be a set of alternatives and $Y = \Delta X$ be the set of all probability distributions over X. The DM is endowed with preference order $\succ$ over X and a probability distribution over $\mathbb{N}$ denoted by σ, that captures her variable "attention". For every sequence, the DM first draws the attention parameter $k \in \mathbb{N}$ using σ and then chooses the $\succ$-maximal alternatives from the k-long segment of the sequence.

Example 6 (Language processing). Let $X = \{0, 1\}$ and $Y = \{\text{TRUE}, \text{FALSE}\}$. Consider a computer program that receives bitstreams or sequences of symbols from X. The bitstreams represent expressions in a natural language (for instance, English) encoded in binary expression i.e. $0's$ and $1's$. For every input bitstream, the program declares it as "TRUE" if it contains a grammatically correct sentence. It outputs "FALSE" otherwise.

Example 7 (Rational choice). Let X be a finite set of alternatives and $Y = X$. The DM is endowed with a preference order $\succ$ over the set X. For any sequence, she picks the $\succ$-maximal alternatives in that sequence.

The outline of rest of the paper is as follows. In Section 1.2, we introduce the setup and prove our main result. Section 1.3 discusses computability of decision rules using Turing machines and finite automatons. In Section 1.4, we introduce choice rules and provide a revealed preference "toolkit" to study choice procedures. Section 1.5 introduces and characterizes two choice procedures of satisficing. In Section 1.6, we introduce configuration dependent rules and characterize rational configuration dependent rules. Section 1.7 discusses some related literature and Section 1.8 concludes. The proofs omitted in the main body of the paper can be found in the Appendix (Section 1.9).

1.2 Main Result

1.2.1 *Preliminaries*

Let X be a non-empty finite set of alternatives. A *sequence* is a map $S : \mathbb{N} \to X$, where $\mathbb{N}$ denotes the set of natural numbers. By $X^{\mathbb{N}}$ we denote the collection of all X-valued sequences. That is, $X^{\mathbb{N}} := \{S \mid S : \mathbb{N} \to X\}$. The term $S(i)$ corresponds to the

i^{th} entry of the sequence S. A *segment* is any map $M : [k] \to X$, where $[k] \coloneqq \{1, \ldots, k\}$ for some $k \in \mathbb{N}$. Let the set of all segments of length k be denoted by $\mathcal{S}_k$ and the set of all segments be denoted by $\mathcal{S}$. Consider any subset E of natural numbers and a sequence S. We define the *restriction* of S to E as the map $S|_E : E \to X$ where $[S|_E](i) = S(i)$ for all $i \in E$. When $E = [k]$ for some $k \in \mathbb{N}$, the segment $S|_{[k]}$ is called the *truncation* of S at k. We will abuse notation and write $S|_k$ instead of $S|_{[k]}$ whenever no confusion arises. For any $S, T \in X^{\mathbb{N}}$ and $k \in \mathbb{N}$, we define the *concatenation* of the segment $S|_k$ and the sequence T to be the sequence $S|_k \cdot T \in X^{\mathbb{N}}$ such that $[S|_k \cdot T](i) = S(i)$ for all $i \in [k]$ and $[S|_k \cdot T](i) = T(i-k)$ for all $i \in \{k+1, \ldots\}$. Concatenation of two segments is defined in a similar manner.

We denote the set of all *decisions* by a non-empty set Y. In particular, Y can be equal to X, a restriction we will impose in Section 4. However, as we can see from the examples in the previous section, Y can be distinct from X. The DM in our model is represented with a decision rule, d, which gives a unique decision for every infinite sequence. Formally, it is defined as follows.

Definition 1. *A decision rule on sequences is any map $d : X^{\mathbb{N}} \to Y$.*

A decision rule is a very general object as the list of examples provided in Section 1.1.2 shows. However, throughout most part of this paper (except for Section 6), we restrict our attention to a subclass of decision rules that capture the notion of endogenous stopping. Recall that in Example 1, for every sequence of alternatives, the cumulative weights as well as the threshold dictate the "stopping" point. Further, since each alternative has a positive weight, the DM stops for *every* sequence. We call such rules *stopping* rules and they are formally defined as follows.

Definition 2. *A decision rule d is a stopping rule if for all $S \in X^{\mathbb{N}}$, there exists a $k \in \mathbb{N}$ such that for all $T \in X^{\mathbb{N}}$,*

$$d(S) = d(S|_k \cdot T)$$

Stopping rules capture the idea that for any given sequence, the DM does not wait indefinitely and "makes up its mind" by a finite amount of time i.e. after viewing a finite initial segment and the subsequent alternatives of the sequence do not affect the decision. To show that not every decision rule is a stopping rule, consider the following simple example. Let $X = \{x^*, y, z\}$ and $Y = \{0, 1\}$. The decision rule d is defined as $d(S) = 1$ if $x^* = S(i)$ for some $i \in \mathbb{N}$ and $d(S) = 0$ otherwise. Consider any sequence

S that does not feature x^* in it. It can be observed that for any $k \in \mathbb{N}$, we can find a $T \in X^{\mathbb{N}}$ that features x^* in it such that the decision for the concatenation of $S|_k$ and T is different than that for S and therefore d is not a stopping rule. If our interpretation of a decision rule is that the sequence is examined by the DM sequentially —in discrete time for instance —then such a decision rule looks implausible since for the sequences that do not feature x^*, the DM will never stop and would have to wait "forever" to make a decision.

It is important to note the stopping point or the "relevant" finite segment for stopping rules can depend on the sequence. Since the set of sequences is infinite, the lengths of these relevant segments are not guaranteed to have a finite upper bound. A subclass of stopping rules for which the these lengths have a finite upper bound are called uniform stopping rules.

Definition 3. *A decision rule d is a uniform stopping rule if there exists a $k \in \mathbb{N}$ such that for all $S, T \in X^{\mathbb{N}}$,*

$$d(S) = d(S|_k \cdot T)$$

While stopping rules require for every sequence, the existence of a finite bound on the "consideration" of the DM, uniform stopping rules require a fixed finite bound on the consideration for *every* sequence. That is, there is a change in the order of quantifiers in the definition of the two subclasses of decision rules. The following is a simple example of a uniform stopping rule: The DM is endowed with a preference order $\succ$ over X, and for any sequence, she considers only the first 10 alternatives if the first element of the sequence is some designated $x^* \in X$ and picks the $\succ$-maximal alternative from them. Otherwise, she looks at the first 20 alternatives and picks the $\succ$-maximal alternative from them.

1.2.2 *The reduction lemma*

Our main result establishes the equivalence of stopping and uniform stopping rules. Before stating and proving the result, we first provide an alternative definition of stopping rules using the following useful object which is defined for any decision rule d.

$$k_d(S) \coloneqq \inf \left\{ k \in \mathbb{N} : d(S) = d(S|_k \cdot T) \text{ for all } T \in X^{\mathbb{N}} \right\}$$

The function $k_d(\cdot)$ is the *stopping time* for the sequence S and captures the smallest truncation of a sequence S beyond which the terms of the sequence do not affect decisions. Using k_d, we redefine stopping and uniform stopping rules as follows (with the convention that $n < \infty$ for all $n \in \mathbb{N}$).

DEFINITION 4. *A decision rule d is a*

(i) Stopping rule if $k_d(S) < \infty$ for every $S \in X^{\mathbb{N}}$.

(ii) Uniform stopping rule if $\sup\{k_d(S) : S \in X^{\mathbb{N}}\} < \infty$.

While it is clear by the definition above that every uniform stopping rule is a stopping rule, we now show that the converse is also true.

THEOREM 1. *Every stopping rule is a uniform stopping rule.*

Proof. Let $d : X^{\mathbb{N}} \to Y$ be a stopping rule. Suppose, for the sake of contradiction, d is not a uniform stopping rule. The proof is organized in steps which are as follows.

Step 1: We iteratively define a sequence of pairs $\{(k_j, \mathcal{A}_j)\}_{j\in\mathbb{N}}$, where $k_j \in \mathbb{N}$ and $\mathcal{A}_j \subseteq X^{\mathbb{N}}$, as follows:

1. Let $k_1 := \inf\{k_d(S) : S \in X^{\mathbb{N}}\}$ and $\mathcal{A}_1 := \{S \in X^{\mathbb{N}} : k_d(S) = k_1\}$.
2. For any $j \in \mathbb{N} \setminus \{1\}$, assuming $(k_l, \mathcal{A}_l)$ have already been defined for every $l \in \{1, \ldots, j-1\}$, let

$$k_j := \inf\{k_d(S) : S \in X^{\mathbb{N}} \setminus \cup_{l=1}^{j-1} \mathcal{A}_l\}, \text{ and}$$
$$\mathcal{A}_j := \{S \in X^{\mathbb{N}} \setminus \cup_{l=1}^{j-1} \mathcal{A}_l : k_d(S) = k_j\}$$

The sets $\mathcal{A}_j$ refer to the set of all the sequences (henceforth inputs[3]) for which the stopping time is k_j. From our supposition that d is stopping rule and d does not have a finite bound on the set of stopping times, the following properties are immediate:

(a) For each $j \in \mathbb{N}$, $k_j \in \mathbb{N}$ and $\mathcal{A}_j \neq \varnothing$.

(b) $k_1 < k_2 < \ldots < k_j < \ldots$ and so on. Further, $k_i \geq i$ for all $i \in \mathbb{N}$.

(c) $\{\mathcal{A}_j : j \in \mathbb{N}\}$ is a partition of $X^{\mathbb{N}}$.

These properties shall be referred to in the rest of the argument.

[3]Since our proof involves constructing sequences of sequences of alternatives and subsequences of those sequences, we use the term "input" to denote a sequence of alternatives to avoid any confusion.

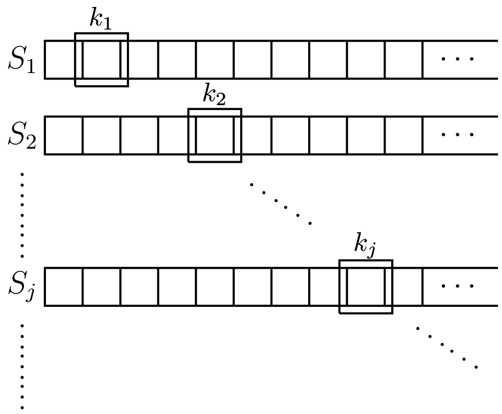


Figure 1: A sequence of inputs $\{S_i\}_{i\in\mathbb{N}}$ with increasing stopping times

Step 2: For every $j \in \mathbb{N}$, pick an arbitrary $S_j \in \mathcal{A}_j$. This generates a sequence $\{S_i\}_{i\in\mathbb{N}}$ of inputs such that the stopping time for each S_j is k_j. By property (b), we know that this corresponds to an increasing sequence of stopping times. Now, we construct a subsequence $\{S_i^*\}_{i\in\mathbb{N}}$ of the above sequence with the following progressive "agreement" property: For all $k \in \mathbb{N}$, we have $S_k^*|_k = S_j^*|_k$ for all $j \geq k$. To do this, we use the following lemma.

Lemma 1. *For any $k \in \mathbb{N}$ and a sequence of inputs $\{T_i\}_{i\in\mathbb{N}}$ where $T_i \in X^{\mathbb{N}}$, there exists a subsequence $\{T_{k_i}\}_{i\in\mathbb{N}}$ such that $T_{k_m}|_k = T_{k_n}|_k$ for all $m, n \in \mathbb{N}$.*

Proof. Consider any $k \in \mathbb{N}$. Since X is finite, the number of possible segments of length k is $|X|^k$. Since $\{T_i\}_{i\in\mathbb{N}}$ is an infinite collection of inputs, by the pigeonhole principle, there exists at least one segment of length k, say M, that is repeated infinitely often and therefore we can construct a subsequence $\{T_{k_i}\}_{i\in\mathbb{N}}$, such that $T_{k_i}|_k = M$ for all $i \in \mathbb{N}$. $\square$

Now using the above lemma, we recursively define an indexed collection of sequences of inputs $\{\{S_{k_i}\}_{i\in\mathbb{N}}\}_{k\in\mathbb{N}}$ as follows:

- For $k = 1$ and $\{S_i\}_{i\in\mathbb{N}}$, applying Lemma 1 we get a subsequence $\{S_{1_i}\}_{i\in\mathbb{N}}$ such that $S_{1_i}|_1 = S_{1_j}|_1$ for all $i, j \in \mathbb{N}$.

- For $k \geq 2$, applying Lemma 1 on the sequence $\{S_{(k-1)_i}\}_{i=2}^{\infty}$, we get a subsequence $\{S_{k_i}\}_{i\in\mathbb{N}}$ such that $S_{k_i}|_k = S_{k_j}|_k$ for all $i, j \in \mathbb{N}$.

Starting from $\{S_i\}_{i\in\mathbb{N}}$ in stage 0, at every stage $k \geq 1$, we generate a sequence of inputs such that all the inputs in the sequence have the same initial k-long segment. Note

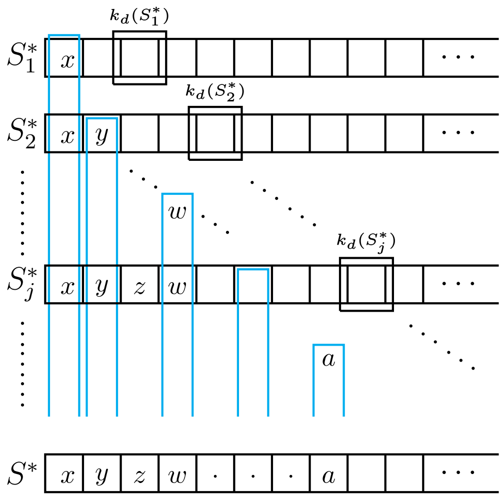


Figure 2: Progressive agreement of $\{S_i^*\}_{i\in\mathbb{N}}$ and the target input S^*

that this indexed collection of sequences is nested i.e. $\{S_{(k+1)_i}\}_{i\in\mathbb{N}}$ is a subsequence of $\{S_{k_i}\}_{i\in\mathbb{N}}$ for all $k \in \mathbb{N}$. Therefore, for any $k, l \in \mathbb{N}$ such that $l > k$, we have $S_{k_i}|_k = S_{l_j}|_k$ for all $i, j \in \mathbb{N}$. In particular $S_{k_1}|_k = S_{l_1}|_k$. Now we define the required sequence of inputs as follows: $S_k^* \coloneqq S_{k_1}$ for all $k \in \mathbb{N}$. That is, S_k^* is equal to the first term (input) of the sequence generated at the k^{th} recursion of the above definition. The sequence of inputs $\{S_i^*\}_{i\in\mathbb{N}}$ thus generated has the property that $S_k^*|_k = S_j^*|_k$ for all $j \geq k$. Further, $\{S_i^*\}_{i\in\mathbb{N}}$ is a subsequence of $\{S_i\}_{i\in\mathbb{N}}$ and hence corresponds to a increasing sequence of stopping times $k_d(S_1^*) < k_d(S_2^*) \ldots < k_d(S_j^*) < \ldots$ where $k_d(S_i^*) \geq k_i \geq i$. Finally, we define the input $S^* \in X^{\mathbb{N}}$ as

$$S^*(i) \coloneqq S_i^*(i) \qquad \forall i \in \mathbb{N}$$

It can be observed that the due to the progressive "agreement", the sequence of inputs $\{S_i^*\}_{i\in\mathbb{N}}$ "converges" to the input S^* i.e. for any $k \in \mathbb{N}$, $S^*|_k = S_j^*|_k$ for all $j \geq k$.

Step 3: Since d is a stopping rule, there must exist $k^* \in \mathbb{N}$ such that

$$d(S^*) = d(S^*|_{k^*} \cdot T) \ \ \forall \ \ T \in X^{\mathbb{N}} \tag{1}$$

Consider any $l > k^*$. Note that since $k_d(S_1^*) < k_d(S_2^*) \ldots < k_d(S_j^*) < \ldots$, there exists $S_l^* \in \{S_i^*\}_{i\in\mathbb{N}}$ such that $k_d(S_l^*) > k^*$. By the definition of S^* and the progressive agreement property of $\{S_i^*\}_{i\in\mathbb{N}}$, we know that $S^*|_k = S_l^*|_k$ for all $k \leq l$. In particular $S^*|_{k^*} = S_l^*|_{k^*}$ and therefore we can write S_l^* as the concatenation of $S^*|_{k^*}$ and $T' \in X^{\mathbb{N}}$, where $T'(i) = S_l^*(k^* + i)$ for all $i \in \mathbb{N}$. But then by (1), we have

$$d(S_l^*) = d([S^*|_{k^*}] \cdot T') = d([S^*|_{k^*}] \cdot T) \quad \forall T \in X^{\mathbb{N}}$$

which implies that k^* is the stopping time for S_l^*, a contradiction. Therefore our initial supposition that d is not a uniform stopping rule is wrong implying that d must be a uniform stopping rule. □

1.2.3 Some remarks

Our proof relies on a diagonalization argument to construct the sequence S^* and there are two critical ingredients in the proof. First, the finiteness of X enables us to establish Lemma 1 and second, the assumption of full domain, $X^{\mathbb{N}}$, allows us to construct the target sequence S^* which leads to the final contradiction. In the Appendix, we discuss a topological approach to our result where we show that the for any stopping rule d, the function $k_d : X^{\mathbb{N}} \to \mathbb{N}$ is continuous when $X^{\mathbb{N}}$ and and $\mathbb{N}$ are endowed with the product topology and the discrete topology respectively. Using the Tychonoff Theorem, we then observe that $X^{\mathbb{N}}$ is compact and hence the function k_d is uniformly continuous. The uniform continuity of k_d gives us the finite bound on the values of k_d.

The reduction lemma is essentially a result about stopping times as defined in probability theory. We provide a formal definition of a stopping time in the Appendix and prove a generalized version of this result concerning a sequence of stopping times. We show that any sequence of stopping times that is pointwise bounded has a uniform bound. This generalized reduction lemma can form the basis for analyzing stochastic stopping rules.

The equivalence of stopping and uniform stopping rules highlights the fact that with the assumption of endogenous stopping, we end up showing that only a finite number of segments are "relevant" in decision making. While this finiteness provides appropriate grounds for testability of various decision procedures (see Section 4), it also provides surprising results when we explore different aspects —such as computability—of decision-making in our setup. We explore such implications in the next section.

1.3 Computability of Decision Rules

It is widely accepted that cognitive limitations and computational constraints have an important role in the decision-making process. While the standard notion of rationality that is synonymous with unrestricted maximization assumes no such constraints, there are a variety of settings where the assumption of infinite processing capabilities is an unrealistic one. As Richter and Wong (1999) remark

> "Can human beings really work with arbitrarily complex preferences, utility functions, and technologies as classical economic theory assumes? . . . real people, using 'realistic' languages, cannot communicate arbitrary real-number quantities and prices. Realism, then, suggests that we restrict ourselves to 'simple' preferences, utility functions, and technologies..."

In order to incorporate computational constraints in our model, a natural first question to ask is that what decision rules are *computable*? To answer this question, we turn to an abstract model of computation: the *Turing machine*. A physical description of a Turing machine involves two objects: A finite state machine and an infinite "tape" which enables it to have an effectively "infinite memory". According to the Church-Turing thesis, any physically realizable computer can be represented using a Turing machine and this makes it the most powerful model of computation known till date. In other words, the question of computational feasibility of a decision problem can be thought of as equivalent to that of its Turing-implementability. Therefore, we call a decision rule computable if it can be implemented by a Turing machine.

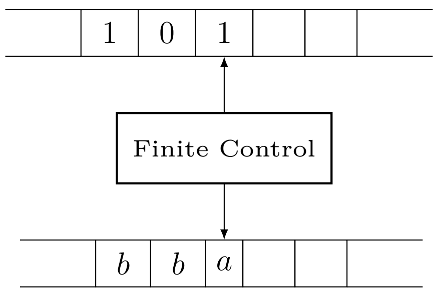


Figure 3: A two-tape Turing machine

What we mean by implementing a decision rule is that there exists a Turing machine such that for any input (a sequence) that is fed (formally defined below) into it, the machine produces the same output as the decision rule on that input. Before describing the decision-making process in our setup using a Turing machine, we first provide its formal definition.

DEFINITION 5. *A Turing machine is a tuple* $TM = (Q, X, \delta, O)$*, where* Q *is a finite set of states,* X *is a finite set of symbols (alternatives),* $\delta : Q \times X^2 \to Q \times X \times \{L, S, R\}^2$ *is a transition function and* $O : \mathcal{S} \to Y$ *is an output function.*[4]

[4]Where $\mathcal{S}$ and Y are as defined in Section 2.

We conceptualize the DM as a Turing machine with a finite number of states denoted by Q and two *tapes* —input and output/working tape—which are infinite one directional line of "cells". Each tape is equipped with a tape head. The tape head of the input tape reads the symbols on the tape one cell at a time whereas the tape head of the output tape can write or rewrite symbols to the tape one cell at a time.

In the standard setup, the inputs to a Turing machine are finite strings from a finite "alphabet" (X). The input in our setup, an infinite sequence, is written on the input tape preceded by a symbol $\triangleleft \notin X$ and the decision-making process is as follows: The symbol $\triangleleft$ *initializes* the machine and it begins in some initial state q_0. Then, it "parses" through the input one at a time using the transition function δ. Depending on the current state and the entries under the two tape heads, the transition function determines the next state, the movement of the tape heads (left, right or stay) and the entry on the output tape. There is a designated set of *terminal* states and once the machine enters a terminal state, it *halts*. The decision is then made using the output function, O, using the segment generated on the output/working tape. Using this notion of a Turing machine, we are now equipped to state a formal definition of computable decision rules.

Definition 6. *A decision rule d is computable, if there exists a Turing machine TM_d such that for all $S \in X^{\mathbb{N}}$, (i) The Turing machine halts; and (ii) $TM_d(S) = d(S)$.*

It is easy to see that not every decision rule is a computable rule. In particular, it is worth observing that rationality —defined as preference maximization —is incompatible with computability i.e. rational choice rules are not computable. To show this, consider a preference order $\succ$ over X and consider any sequence $S \in X^{\mathbb{N}}$ such that it does not feature the $\succ$-maximal element in it. Then, no Turing machine will halt for this input. This is in line with Kramer (1967) who shows that when the DM suffers from computational constraints, it is impossible to display fully rational behavior.

Having defined computability of decision rules, we now introduce another model of computation which is simpler and has been widely used to model various aspects of bounded rationality: a finite automaton. In the context of repeated games, automatons have been used to incorporate the cost of complexity of strategies (see Rubinstein (1986)) and in the context of individual decision-making, they have used been describe the procedural aspects of decision-making (Salant (2011)). It is formally defined as follows.

Definition 7. *An automaton is a tuple $A = (Q, X, \delta, O)$ where Q is a finite set of* *states, X is a finite set of symbols (alternatives), $\delta : Q \times X \to Q$ is a transition function and $O : F \to Y$ is an output function where $F \subset Q$ is the set of terminal states.*

The DM is conceptualized as an automaton in a similar way as a Turing machine. The input is written on an input tape. It starts in an initial state $q_0 \in Q$ and reads elements of an input one at a time. However, an important difference is that the tape head can move only in one direction. For every input element and the current state, the transition function determines the next state of the automaton and the tape head moves to the next element. Within the set of states is a designated set of *terminal* states, denoted by F. Once the automaton enters one of these states, it halts. An output function O then produces a decision based on the terminal state.[5]

In the theory of computation, an automaton is a simpler model since it does not have an infinite tape to simulate an effectively infinite memory. The following example illustrates this difference: Let $X = \{a, b, \varnothing\}$, $Y = \{\text{yes}, \text{no}\}$ and $\mathcal{S}^o$ be the set of all finite segments that comprise of a and b with the last element being $\varnothing$ (indicating the end of the segment). The decision rule outputs "yes" for any $S \in \{a^n b^n \varnothing : n \in \mathbb{N}\}$ i.e. any segment that comprises of n number of a's followed by n number of b's for any $n \in \mathbb{N}$ and it outputs "no" otherwise. Such a decision rule can be implemented using a Turing machine. However, it cannot be implemented by a finite automaton. This example also illustrates the fact that automaton-implementable rules over *lists* form a strict subclass of computable rules.

Since automatons are simpler model of computation than Turing machine, we call decision rules *simply computable* if they can be implemented by an automaton. Formally, they are defined as follows.

Definition 8. *A decision rule d is simply computable if there exists an automaton A_d such that for all inputs $S \in X^{\mathbb{N}}$, (i) The automaton halts and; (ii)$A_d(S) = d(S)$.*

To illustrate a simply computable decision rule, consider the DM in Example 1. Suppose $X = \{x, y\}$, the weight is 1 for both the alternatives and the threshold is 2. Then this decision rule can be implemented using an automaton with 5 states, excluding the initial state, q_0 (see Figure 4). Our notions of computable and simply computable rules are closely linked to stopping rules and uniform stopping rules and using the reduction lemma, we are able to show that all of them are in fact equivalent.

[5]In Salant (2011), the output function requires both the state as well as the symbol under the tape head to produce the output. Such machines are called Mealy machines. Whereas our formulation is similar to the one in Rubinstein (1986) and such machines are called Moore machines.

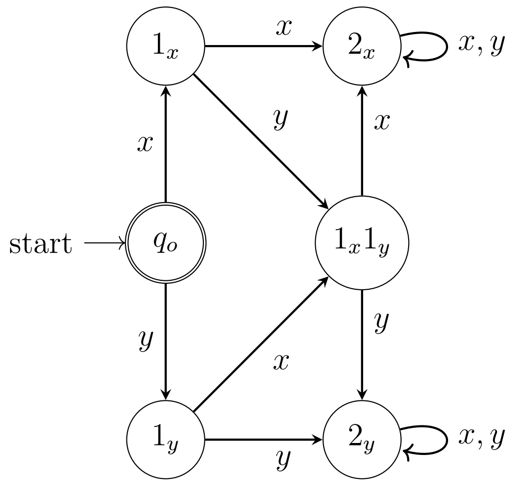


Figure 4: An automaton-computable decision rule

Theorem 2. *Every computable rule is simply computable.*

Proof. Let d be a computable decision rule implementable by a Turing machine TM_d. Consider any arbitrary input $S \in X^{\mathbb{N}}$. Since the Turing machine halts for S, there exists $k \in \mathbb{N}$ such that TM_d does not examine alternatives in S beyond $S|_k$. Consider any $S' \in X^{\mathbb{N}}$ such that $S'|_k = S|_k$. The Turing machine does not examine alternatives beyond k in S' as well and since $S'|_k = S|_k$, we have $d(S') = d(S)$. Therefore d is a stopping rule. By the reduction lemma, it is a uniform stopping rule. So, there exists $k^* \in \mathbb{N}$ be such that for all $S, T \in X^{\mathbb{N}}$, we have $d(S) = d([S|_{k^*}] \cdot T)$. An automaton with at most $\sum_{j=1}^{k^*-1} |X|^j$ non-terminal states (one for each segment of length less than k^*) and $|X|^{k^*}$ terminal states (one for each segment of length k^*) can implement this decision rule. Therefore, it is simply computable. $\square$

Remark. The construction of the automaton in the above proof is the most "inefficient" in terms of the state complexity i.e. the number of states. This is the largest number of states required to implement a uniform stopping rule/simply computable rule. To illustrate this fact, consider the example above. Going by the construction given in the proof, the automaton will have 6 states instead of 5, excluding the initial state (the additional state being $1_y 1_x$ which will be different from $1_x 1_y$).

1.4 Choice Rules

One of the central objects of study in abstract choice theory is a choice function. The domain of a choice function is a collection of "menus" and for every menu it gives the choice set, namely the set of chosen or "choosable" alternatives *from* that menu.[6] In the classical theory, these menus correspond to *sets* of alternatives. The analogue of a menu in our model is an infinite sequence and that of a single-valued choice function is what we term as a choice rule. These form a subclass of decision rules that require the DM to choose an alternative from within the sequence. They are formally defined as follows.

Definition 9. *A choice rule d is a map $d : X^{\mathbb{N}} \to X$ such that for all $S \in X^{\mathbb{N}}$, $d(S) = S(i)$ for some $i \in \mathbb{N}$.*

If we restrict our attention to choice rules that are also stopping rules, we get the further restriction that for any sequence S, the choice must lie within the initial $k_d(S)$-long segment. An import of Theorem 2 is that stopping rules are equivalent to computable decision rules. While studying choice rules, we maintain that the assumption of computability is a *normative* one. This is in line with the interpretation that in our model, the DM encounters alternatives sequentially, one at a time. Therefore the assumption of stopping or equivalently that of computability is a plausible one. In this section, we first provide a characterization of computable choice rules and provide a "toolkit" for conducting revealed preference analysis in our setup. Then we do a revealed preferences analysis of two natural choice procedures that capture the idea of "satisficing" behavior.

1.4.1 *Continuity and computability*

With the added structure to the decision space in the case of choice rules, we provide a characterization of computable choice rules via their continuity with respect to a natural topological structure on the domain and the co-domain. This is captured in the following result, a proof of which is relegated to the appendix.

Theorem 3. *Consider a non-constant choice rule d and assume X and $X^{\mathbb{N}}$ are endowed with the discrete and the product topology respectively. Then d is computable if and only if it is continuous.*

The domain of inputs, $X^{\mathbb{N}}$, is often referred to as a Cantor space and the product topology on it as Cantor topology. In the proof of the result, we first show the structure of basic open sets in the product topology, which are referred to as open cylinders. These correspond to all inputs which "agree" on one location. Finite intersections of such open cylinders are called cylinder sets and they form the basis for this topology. Therefore the

[6] Multi-valued functions are often called choice correspondences.

behavioral interpretation is that a DM considers two sequences “approximately” same by comparing only finitely many initial locations. Continuity of the choice rule then implies that the DM cannot display “jumps” for close enough choice problems.

When restricting our attention to choice rules, the reduction lemma allows us to succinctly represent computable choice rules via finite *trees*. To show this, we revisit Example 1 again with $X = \{x, y\}$, both the alternatives having a weight of 1 and the threshold weight being 2. The root node of the tree is the “null” symbol and every path from the root node to a terminal node corresponds to a $k_d(\cdot)$ long segment. Further, in the case of the procedure in this example, the terminal node indicates the choice from the $k_d(\cdot)$ segment. This provides a complete descirption of the choice rule (see Figure 5).

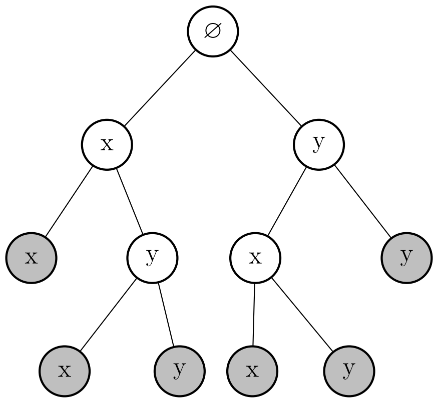


Figure 5: Tree representation of a computable choice rule

1.4.2 *A revealed preference toolkit*

In the theory of choice from sets, the analysis of different choice procedures involves imposing some consistency properties—called axioms—on choice functions. These axioms are often in the form of “contraction” or “expansion” properties i.e. consistency of choices across menus that are related via set inclusion. In order to conduct an axiomatic analysis of choice procedures in our setup, we require a suitably adapted “language” to state such axioms on choice rules. To that end, we introduce two useful informational concepts of *sufficiency* and *minimal sufficiency* of segments. In order to define these formally, we require some notation. Recall that a segment is any map $M : [k] \to X$ where $k \in \mathbb{N}$ and the set of all segments is denoted by $\mathcal{S}$. Denote domain of a segment M as $\text{dom}(M)$. We define a strict partial order[7] over the set of all segments $\mathcal{S}$ as follows: for

[7] An asymmetric and transitive binary relation.

any $M, M' \in \mathcal{S}$, let $M \triangleright M'$ if and only if (i) $\mathrm{dom}(M') \subsetneq \mathrm{dom}(M)$ and (ii) $M(i) = M'(i)$ for all $i \in \mathrm{dom}(M')$. The relation $\triangleright$ is thus the "extending" relation and $M \triangleright M'$ can be interpreted as the segment M "extends" the segment M'. A sufficient segment is defined as follows.

Definition 10. *For a decision rule d, a segment M is sufficient if $d(M \cdot T) = d(M \cdot T')$ for all $T, T' \in X^{\mathbb{N}}$.*

The intuitive content of the above definition is as follows. As the DM faces an sequence S, there comes a point $k \in \mathbb{N}$ when the segment $M = S|_k$ has enough information for the decision maker to have "made up her mind" i.e. M is informationally "sufficient" to enforce a decision. However, the acquired information will *not* be sufficient until a certain point. This motivates the notion of *minimal sufficiency.*

Definition 11. *For a decision rule d, a segment M is minimal sufficient if it is sufficient and for any M' such that $M \triangleright M'$, the segment M' is not sufficient.*

Minimal sufficiency captures the idea of the "critical" length of a segment to enforce a decision. By critical, we mean that if the segment is smaller than that length, it can no longer guarantee the same decision for all concatenated sequences. Note that the definition of stopping rules indicates that every sequence must have a corresponding *minimal sufficient* segment that "implements" the decision. For a given stopping rule, d, let the class of all *sufficient* and *minimal sufficient* segments be denoted by $\mathscr{S}$ and $\mathscr{MS}$ respectively. If $M = S|_k$ for some $k \in \mathbb{N}$ and M is a sufficient segment for a decision rule d, then we will abuse notation and denote the decision for S by $d(M)$ i.e. $d(S) = d(M)$.

To illustrate the idea of sufficiency and minimal sufficiency, let us consider Example 1 again. Let $X = \{a, b, c\}$ and all alternatives have weight 1. Suppose the DM has a threshold value of 3 and consider the sequence $S = (a\ b\ c\ a\ b\ c..)$ i.e. it consist of "cycles" of alternatives a, b and c. Here, the minimal sufficient segment is of length 7 i.e. where a is the first alternative to appear 3 times. Any intial segment of S with length less than 7 is not minimal sufficient and any segment with length more than 7 is sufficient.

Since the set of sequences is infinite, in practice, the identification of sufficient (and minimal sufficient) of segments is not possible for a given stopping rule. In order to overcome this problem of identification, we introduce a new object —a *decision procedure* —that captures the "dynamic" aspect of decision-making. Denote by $\star$, a symbol not in the decision space Y, representing "indecision". Decision procedures are maps that are defined on the set of all finite segments. They map every finite segment to either "indecision" or some decision in Y.

Definition 12. *A decision procedure is any map* $d_* : \mathcal{S} \to Y \cup \{\star\}$ *such that*

(i) If $M' \triangleright M$ *and* $d_*(M') = \star$*, then* $d_*(M) = \star$

(ii) If $M' \triangleright M$ *and* $d_*(M) \in Y$*, then* $d_*(M) = d_*(M')$

(iii) For any sequence of segments $M_1, M_2, \ldots$ *satisfying* $M_{k+1} \triangleright M_k$ *for all* $k \in \mathbb{N}$*, there exists* $k \in \mathbb{N}$ *such that* $d_*(M_k) \in Y$

A decision procedure can be thought of as a dynamic representation of a stopping rule. If a DM is represented by a decision procedure, the three consistency requirements can be interpreted in the following manner. First, if the DM has not made a decision at a given point (in time or space), i.e. at a segment, then she would have not made a decision at any preceding point as well i.e. at any sub-segment of that segment. Second, if the DM has made a decision at a given point, then for any subsequent point as well she makes a decision. Further, she makes the same decision at any subsequent point as well. Third, for any progressively increasing sequence of segments, she makes a decision at *some* point along the sequence.

The connection between stopping rules and decision procedures is made precise by defining a map that outputs a decision procedure for every stopping rule. Let $\mathcal{D}_s$ and $\mathcal{D}_*$ be the set of all stopping rules and decision procedures respectively. To each stopping rule d, associate the corresponding map $d_*^\dagger : \mathcal{S} \to Y \cup \{\star\}$ as follows:

1. For any segment $M \in \mathcal{S}$, let $d_*^\dagger(M) := \star$ if there exists $S \in X^{\mathbb{N}}$ and $k \in \mathbb{N}$ such that $k < k_d(S)$ and $M = S|_k$.
2. For any segment $M \in \mathcal{S}$, let $d_*^\dagger(M) := d(S)$ if there exists $S \in X^{\mathbb{N}}$ such that $k \geq k_d(S)$ and $M = S|_k$.

Lemma **2.** *The map* $d_*^\dagger$ *is well defined and* $d_*^\dagger \in \mathcal{D}_*$.

Proof. To show $d_*^\dagger$ is well defined, consider any arbitrary $M \in \mathcal{S}$. Suppose there exists $S \in X^{\mathbb{N}}$ and $k \in \mathbb{N}$ such that $k < k_d(S)$ and $M = S|_k$. Then by definition $d_*^\dagger(M) = \star$. Assume for contradiction that there exists $S' \in X^{\mathbb{N}}$, $S' \neq S$ such that $M = S'|_k$ and $k \geq k_d(S')$. Now, since $S'|_{k_d(S')} = S|_{k_d(S')}$, by the definition of a stopping rule, we must have $k_d(S') = k_d(S)$, a contradiction. Now, suppose there exist $S, S' \in X^{\mathbb{N}}$ and $k \in \mathbb{N}$ such that $S|_k = S'|_k = M$ and $\max(k_d(S), k_d(S')) \leq k$. Since $S|_k = S'|_k = M$, by the definition of stopping rules we must have $k_d(S) = k_d(S')$ and hence $d(S) = d(S') = d(M)$.

Therefore, $d_*^\dagger$ is well-defined.
Now, to show $d_*^\dagger$ is a decision procedure, consider an arbitrary $M \in \mathcal{S}$. If $d_*^\dagger(M) = \star$, then consider any M' such that $M \triangleright M'$. Since there exists $S \in X^{\mathbb{N}}$ such that $k < k_d(S)$ and $M = S|_k$, we know that $M = S|_{k'}$ for some $k < k$ and hence we have $d_*^\dagger(M) = \star$. Now, suppose $d_*^\dagger(M) = d(S)$ for some $S \in X^{\mathbb{N}}$. We know that $M = S|_k$ for some k and $k \geq k_d(S)$. Consider any $M' \in \mathcal{S}$ such that $M' \triangleright M$. For any $S' \in X^{\mathbb{N}}$ and $k' > k$ such that $S'|_{k'} = M$, we know that $S'|_{k_d(S)} = S|_{k_d(S)}$ Therefore we must have $k_d(S') = k_d(S)$ and hence $d(S) = d(S') = d_*^\dagger(M')$. Finally consider any sequence of segments $M_1, M_2, \dots$ such that $M_{k+1} \triangleright M_k$ for all $k \in \mathbb{N}$. Now, since every stopping rule is a uniform stopping rule, there exists a $S \in X^{\mathbb{N}}$ and $k \in \mathbb{N}$ such that $k \geq k_d(S)$ such that $M_{k'} = S|_k$ for some $k' \in \mathbb{N}$. Therefore, we have $d_*^\dagger(M_{k'}) = d(S) \in Y$. $\square$

Using the above lemma, we let $\eta : \mathcal{D}_s \to \mathcal{D}_*$ be defined as:

$$\eta(d) \coloneqq d_*^\dagger \text{ for every } d \in \mathcal{D}_S$$

This map provides natural way to assign a unique decision procedure for every stopping rule. Further, for every decision procedure there exists a unique stopping rule. This claim is established by the following result.

Proposition 1. *The map $\eta : \mathcal{D}_s \to \mathcal{D}_*$ is a bijection.*

Proof. To show that η is one-to-one, consider $d, d' \in \mathcal{D}_s$ such that $d \neq d'$. Therefore there exists $S \in X^{\mathbb{N}}$ such that $d(S) \neq d'(S)$. Let $k \coloneqq \max(k_d(S), k_{d'}(S))$ and consider $M = S|_k$. Let $\eta(d) = d_*^\dagger$ and $\eta(d') = d_*'^\dagger$. By definition $d_*^\dagger(M) = d(S) \neq d(S') = d_*'^\dagger(M)$. Therefore, η is one-to-one.
To show that η is onto, consider any arbitrary $d_* \in \mathcal{D}_*$. We need to define $d \in \mathcal{D}_s$ such that $\eta(d) = d_*$. Define $d : X^{\mathbb{N}} \to Y$ as follows. Consider any $S \in X^{\mathbb{N}}$ and the sequence of segments $S|_1, S|_2, \dots$. Note that $S|_{k+1} \triangleright S|_k$ for all $k \in \mathbb{N}$ and by definition of a decision procedure, there exists $k \in \mathbb{N}$ such that $d_*(S|_k) \in Y$. Let $k^* = \inf\{k \in \mathbb{N} : d_*(S|_k) \in Y\}$. Define $d(S) \coloneqq d_*(S|_{k^*}) \in Y$. By definition of d_*, we have $d(S|_{k^*} \cdot T) = d(S)$ for all $T \in X^{\mathbb{N}}$ and $k_d(S) = k^*$. Therefore d is a stopping rule. Also, by the definition of η, we have $\eta(d) = d_*$. Therefore, η is onto. $\square$

We have shown that there is a natural bijection between the class of stopping rules and that of decision procedures. While mathematically equivalent, decision procedures and stopping rules are conceptually different objects. Stopping rules process entire infinite sequences whereas decision procedures show how the DM processes information when the infinite sequences are presented "gradually" in a dynamic manner. To study physical settings, decision procedures provide a more realistic model of a DM. The minimal sufficiency and sufficiency of segments are naturally defined for decision procedures as follows: For a decision procedure d_*, a segment M is

(i) *Minimal sufficient* if $d_*(M) \in Y$ and $d_*(M') = \star$ for all M' such that $M \triangleright M'$.

(ii) *Sufficient* if it is minimal sufficient or there exists $M' \in \mathcal{S}$ such that M' is minimal sufficient and $M \triangleright M'$.

Decision procedures are useful from the revealed preference perspective as they enable complete identification of minimal sufficient and sufficient segments in finitely many steps. For instance, given a tree representation of a decision procedure that corresponds to a choice rule, a Depth First Search (DFS) algorithm will output the class of all minimial sufficient segments —and consequently sufficient segments —in finite time. As we will see in the next section, these fully identifiable segments will help us formulate axioms to behaviorally characterize some natural choice procedures.

1.5 Stopping via Satisficing

Satisficing, first introduced by Herbert Simon (see Simon (1955)) is an influential model of decision-making and has been studied widely in the literature (see Kovach and Ülkü (2020), Aguiar et al. (2016), Tyson (2015) and Papi (2012), among others). The basic idea underlying satisficing is that the due to factors like computational constraints, complexity of the choice problem etc., a DM may not resort to optimizing behavior. Instead, based on a binary classification of the alternatives into acceptable/satisfactory and non-acceptable/unsatisfactory, she may select any alternative belonging to the former category. Satisficing behavior is often modeled as a DM examining alternatives sequentially until a "good enough" alternative is observed. While some existing models endogenize the search order of the DM (see Aguiar et al. (2016)), others treat it as observable in the form of a list and vary the threshold (see Kovach and Ülkü (2020)). Our setup provides a natural way to study satisficing behavior and in this section, we propose two simple adaptations of it and provide their behavioral characterization. The first one is a formalization of Example 1 and the second is a formalization of Example 2.[8]

[8]The choice procedures formulated in this section form a subclass of stopping rules. Since stopping

1.5.1 *Cardinal Satisficing*

The DM is equipped with two objects. The first one is a *weight* function $w : X \to \mathbb{R}_{++}$ that assigns a positive real number to every alternative. The weights can be thought of as some scores the DM assigns to the alternatives that are indicative of the relative importance of alternatives. For instance, while seeking movie recommendations, a DM may give a higher score to "action" movies over the ones belonging to the genre "drama". The second object is a *threshold* weight $v \in \mathbb{R}_+$. This threshold corresponds to the satisficing component that the DM uses to make decisions.

The DM uses the following procedure to make choices. For any sequence, she "parses" through it sequentially, maintaining a count of the cumulative weight of each alternative in a "register". As soon as she encounters an alternative whose cumulative weight crosses the threshold, she stops and selects it. We call this procedure as a Cardinal Satisficing Rule (CSR).[9] The reason this procedure can be classified as satisficing behavior is that it allows for the choice of "sub-optimal" alternatives. For instance, for low threshold levels, alternatives with lower weights can be chosen if they are presented sufficiently many times to the DM before the alternatives with higher weights.

In order to formally define this procedure, denote for any given sequence S and a position $N \in \mathbb{N}$, the cumulative weight of an alternative x as

$$W_S^N(x) := |\{i \in [N] : S(i) = x\}|.w(x)$$

Now, we can define CSR formally as follows.

Definition 13. *A computable choice rule d is a Cardinal Satisficing Rule (CSR) if there exists $v \in \mathbb{R}_+$ and $w : X \to \mathbb{R}_{++}$ such that for any $S \in X^{\mathbb{N}}$,*

$$d(S) = x^*(S)$$

where $x^*(S) \in X$ is the unique alternative satisfying the following condition: $W_S^N(x) \geq v > W_S^N(y)$ for all $y \neq x$ and some $N \in \mathbb{N}$.

Note that since the weights assigned to alternatives are positive real numbers, for any sequence, there exists some position in it such that exactly one alternative's cumu-

rules are equivalent to decision procedures, these can be formulated as decision procedures as well and all the results go through. However, for expositional and notational convenience, we will operate in the domain of stopping rules.

[9] The use of the word "cardinal" stems from the fact that the intensity of the weights can affect the choice behavior.

lative weight crosses (weakly) the threshold at that position. This defines the stopping condition of the DM. This procedure is behaviorally characterized by two axioma. In order to state the first axiom, we introduce the concept of a *favorable transformation* of a sequence with respect to an alternative. Intuitively, this involves bringing an alternative "closer" to the DM by transforming that sequence into a new one. That is, by lowering its position, an alternative is examined earlier than it was examined previously. There are two ways to favorably transform an sequence with respect to an alternative. The first way is to interchange the position of that alternative with another alternative that precedes it in the input. Formally, for any sequence S and $k \in \mathbb{N}$, let $\hat{S}^k \in X^{\mathbb{N}}$ be the sequence which is defined as follows:

$$\hat{S}^k(i) := \begin{cases} S(k+1) & \text{if } i = k; \\ S(k) & \text{if } i = k+1; \\ S(i) & \text{otherwise.} \end{cases}$$

That is, the sequence $\hat{S}^k$ is obtained from S by interchanging its k^{th} and $(k{+}1)^{st}$ elements. We call $\hat{S}^k$ a *favorable shift* of S with respect to an alternative x if $S(k+1) = x$. Denote the class of all favorable shifts of S with respect to an alternative x by $\mathcal{FS}(S, x)$. The second way to bring an alternative closer to the DM is by deleting another alternative. Formally, for any sequence S and $k \in \mathbb{N}$ let $\tilde{S}^k \in X^{\mathbb{N}}$ be the sequence defined as

$$\tilde{S}^k(i) := \begin{cases} S(i) & \text{if } i < k; \\ S(i+1) & \text{if } i \geq k \end{cases}$$

The sequence $\tilde{S}^k$ is obtained from S by dropping the alternative located at the k^{th} position. We call $\tilde{S}^k$ as a *favorable deletion* of S with respect to an alternative x if $S(k) \neq x$. Denote the class of all favorable deletions of S with respect to an alternative x be denoted by $\mathcal{FD}(S, x)$.

For any $S \in \mathcal{S}$ and $x \in X$, a *favorable transformation* of S with respect to x is a favorable shift *or* a favorable deletion. The class of all favorable transformations of S with respect to x shall be denoted by $\mathcal{F}(S, x)$. Therefore, $\mathcal{F}(S, x) = \mathcal{FS}(S, x) \cup \mathcal{FD}(S, x)$ by definition. Our first condition requires that the stopping rule should be "monotone" when it comes to favorable transformations with respect to the chosen alternatives. That is, it requires the DM to make the same choice if the chosen alternative is brought "closer" to him via a favorable tranformation.

Monotonicity: A decision rule satisfies monotonicity if for any $S, S' \in X^{\mathbb{N}}$ such that $S' \in \mathcal{F}(S, x)$,

$$[d(S) = x] \implies [d(S') = x]$$

The second condition is about the effect on choice when a sufficient segement is concatenated to any truncation of a minimal sufficient segment. It states that if a minimal sufficient segment M "implements" an alternative x and another sufficient segment N that does not contain x implements some other alternative, then concatenating any truncation of M with N prevents x from being chosen. In other words, it asserts that a sufficient segment not containing an alternative can "dominate" a non-minimal sufficient segment in an informational sense. We say that for an alternative x and a segment M, $x \notin M$ if $M(i) \neq x$ for all $i \in \text{dom}(M)$. That is, $x \notin M$ when x does not appear in any position of the segment M.

Informational Dominance: A decision rule d satisfies informational dominance if for any $M \in \mathscr{M}\mathscr{S}$ and $N \in \mathscr{S}$ such that $d(M) = x$ and $d(N) \neq x$ and for any M' such that $M \triangleright M'$,

$$[x \notin N] \implies [x \neq d([M' \cdot N])]$$

Note that since the segment $[M' \cdot N]$ contains a sufficient segment within it, it must be a sufficient segment itself. To illustrate this condition, consider a DM who is a cardinal satisficer that assigns weight 1 to each alternative in $X = \{a, b, c\}$ and has a threshold of weight of 3. The segment $M = (a\ b\ c\ a\ b\ c\ a)$ is a minimal sufficient segment with $d(M) = a$. Consider another segment $N = (b\ b\ c\ b\ b)$. Note that this a sufficient segment since b appears 3 times in it. Now consider an arbitrary truncation M' of M. Informational dominance requires that for any concatenation of M' with the segment N, the choice cannot be equal to a. In this case, for any truncation M', we can see that $d(M' \cdot N) = b$ since b is the first alternative whose cumulative weight reaches 3. We now show that these two conditions are characterize cardinal satisficing behavior.

Theorem 4. *A computable choice rule is a CSR if and only if it satisfies Monotonicity and Informational Dominance.*

1.5.2 *Ordinal Satisficing*

Now we turn to the second adaptation of satisficing behavior in our setup. The DM is represented by three objects: (i) A ranking over the set of alternatives X, denoted by

$\succ$ which is a preference order, (ii) A threshold alternative $a^* \in X$ which is used for the binary classification of the set of alternatives into satisfactory and unsatisfactory; and (iii) an attention parameter $k \in \mathbb{N}$ that specifies the relevant segment for any sequence.

The DM uses the following procedure to make a choice. For any sequence, she "parses" through it sequentially. She stops if she encounters a^* or an alternative that is ranked above a^*. Otherwise she stops after observing the first k alternatives and chooses the $\succ$-maximal one from the set of observed alternatives[10]. This is in contrast with the satisficing model discussed in Rubinstein (2012) where a DM chooses the last alternative from the list if it contains no alternative ranked above the threshold. To illustrate this procedure, consider an example where $X = \{a, b, c, d\}$ and the DM's preferences are $a \succ b \succ c \succ d$, the attention parameter k is 2 and the threshold alternative is b. Consider a sequence $S = (c\ d\ a\ a \dots)$. Since the first two positions do not contain any satisfactory alternative, the choice is c whereas the choice from the sequence $S' = (a\ b\ c\ c \dots)$ is a.

In order to formally define our procedure, we let for any alternative $x \in X$ its upper and lower contour set with respect to $\succ$ be denoted by $U(x)$ and $L(x)$, respectively. That is, $U(x) := \{y \in X : y \succ x\}$[11] and $L(x) := X \setminus U(x)$. Now, we formally define this procedure.

Definition 14. *A computable choice rule d is an Ordinal Satisficing Rule (OSR) if there exists $(\succ, a^*, k)$ such that for any $S \in X^{\mathbb{N}}$,*

$$d(S) = \begin{cases} S(i) & \text{where } i \in [k], S(i) \in U(a^*), S(j) \in L(a^*)\ \forall j < i; \\ \max(\bigcup_{i \in [k]} S(i), \succ) & \text{if } S(i) \in L(a^*)\ \forall\ i \in [k] \end{cases}$$

Consider two extreme cases of this procedure. The first is when everything is satisfactory. That is, $x \succ a^*$ for all $x \in X$). In this case, the DM always chooses the first alternative that is presented to her. On the other hand, if a^* is the $\succ$-maximal alternative, then this procedure corresponds to "attention-contrained" *rational* behavior. That is, within the limited attention span of the DM, she always chooses the best alternative. Therefore rational behavior is only a special case which is in contrast with satisficing over sets where it is indistinguishable from preference maximization irrespective of the threshold (See Rubinstein (2012)).

This procedure is behaviorally characterized by three axioms. Before we state these

[10]For any set A, the $\succ$-maximal set, denoted by $\max(A, \succ)$ is defined as $\max(A, \succ) := \{x \in A | \neg y \succ x\ \forall\ y \in A \setminus \{x\}\}$.

[11]Note that since $\succ$ is assumed to be reflexive, $x \in U(x)$ for all $x \in X$.

axioms, we need to define the concept of a *decisive* alternative. The idea behind a decisive alternative is that whenever it is present in a minimal sufficient segment, it is chosen. Intuitively, it dominates attention of the DM and enforces its choice. For an arbitrary choice rule, the set of decisive alternatives may be empty. However, we will show that in the case of OSR, it turns out to be non-empty. Further, in the special case where the DM is an *attention-constrained* preference maximizer i.e. she chooses the $\succ$-maximal alternative after viewing a fixed length of alternatives, this set will be a singleton.

Definition 15. *For a decision rule d, an alternative x is decisive if for all $M \in \mathscr{MS}$,*

$$[x \in M] \implies [x = d(M)]$$

We say that an alternative is *non-decisive* if it is not decisive. Denote by D and $D' = X \setminus D$, the set of all decisive alternatives and non-decisive alternatives, respectively. Now, based on the notion of a decisive alternative, let $\mathscr{MS}_{D'}$ and $\mathscr{MS}_D$ denote the classes of all minimal sufficient segments that do not contain a decisive alternative and that do contain a decisive alternative, respectively. That is, $\mathscr{MS}_{D'} \coloneqq \{M \in \mathscr{MS} : M \subseteq D'\}$ and $\mathscr{MS}_D \coloneqq \mathscr{MS} \setminus \mathscr{MS}_{D'}$.

Our first condition is an adaptation of condition α (also called Chernoff's condition) for single valued choice functions to our setup. In the case of menus as sets, condition α requires that if an alternative that is chosen from a menu, say A and it is present in a smaller menu, say B where $B \subset A$, then it must be chosen from B also. In our setup, we say that if an alternative is chosen in minimal sufficient segment and it is present in another minimal sufficient segment whose range is contained within the range of the former segment, then it must be chosen in the new segment as well. Notice that this is true for a decisive alternative by definition, therefore we impose it for only minimal sufficient segments belonging to $\mathscr{MS}_{D'}$. For any two segments M and M', if $x \in M$ implies $x \in M'$, we abuse notation and write $M \subseteq M'$.

Sequential-α: A choice rule d satisfies sequential-α if for any $M, M' \in \mathscr{MS}_{D'}$ such that $M \subseteq M'$,

$$[d(M') \in M] \implies [d(M') = d(M)]$$

Our next axiom is an adaptation of the No Binary Cycles (NBC) condition on choice functions over sets. NBC requires that binary choices cannot display cycles implying that in the case of single valued choice functions, the pairwise revealed relation must be transitive. The analaogue of a binary menu in our setup is a minimal segment that has only two alternatives in it. For any two distinct alternatives x, y, denote by M_{xy} any segment that has only x and y in it. Therefore, our NBC condition prevents choices from these segments to display cycles.

SEQUENTIAL-NBC: A choice rule d satisfies sequential-NBC is for any $x, y, z \in X$ and $M_{xy}, M_{yz}, M_{xz} \in \mathscr{MS}_{D'}$,

$$[d(M_{xy}) = x, d(M_{yz}) = y] \implies [d(M_{xz}) \neq z]$$

Our final axiom is about the effect of replacing alternatives on the informational content of a minimal sufficient segment. It requires that if the occurence of a non-decisive alternative in a minimal sufficient segment is replaced by some other non-decisive alternative, its informational content should remain the same. In other words, the new segment should still be a sufficient segment. Again, this applies to only segments that do not have decisive alternatives since it holds for decisive alternatives by definition.

REPLACEMENT: A choice rule d satisfies replacement if for any $x, y \in D'$ and $M, M' \in \mathcal{S}_m$ such that $M(i) = x$, $M'(i) = y$ for some $i \in [m]$ and $M(j) = M'(j)$ for all $j \neq i$, $j \in [m]$

$$[M \in \mathscr{MS}_{D'}] \implies [M' \in \mathscr{S}]$$

To illustrate, consider the following choice procedure that *violates* this condition: Let $X = \{x, y, x\}$ and d be defined as follows

$$d(S) = \begin{cases} S(3) & \text{if } S(3) \in \{x, y\} \\ S(5) & \text{otherwise} \end{cases}$$

For any sequence S, the DM looks at the third location. If it is either x or y, then she chooses it, otherwise she picks the alternative at the fifth location. The minimal sufficient segments are of size 3 or 5. Further, no alternative is decisive. Now consider a segment $M = (x\ y\ x)$. It is clear that this segment is a minimal segment since for all $S \in X^{\mathbb{N}}$, we have $d([M \cdot S]) = x$. Now, create M' by replacing x in the third location with z. That is, $M' = (x\ y\ z)$. The segment M' is no longer sufficient since for any sequence $S = (x\ y\ z\ x\ y\ \ldots)$, we have $d(S) = y$, whereas for any $S' = (x\ y\ z\ x\ z\ \ldots)$, we have $d(S') = z$. Now, we are ready to state our result.

THEOREM 5. *A computable choice rule is an OSR if and only if it satisfies Sequential-α, Sequential-NBC and Replacement.*

1.6 Configuration Dependent Rules

In this section, we discuss a class of choice rules that have a conceptually different interpretation than the foregoing discussion. Each sequence is thought of as a menu that is composed of finitely many sequences, one associated with each alternative that appears in it. The DM decomposes every sequence into this collection of "simple" sequences. Using this decomposition, the DM makes the choice. To fix ideas, consider the following scenario. There is finite collection of companies denoted by X. At every time period t, a government agency leases out a production facility to exactly one of them. The company that gets to use the facility at time period t generates a value of 1 in that period for its shareholders and rest of the companies generate a value of 0. A research agency produces a *forecast* which is an infinite sequence that indicates its assessment of which company will get to use the facility in each time period. A DM is endowed with a *discount* factor δ and wishes to invest in one of the companies using this forecast. She uses the following procedure. For every forecast, she looks at the infinite "value stream" associated with each company/stock and selects the stock with the highest sum of discounted values.[12]

The above defined procedure generates a choice rule which for every forecast i.e. an infinite sequence, selects an alternative from that sequence. We define a broad class of choice rules within which such a choice rule lies. We call these rules configuration dependent rules. The underlying idea is that the decision is made using the "configuration" of the alternatives i.e. the pattern of their occurrence in a sequence. Configuration dependent rules subsume many possible behaviors. One can think of rules that utilize the information on positioning of alternatives to make choices such as ones that focus on frequency of alternatives appearing in the sequence.

Configuration dependent rules can be formally defined using a configuration *selection function*. A configuration is any sequence $b \in \{0,1\}^{\mathbb{N}}$ i.e. a sequence of 0's and 1's. Let $b(i)$ denote the i^{th} component of the configuration b. In order to define a configuration selection function, we first need to introduce the idea of a feasible collection of configurations. We call a collection of configurations B *feasible* if it satisfies the following condition:

$$|\{b \in B | b(i) = 1\}| = 1 \quad \forall i \in \mathbb{N}$$

[12] This choice procedure relates to the ranking of infinite utility streams by the discounting criterion. There is a large literature on ranking infinite utility streams that originated with Koopmans (1960).

In words this condition says that for any arbitrary position $i \in \mathbb{N}$, there is exactly one configuration in the collection that contains 1 at its i^{th} position. Denote by $\mathbb{B}$ the set of all configurations and $\mathscr{B}$ the set of all feasible collections of configurations. It is worth noting that every X-valued sequence S can be viewed as a feasible collection of configurations, one associated with each x that appears in S. A visual representation of a sequence as a feasible collection of configurations is given via a configuration matrix.

$$
\begin{array}{c|cccc|}
 & b_x & b_y & b_z & \cdots\cdots \\
\hline
1 & 0 & 1 & 0 & \cdots 0 \cdots \\
2 & 1 & 0 & 0 & \cdots 0 \cdots \\
3 & 0 & 0 & 1 & \cdots 0 \cdots \\
\vdots & 0 & 0 & 0 & \cdots 1 \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

Figure 6: A configuration matrix

Now, we formally define a configuration selection function that selects from every feasible collection of configuration, a configuration from it as follows.

Definition 16. *A configuration selection function is a map $f : \mathscr{B} \to \mathbb{B}$ such that $f(B) \in B$ for all $B \in \mathscr{B}$.*

Define $x(S) \coloneqq \{a \in \{0,1\}^{\mathbb{N}} : a(i) = 1 \text{ if } S(i) = x,\ 0 \text{ otherwise}\}$. This corresponds to the configuration of the alternative x in the sequence S i.e. the configuration associated with x in S. We know that every S corresponds to a feasible collection of configurations. Denote by $B(S)$ the feasible collection of configurations generated by S where the number of configurations is equal to the number of alternatives appearing in S. To illustrate this, consider the sequence $S = (x\ y\ z\ x\ y\ z ...)$ i.e. the sequence comprising of "cycles" of x, y and z. This will generate three configurations: $b_x = (1\ 0\ 0\ 1\ 0\ 0 \ldots)$, $b_y = (0\ 1\ 0\ 0\ 1\ 0 \ldots)$ and $b_z = (0\ 0\ 1\ 0\ 0\ 1 \ldots)$. Note that $\{b_x, b_y, b_z\}$ form a feasible collection of configurations. One can think of configuration dependent rules "encrypting" any given sequence into a feasible collection of configurations and feeding this collection into a configuration selection function which then selects one configuration out of the ones fed into it. This selected configuration is then "decrypted" into an alternative which is the final choice. We formally define configuration dependent rules using a configuration selection function as follows.

Definition 17. *A choice rule d is a configuration dependent rule if there exists a configuration selection function f such that $d(S) = x$ if and only if $f(B(S)) = x(S)$ for all $S \in X^{\mathbb{N}}$.*

Configuration dependent rules are characterized by a neutrality axiom. It states that if a sequence is "transformed" into a new sequence by relabeling the alternatives, then the choice from the new sequence must respect this transformation. In other words, the choice rule is "neutral" with respect to the identity of the alternatives. It is formally stated as follows.

Neutrality: A choice rule d satisfies neutrality if for any bijection $\sigma : X \to X$ and $S, S' \in X^{\mathbb{N}}$,

$$[S'(i) = \sigma(S(i)) \; \forall i \in \mathbb{N}] \implies [d(S') = \sigma(d(S))]$$

Theorem 6. *A choice rule is a configuration dependent rule if and only if it satisfies Neutrality.*

We now introduce a sub-class of configuration rules that require the DM to maximize a preference order $\succ$ over the set of all configurations. We call such rules *rational* configuration dependent rules. For any sequence, the DM looks at the corresponding feasible collection of configurations and chooses the alternative that corresponds to the $\succ$-maximal configuration. This procedure can be formally defined as follows

Definition 18. *A choice rule d is a rational configuration dependent rule if there exists a preference order $\succ$ over $\{0,1\}^{\mathbb{N}}$ such that for all $S \in X^{\mathbb{N}}$*

$$d(S) = \{x : x(S) \succ y(S) \text{ for all } y \neq x, y = S(i) \text{ for some } i \in \mathbb{N}\}$$

Note that the example discussed above corresponds to a special kind of preference relation over configurations. As mentioned above, configuration dependent rules can be used to describe behavior where the information about location of alternatives can be used to make choices. Rational configuration-dependent rules, in particular, are useful to this effect. For instance, consider a DM that always picks the second alternative from a sequence. Let O_1 and O_2 form a partition of $\{0,1\}^{\mathbb{N}}$ where $O_1 \coloneqq \{a \in \{0,1\}^{\mathbb{N}} : a(2) = 1\}$ i.e. the set of all configurations that have 1 at its second position and $O_2 \coloneqq \{0,1\}^{\mathbb{N}} \setminus O_1$. Then such behavior can be explained as a rational configuration dependent rule by any preference order $\succ$ over $\{0,1\}^{\mathbb{N}}$ with $b \succ b'$ for any $b \in O_1$ and any $b' \in O_2$.

Rational configuration-dependent rules are characterized using a condition that resembles the well-known Strong Axiom of Revealed Preference (SARP) due to Houthakker

(1950). To state this condition, we need the notion of an equivalence relation between two sequences with respect to an alternative.

Definition 19. *For any $x \in X$, let $\sim_x \in X^{\mathbb{N}} \times X^{\mathbb{N}}$ such that $S \sim_x S'$ if and only if $S(i) = x \implies S'(i) = x$ for all $i \in \mathbb{N}$.*

The above defined binary relation says that two sequences are related via the relation $\sim_x$ if the configuration of the alternative x is the same for both. To illustrate, consider $S, S' \in X^{\mathbb{N}}$ such that $S = (a\ b\ x\ a\ b\ x \ldots)$ and $S = (a\ c\ x\ a\ c\ x \ldots)$.That is, S consists of repeating "cycle" of a, b and x whereas S' consists of a repeating "cycle" of a, c and x. Then we say $S \sim_x S'$. It is easy to see that $\sim_x$ is an equivalence relation. Now, we are ready to state our condition.

Acylicity: A choice rule d satisfies acyclicity if for any $x_1, x_2 \ldots x_n \in X$ and $S_1, S_2 \ldots S_n \in X^{\mathbb{N}}$ such that $S_j \sim_{x_{j+1}} S_{j+1}$ for all $j \in \{1, \ldots, n-1\}$ and $S_n \sim_{x_1} S_1$

$$[d(S_1) = x_1, \ldots, d(S_{n-1}) = x_{n-1}] \implies [d(S_n) \neq x_n]$$

The choice of x over y in a sequence reveals that x's configuration is directly "revealed preferred" over y's configuration. A chain of such direct revelations constitutes an indirect revealed preference. The acyclicity condition says that if a configuration is directly or indirectly revealed preferred to another configuration, then the converse cannot hold. Now, we show that neutrality and acylicity characterize rational configuration dependent rules.

Theorem 7. *A choice rule d is a rational configuration-dependent rule if and only if it satisfies Neutrality and Acyclicity.*

Before ending this section, we would like to make two comments. First, while comparing configurations, a DM may use the sum of discounted values or the limiting frequency of the configurations as a criteria to generate a ranking over them. In such cases, the DM can possibly declare indifference between multiple configurations. With a minor adaptation by using the primitive as a weak order[13] in place of a preference order, our choice rules can be extended to such cases where the DM can potential choose a set of maximal alternatives rather than a single alternative. Second, while the choice rules discussed in this section may not be computable, they can be studied within the domain of

[13]A weak order is a reflexive, complete and transitive binary relation but not necessarily antisymmetric.

computable decision rules as well. A suitable modification of the neutrality and acylicity conditions characterize configuration dependent (and rational configuration dependent) rules that are also computable.

1.7 Related Literature

The idea that a DM may observe alternatives in the form of a list i.e. an ordered set was first formalized in choice theory by Rubinstein and Salant (2006). Based on their framework, a variety of models incorporating order effects in choice have been introduced in the literature (see for instance Horan (2010), Guney (2014) and Dimitrov et al. (2016)). Satisficing, first introduced by Simon (1955) has been an influential idea in the choice theoretic literature and there have been several adaptations of it. The list setup provides a natural framework to study satisficing behavior. Kovach and Ülkü (2020) introduce one such model. In their model, the DM makes her choice in two stages. In the first stage, she searches through the list till she sees k alternatives. In the second stage, she chooses from the alternatives she has seen. Another adaptation of satisficing was introduced in Manzini et al. (2019). Their model is interpreted as one of *approval* as against choice. Since our framework is a generalization of lists, satisficing heuristics are a natural choice of study and we introduced two satisficing procedures in Section 5. However, it is worth pointing out that certain heuristics such as the "last satisficing" rule i.e. choosing the "last" alternative of the list in case no satisfactory alternative appears in the list cannot be formulated in our setup.

Computational aspects of decision-making have been an area of interest to economic theorists. As has been pointed out in Richter and Wong (1999)), computability-based economic theories also provide foundations for complexity analysis. The idea of bounded rationality has been closely linked to computational limitations of an economic agent (see Futia (1977)). For instance, in the analysis of repeated games, the strategies of players are implemented using finite automata (Rubinstein (1986)). Further, this implementation of strategies is also assumed to be costly in the state complexity of the automata implementing them (Abreu and Rubinstein (1988)). In the context of finitely repeated games, the imposition of bounds on the complexity of strategies of the players is shown to justify cooperation (see Neyman (1985)).

The discussion of the role of computational constraints in individual decision-making goes back to Simon (1955). He remarks "…limits on computational capacity may be important constraints entering into the definition of rational choice under particular circumstances". To capture the finite informational processing capacity, Kramer (1967) models the DM as a finite automaton and shows the incongruence of rational decision-making given this behavioral restriction on the DM. On the other hand, Salant (2003) shows that with the finite automaton model of decision-making, implementation of rational choice functions over menus (sets of alternatives) is computationally efficient. He further shows that implementation of other choice functions is much more computationally demanding.

While modeling DMs using an automaton has been a popular approach to model aspects of bounded rationality, to model the computational aspects of decision-making, the Turing machine is a more appropriate device. It is a more powerful model of computation than a finite automaton and can be thought of as a precise way of mathematically describing an algorithm. Turing machines embody the idea of computability in the truest sense. Richter and Wong (1999) use the idea of a Turing machine to define computable preferences and show that computable preferences have computable utility representations. Camara (2021) models the DM as a Turing machine in the environment of decision-making under risk. He introduces the notion of computational tractability. A decision problem is *intractable* if it cannot be implemented by an algorithm in a "reasonable" amount of time. He shows that expected utility maximization is intractable unless the utility function satisfies a strong separability property.

As discussed above, an application of our model is when alternatives come in the form of streams of recommendations. The idea that recommendations influence choices has been widely accepted. Cheung and Masatlioglu (2021) have introduced a model of decision-making under recommendation. However, in their setup, the decision maker observes sets of alternatives and hence is different from our setup. Our object of interest, infinite sequences, in the context of choices has been previously studied by Caplin and Dean (2011). Our model differs from their model in terms of incorporating sequences in the domain of choice functions whereas they enrich the observable choice data by incorporating sequences as the output of the choice function and interpret these sequences as provisional choices of the DM with contemplation time.

1.8 Concluding Remarks

In this paper, we introduced a new model of decision-making that considers infinite sequences as primitives as against sets or finite lists. This model provides a natural setting to study decision-making situations where the DM faces alternatives sequentially and also provides a generalization of the framework on choice over lists introduced in Rubinstein and Salant (2006). Further, our model allows a study of situations where the decision to stop examining alternatives is completely endogenous to the DM. To that end, we introduced a natural subclass of decision rules called stopping rules that require the DM to decide after viewing a finite segment of an every sequence. Our main result —the reduction lemma —showed that the class of stopping rules is equivalent to its seemingly stricter subclass—that of uniform-stopping rules. We introduced the notion of computability of a decision rule using Turing machines and showed that any computable decision rule can be implemented by a finite automaton —a result that does not hold in the setup of decisions over finite lists.

The reduction lemma allows us to develop a language to formulate *testable* conditions for studying different choice procedures. This involves defining the informational concepts of sufficiency and minimal sufficiency of finite segments in decision-making. With a dynamic representation of stopping rules in the form of decision procedures, we showed that these segments are completely identifiable in practice. To demonstrate the applicability of our model, we introduced some natural choice procedures and provided their behavioral characterizations. Future work will involve studying stochastic choice rules and examining potential applications of our main result.

1.9 Appendix

1.9.1 *A generalized reduction lemma*

We prove a result on *stopping times*.[14] A map $\tau : X^{\mathbb{N}} \to \mathbb{N} \cup \{\infty\}$ is a stopping time if it satisfies the following consistency property: for all $S, S' \in X^{\mathbb{N}}$,

$$[S|_{\tau(S)} = S'|_{\tau(S)}] \implies [\tau(S) = \tau(S')]$$

Observe that in the proof of the reduction lemma, the function k_d is a stopping time in the sense of the definition given above. It is immediate from the reduction lemma that if $\tau(S) < \infty$ for all $S \in X^{\mathbb{N}}$, then $\sup\{\tau(S) : S \in X^{\mathbb{N}}\} < \infty$. We extend this result to any sequence of stopping times that is pointwise bounded which is formally defined as follows.

Definition 20. *A sequence of stopping times $(\tau_i)_{i\in\mathbb{N}}$ is pointwise bounded if* $\sup\{\tau_i(S) : i \in \mathbb{N}\} < \infty$.

Analogous to a uniform stopping rule, we can define a sequence of stopping times that has a uniform bound.

Definition 21. *A sequence of stopping times $(\tau_i)_{i\in\mathbb{N}}$ has a uniform bound if* $\sup\{\tau_i(S) : S \in X^{\mathbb{N}},\ i \in \mathbb{N}\} < \infty$.

While it is immediate by definition that any sequence of stopping times that has a uniform bound is pointwise bounded, we now show that the converse is also true.

Theorem 8. *A sequence of stopping times $(\tau_i)_{i\in\mathbb{N}}$ that is pointwise bounded has a uniform bound.*

Proof. Consider a sequence of stopping times $(\tau_i)_{i\in\mathbb{N}}$ that is pointwise bounded. Now, define $\tau^* : X^{\mathbb{N}} \to \mathbb{N}$ as $\tau^*(S) = \sup\{\tau_i(S) : i \in \mathbb{N}\}$. First, we observe that for every $S \in X^{\mathbb{N}}$, $\tau^*(S) = \tau_j(S)$ for some $j \in \mathbb{N}$. That is, the supremum of the set $\{\tau_i(S) : i \in \mathbb{N}\}$ is $\tau_j(S)$. Now, we show that τ^* is a stopping time. Consider any arbitrary $S, S' \in X^{\mathbb{N}}$ such that

$$S|_{\tau^*(S)} = S'|_{\tau^*(S)}$$

[14] A definition of stopping times in probability theory requires a filtration which is a totally ordered collection of σ-algebras. As it turns out, with respect to a natural filtration that is defined using the sequence of projection maps $(\pi_i)_{i\in\mathbb{N}}$ (where $\pi_i : X^{\mathbb{N}} \to X$), that definition coincides with our definition.

Notice that since $\tau^*(S) = \sup\{\tau_i(S) : i \in \mathbb{N}\} = N$ for some $N \in \mathbb{N}$ and $\tau^*(S) = \tau_j(S)$ for some $j \in \mathbb{N}$, we have $\tau_k(S) \leq N$ for all $k \in \mathbb{N}$. This implies that for all $k \in \mathbb{N}$, we have

$$S|_{\tau_k(S)} = S'|_{\tau_k(S)}$$

Therefore, we must have $\tau_k(S) = \tau_k(S')$ for all $k \in \mathbb{N}$ implying $\sup\{\tau_i(S) : i \in \mathbb{N}\} = \sup\{\tau_i(S') : i \in \mathbb{N}\}$. Since $\sup\{\tau_i(S') : i \in \mathbb{N}\} = \tau^*(S')$, we get $\tau^*(S) = \tau^*(S')$. Since S and S' were chosen arbitrarily, we have shown that τ^* is a stopping time. Further, for every $S \in X^{\mathbb{N}}$, $\tau^*(S) < \infty$ and by the argument in the proof of the reduction lemma, we know that $\sup\{\tau^*(S) : S \in X^{\mathbb{N}}\} < \infty$. Therefore, it follows that the sequence of stopping times $(\tau_i)_{i \in \mathbb{N}}$ has a uniform bound. $\square$

1.9.2 *Omitted Proofs*

Proof of Theorem 3

First, we describe the structure of product topology on $X^{\mathbb{N}}$ where X is endowed with the discrete topology. We know that $\Pi_{X^{\mathbb{N}}}$, the product topology, is the smallest topology with respect to which the projection maps are continuous. Consider any map $M : \{1, \ldots, N\} \to X$ where $N \in \mathbb{N}$ and define the set $B(M)$ as:

$$B(M) = \{S \in X^{\mathbb{N}} : \text{ for all } i \in \{1, \ldots, N\},\ S(i) = M(i)\}$$

Let $\mathcal{B}_{X^{\mathbb{N}}}$ be the class of all such sets. Note that a for any $N \in \mathbb{N}$, the number of possible maps $M : \{1, \ldots N\} \to X$ is $|X|^N$. These sets are what can be interpreted as "open balls" in $X^{\mathbb{N}}$. Let $\mathscr{T}_{X^{\mathbb{N}}}$ be the class of unions of arbitrary subcollections of $\mathcal{B}_{X^{\mathbb{N}}}$.

Lemma **3.** *$\mathscr{T}_{X^{\mathbb{N}}}$ is the product topology on $X^{\mathbb{N}}$*

Proof. First, we show that $\mathscr{T}_{X^{\mathbb{N}}}$ is indeed a topology over $X^{\mathbb{N}}$. Notice that $\mathscr{T}_{X^{\mathbb{N}}}$ is closed under arbitrary unions by definition. To show that it is closed under finite intersections, let $\bigcap_{i=1}^{K} B_i$ be a finite intersection such that $B_i \in \mathscr{T}_{X^{\mathbb{N}}}$ for all $i \in \{1, \ldots K\}$. Note that each B_i is a union of some subcollection of $\mathcal{B}_{X^{\mathbb{N}}}$ and therefore we can write $B_i = \bigcup_{j_i \in \mathcal{J}_i} B_i^{j_i}$, with $\mathcal{J}_i$ being some indexed set, where each $B_i^{j_i}$ corresponds to an "open ball" i.e. is a set of the form $B(M)$ for some $M : \{1, \ldots, N\} \to X$ and $N \in \mathbb{N}$. Using the definition of $B(M)$, we know that there exist sets $A_1^i, A_2^i \ldots$ with $A_j^i \subseteq X$ for all $j \in \mathbb{N}$ such that

$$B_i = \{S \in X^{\mathbb{N}} : \text{ for all } j \in \mathbb{N}, S(j) \in A_j^i\}$$

So, we can write $\bigcap_{i=1}^{K} B_i$ as

$$\bigcap_{i=1}^{K} B_i = \{S \in X^{\mathbb{N}} : \text{ for all } j \in \mathbb{N}, S(j) \in \bigcap_{i=1}^{K} A_j^i\}$$

Clearly, $\bigcap_{i=1}^{K} B_i = \bigcup_{M \in \mathcal{M}_i} B(M)$ for some collection of maps $\mathcal{M}_i$. Therefore, $\mathscr{T}_{X^{\mathbb{N}}}$ is closed under finite intersection. Finally, $\mathscr{T}_{X^{\mathbb{N}}}$ contains $X^{\mathbb{N}}$ and $\varnothing$ as its elements. That $\varnothing \in \mathscr{T}_{X^{\mathbb{N}}}$ holds follows from the fact that $\varnothing$ is the union of elements from the empty subcollection of $\mathcal{B}_{X^{\mathbb{N}}}$. Further, $X^{\mathbb{N}}$ is the union of elements from the full collection $\mathcal{B}_{X^{\mathbb{N}}}$. Thus, $\mathscr{T}_{X^{\mathbb{N}}}$ is a topology over $X^{\mathbb{N}}$.

Now, we argue: $\Pi_{X^{\mathbb{N}}} \subseteq \mathscr{T}_{X^{\mathbb{N}}}$. For this, fix an arbitrary $i_* \in \mathbb{N}$ and $A \subseteq X$. If $A = \varnothing$, then $\pi_{i_*}^{-1}(A) = \varnothing$. As $\varnothing \in \mathscr{T}_{X^{\mathbb{N}}}$, $\pi_{i_*}^{-1}(A) \in \mathscr{T}_{X^{\mathbb{N}}}$ if $A = \varnothing$. However, if $A \neq \varnothing$, then observe:

$$\pi_{i_*}^{-1}(A) = \bigcup \Big\{ B(M) : M \in X^{\{1,\ldots,i_*\}} \ ; \ M(i_*) \in A \Big\}.$$

Thus, if $A \neq \varnothing$, then $\pi_{i*}^{-1}(A) \in \mathscr{T}_{X^{\mathbb{N}}}$. That is, $\pi_i^{-1}(A) \in \mathscr{T}_{X^{\mathbb{N}}}$ for every $A \subseteq X$. Hence, $\{\pi_i^{-1}(A) : i \in \mathbb{N} \ ; \ A \subseteq X\} \subseteq \mathscr{T}_{X^{\mathbb{N}}}$ and we have already shown that $\mathscr{T}_{X^{\mathbb{N}}}$ is a topology over $X^{\mathbb{N}}$. Further, by definition, $\Pi_{X^{\mathbb{N}}}$ is the smallest topology that satisfies $\{\pi_i^{-1}(A) : i \in \mathbb{N} \ ; \ A \subseteq X\} \in \Pi_{X^{\mathbb{N}}}$. Therefore, we obtain: $\Pi_{X^{\mathbb{N}}} \subseteq \mathscr{T}_{X^{\mathbb{N}}}$.

Finally, we argue: $\mathscr{T}_{X^{\mathbb{N}}} \subseteq \Pi_{X^{\mathbb{N}}}$. For this, fix an arbitrary $I \in \mathbb{N}$ and consider an arbitrary map $M : \{1, \ldots, I\} \to X$. For each $i \in \{1, \ldots, I\}$, let $A_i \coloneqq \{M(i)\}$. Then, we have the following:

$$B(M) = \bigcap \Big\{ \pi_i^{-1}(A_i) : i = 1, \ldots, I \Big\}.$$

Since $\Pi_{X^{\mathbb{N}}}$ is a topology and $\{\pi_i^{-1}(A) : i \in \mathbb{N} \ ; \ A \subseteq X\} \subseteq \Pi_{X^{\mathbb{N}}}$, it follows that $B(M) \in \Pi_{X^{\mathbb{N}}}$. Thus, $\Pi_{X^{\mathbb{N}}}$ is a topology over $X^{\mathbb{N}}$ such that $\mathcal{B}_{X^{\mathbb{N}}} \subseteq \Pi_{X^{\mathbb{N}}}$. Moreover, $\mathscr{T}_{X^{\mathbb{N}}}$ is the smallest topology over $X^{\mathbb{N}}$ such that $\mathcal{B}_{X^{\mathbb{N}}} \subseteq \mathscr{T}_{X^{\mathbb{N}}}$. Hence, we conclude: $\mathscr{T}_{X^{\mathbb{N}}} \subseteq \Pi_{X^{\mathbb{N}}}$.

□

Now to show d is a stopping rule if and only if it is continuous, first, assume that $d : X^{\mathbb{N}} \to X$ is continuous. Fix an arbitrary $S_* \in X^{\mathbb{N}}$ and let $y_{S_*} \coloneqq d(S_*)$. Now, we know that $\{y_{S_*}\}$ is open in the discrete topology over X. By continuity of the map d, the following set:

$$d^{-1}(\{y_{S_*}\}) \coloneqq \{S \in X^{\mathbb{N}} : d(S) = y_{S_*}\}$$

satisfies $d^{-1}(\{y_{S_*}\}) \in \Pi_{X^{\mathbb{N}}}$. By the lemma above and the definition of $\mathscr{T}_{X^{\mathbb{N}}}$, there exists $M : \{1, \ldots, k\} \to X$ such that $S_* \in B(M) \subseteq d^{-1}(\{y_{S_*}\})$. Now, $S_* \in B(M)$ implies: $M = S_*|_k$ and $B(M) = \{S_*|_k \cdot T : T \in X^{\mathbb{N}}\}$. Since $B(M) \subseteq d^{-1}(\{y_{S_*}\})$, it follows: $d(S_*|_k \cdot T) = y_{S_*}$ for all $T \in X^{\mathbb{N}}$. Since $y_{S_*} = d(S_*)$ and S_* was arbitrary, we have established: if the map $d : X^{\mathbb{N}} \to X$ is continuous, then it is a stopping rule.

Now, assume that $d : X^{\mathbb{N}} \to X$ is a stopping rule. Since X has the topology 2^X, we must argue that $d^{-1}(A) \coloneqq \{S \in X^{\mathbb{N}} : d(S) \in A\} \in \Pi_{X^{\mathbb{N}}}$ for any $A \subseteq X$. Since d^{-1} preserves arbitrary unions and $\Pi_{X^{\mathbb{N}}}$ is closed under arbitrary unions, it is enough to argue that $d^{-1}(\{y\}) \in \Pi_{X^{\mathbb{N}}}$ for any $y \in X$. So, fix an arbitrary $y_* \in X$. If $d^{-1}(\{y_*\}) = \varnothing$, then we have nothing more to argue as $\varnothing \in \Pi_{X^{\mathbb{N}}}$. Hence, assume that $d^{-1}(\{y_*\}) \neq \varnothing$. Consider an arbitrary $S_* \in d^{-1}(\{y_*\})$. Since d is a stopping rule, there exists $k(S_*) \in \mathbb{N}$ such that: $d(S) = y_*$ for every $S \in B(S_*|_{k(S_*)})$. This is because

$B(S_*|_{k(S_*)}) = \{S_*|_{k(S_*)} \cdot T : T \in X^{\mathbb{N}}\}$. Thus, we have:

$$\bigcup \Big\{ B(S|_{k(S)} : S \in d^{-1}(\{y_*\})) \Big\} = d^{-1}(\{y_*\}).$$

Hence, $d^{-1}(\{y_*\}) \in \mathscr{T}_{X^{\mathbb{N}}}$ by definition of $\mathscr{T}_{X^{\mathbb{N}}}$. By the lemma above, it follows that $d^{-1}(\{y_*\}) \in \Pi_{X^{\mathbb{N}}}$. Since $y_* \in Y$ was arbitrary, we have: $d^{-1}(A) \in \Pi_{X^{\mathbb{N}}}$ for any $A \in 2^X$. Thus, if the $d : X^{\mathbb{N}} \to X$ is a stopping rule, then it is continuous. $\square$

Proof of Theorem 4

We first prove the necessity. Suppose d is a Cardinal Satisficing Rule with $v \in \mathbb{R}_+$ and $w : X \to \mathbb{R}_{++}$. Consider an arbitrary $S \in X^{\mathbb{N}}$ with $d(S) = x$. Then there exists $N_1 \in \mathbb{N}$ such that $W_S^{N_1}(x) \geq v > W_S^{N_1}(y)$ for all $y \neq x$. Consider any $S' \in \mathcal{FS}(S, x)$ where S' and S differ on k and $k+1^{\text{th}}$ position for some $k \in \mathbb{N}$. Suppose $N_1 = k+1$ and $S(k) \neq x$ (and $S(k+1) = x$). In this case, we get $W_{S'}^{N_1-1}(x) \geq v > W_{S'}^{N_1-1}(y)$ for all $y \neq x$ and we have $d(S') = x$. For all other cases, we have $W_{S'}^{N_1}(x) \geq v > W_{S'}^{N_1}(y)$ and therefore we get $d(S') = x$. A similar argument holds for any $S' \in \mathcal{FD}(S, x)$. Therefore $d(S) = d(S')$ for all $S' \in \mathcal{F}(S, x)$. Since S was arbitrary, we have shown that d satisfies Monotonicity. To show that d satisfies Informational Dominance, consider a minimal sufficient segment M with $d(M) = x$ and a sufficient segment N such that $d(N) = y \neq x$ and $x \notin N$. Assume for contradiction that $d([M'.N]) = x$ for some segment M' such that $M \triangleright M'$. Since M' is not sufficient, the cumulative weight of x in M' is less than v. Since $x \notin N$, the cumulative weight of x in $[M' \cdot N]$ is less than v, a contradiction. Therefore $x \neq d([M' \cdot N])$.

Now, we prove the sufficiency. Let d satisfy Monotonicity and Informational Dominance. First, we construct the "revealed" critical frequency of each alternative. Fix $x \in X$. Note, by definition of a choice rule, $d(S^x) = x$ for the constant sequence $S^x = (x, x \ldots)$. Since d is a stopping rule, there exists $k \in \mathbb{N}$ such that $d(S^x) = d([S^x|_k] \cdot T)$ for all $T \in X^{\mathbb{N}}$. Let $n_x := \inf\{k \in \mathbb{N} : d(S^x) = d([S^x|_k] \cdot T)\ \forall T \in X^{\mathbb{N}}\}$. Since $\mathbb{N}$ is well-ordered, we know that $n_x \in \mathbb{N}$. Consider an arbitrary sequence S with $d(S) = x$. For any $i \in \mathbb{N}$ and the segment $S|_i$ of S, denote by $\#x(S|_i)$ the number of appearances of x in it. That is, $\#x(S|_i) := |\{j \in [i] : [S|_i](j) = x\}|$. Now, denote by $i(S, a)$ the position at which an alternative a reaches n_a appearances in S. That is, $i(S, a) := \{i \in \mathbb{N} : \#a(S|_i) = n_a\}$. In case alternative a does not reach n_a appearances in S, let $i(S, a) = \infty$ with the convention that $n < \infty$ for all $n \in \mathbb{N}$.

We show that $i(S, x) < i(S, y)$ for all $y \neq x$. Assume for contradiction that $i(S, y) < i(S, x)$ for some $y \neq x$. Let S' be a sequence generated from S by deleting all the first terms in the first $i(S, x)$ positions that are not equal to x or y. That is, S' is generated by finitely many favorable deletions with respect to x and y. By Monotonicity, we have $d(S') = x$. Note that first $i(S', x)$ terms contain n number of y's and n_x number of x's ($n + n_x = i(S', x)$) where $n \geq n_y$. Now, consider finitely many favourable shifts of S' with respect to x to generate S'' such that its first n_x terms are all x followed by n terms that are y. Again, by Monotonicity, we have $d(S'') = x$. Denote this initial segment of

x's as $[M \cdot x]$ where M is $(n_x - 1)$ long segment of x's and the n long segment of y's as N. So, we can write $S'' = [Mx \cdot N] \cdot T$ where $T \in X^{\mathbb{N}}$ and $T(j) = S'(i(S', x) + j)$ for all $j \in \mathbb{N}$. By the definition of n_x we know that there exists some $T \in X^{\mathbb{N}}$ such that $d(M \cdot T) \neq x$. Also, by the definition of n_y, we know that $d(N.T) = y$ for all $T \in X^{\mathbb{N}}$. In other words, Mx is a minimal sufficient segment, M is not a minimal sufficient segment and N is a sufficient segment. Using Informational Dominance, we know that $d([M \cdot N]) \neq x$. It must be that $d([M \cdot N \cdot x \cdot T]) = y$ for all $T \in X^{\mathbb{N}}$. Suppose not i.e. $d(M \cdot N \cdot x \cdot T) = z$ for some $z \neq x, y$ and $T \in X^{\mathbb{N}}$. Then, by Monotonicity, it must be that $d(NxT) = z$, a contradiction since N contains n_y first y's. Therefore $d(M \cdot N \cdot x \cdot T) = y$ for all $T \in X^{\mathbb{N}}$. Now, notice that we can generate the sequence S' by successively moving y's to the left i.e. a finitely many favourable shifts with respect to y and again, by Monotonicity, we have $d(S') = y$, a contradiction. Therefore, we get $i(S, x) < i(S, y)$.

Now, we define $w(x) \coloneqq \frac{1}{n_x}$ for all $x \in X$ and let $v = 1$. Consider the computable choice rule d^* such that $d^*(S) = \{x : W_S^N(x) \geq v > W_S^N(y)\}$ for all $S \in X^{\mathbb{N}}$. Note that since $w(x) > 0$, d^* is indeed a computable choice rule. We will show that $d^* = d$. Consider any arbitrary $S \in X^{\mathbb{N}}$ and let $d(S) = z$. We know that $i(S, z) < i(S, y)$ for all $y \neq z$. Let $i(S, z) = N$. By construction, we know that $W_S^N(z) = 1$ and $W_S^N(z) \geq v > W_S^N(y)$ for all $y \neq z$ and therefore $d^*(S) = z$. Since S was chosen arbitrarily, we have shown that $d^* = d$. □

Proof of Theorem 5

We will first show the necessity part. Suppose d is OSR with $(\succ, a^*, k)$. First, note that $D = U(a^*)$ and length of a minimal sufficient segment M is k if $M \subset D'$. Also, the length of any minimal sufficient segments is not greater than k. Since $\succ$ is a preference order, d satisfies Sequential-α and Sequential-NBC. To show that d satisfies Replacement, consider any arbitrary $M \in \mathscr{MS}_{D'}$ and $M' \in \mathcal{S}_m$ such that $M(i) = x$, $M'(i) = y$ for some $i \in [m]$ and $M(j) = M'(j)$ for all $j \neq i$. W.L.O.G let z be the $\succ$-maximal alternative in M'. For any $S \in X^{\mathbb{N}}$, we know that $d(M' \cdot S) = z$ and therefore $M \in \mathscr{S}$

Now, to show the sufficiency, suppose d satisfies Sequential-α, Sequential-NBC and Replacement. We will proceed in several steps as follows:

Step 1: First, we show that $|M| = |M'|$[15] for any $M, M' \in \mathscr{MS}_{D'}$. Suppose not. W.L.O.G let $|M| > |M'|$. Consider the segment $M|_{|M'|}$. Since X is finite, we can reach from M' to $M|_{|M'|}$ in finite number of "steps" of replacement i.e. there exists a chain of segments $M_1, \ldots, M_n$ with $M_1 = M'$ and $M_n = M|_{|M'|}$ such that $|\{i : M_j(i) \neq M_{j+1}(i)\}| = 1$ for all $j \in \{1, \ldots n-1\}$. In other words, M_j and M_{j+1} differ only in one position for all segments in the chain. By repeated application of Replacement, we know all the segments in the chain are sufficient and therefore $M|_{|M'|} \in \mathscr{S}$. Since $|M_{|M'|}| < |M|$ and $[M|_{|M'|}](i) = M(i)$ for all i, this is a contradiction to $M \in \mathscr{MS}$. We have established that all minimal sufficient segments that do not contain any decisive alternatives are of the same length. Let that length be denoted by $i^{D'}$.

Step 2: Consider an arbitrary $M \in \mathscr{MS}_{\mathscr{D}}$ and let $i^D := \inf\{i \in \mathbb{N} : M(i) \in D\}$. That is, i^D denotes the location of first occurence of a decisive alternative in M. We will show that $|M| \leq i^{D'}$ and $|M| = i^D$. Suppose not i.e. there exists a $M \in \mathscr{MS}_D$ such that $|M| > i^{D'}$. Consider an arbitrary $M' \in \mathscr{MS}_{D'}$. By the previous step, we know that $|M'| = i^{D'}$. Since X is finite, as in the previous step, consider a chain of segments $M_1 \ldots, M_n$ such that $M_1 = M'$ and $M_n = M|_{|M'|}$ such that every successive element in the chain differs by an alternative in exactly one position. By repeated application of Replacement, we know that M_n is a sufficient segment. Since $|M_n| < M$ and $M_n(i) = M(i)$ for all i, this is a contradiction to $M \in \mathscr{MS}$. Therefore, $|M| \leq i^{D'}$. Now, we will show that $|M| = i^D$. Assume for contradiction that $|M| > i^D$ (note that the argument for the case $|M| < i^D$ is trivial by the definition of $\mathscr{MS}_D$). W.L.O.G, let $M(i^D) = x$. By the definition of D, we know that $d(M) = x$. Since $M \in \mathscr{MS}$,

[15]We abuse notation and denote the length of the segment M by $|M|$.

there exists a sequence T such that $d([M|_{|M|-1} \cdot T]) \neq d(M) = x$. Let $\bar{M}$ be the minimal sufficient segment of the sequence $[M|_{|M|-1} \cdot T]$. By the definition of D, $|\bar{M}| < i^D$. Since $|\bar{M}| < |M|$ and $\bar{M}(i) = M(i)$ for all i, we have a contradiction to $M \in \mathscr{MS}$. Therefore, we must have $|M| = i^D$.

Step 3: Since d is a computable rule (and therefore a stopping rule), it is completely specified by the choices on its minimal sufficient segments. By the previous two steps, we know that the length of any minimal sufficient segment is at most $i^{D'}$. Let $k = i^{D'}$. There are two possible cases:
(i) $k = 1$. This implies that size of all minimal sufficient segment is 1. Further, this implies that every alternative is decisive i.e. $D = X$. Consider an arbitrary preference order $\succ$ on X and let $a^* := \min(X, \succ)$.[16] It is easy to see that $(k, \succ, a^*)$ rationalize d.
(ii) $k \geq 2$. Consider any $x, y \in D'$. Define $\succ$ as follows: $x \succ y$ iff there exists $M \in \mathscr{MS}_{D'}$ such that $x, y \in M$ and $d(M) = x$. We first show that $\succ$ is a preference order over D'. Reflexivity follows from the definition. Assume for contradiction that $\succ$ is not antisymmetric. Then there exists distinct $x, y \in D'$ such that $x \succ y$ and $y \succ x$. By definition, there exists $M, M' \in \mathscr{MS}_{D'}$ such that $x, y \in M$, $x, y \in M$, $d(M) = x$ and $d(M') = y$ Consider $M'' \in \mathscr{MS}_{\mathscr{D}'}$ such that $x, y \in M''$ and $z \notin M''$ for all $z \neq x, y$. That is, M'' consists of only x and y (such M'' exists due to the assumption that $k \geq 2$ and step 1). By Sequential-α, we have $d(M'') = x$ and $d(M'') = y$, a contradiction. Therefore $\succ$ is antisymmetric. Now, consider any distinct $x, y \in D'$ and $M \in \mathscr{MS}_{D'}$ such that $x, y \in M$ and $z \notin M$ for all $z \neq x, y$. By definition of $\succ$ and antisymmetry we have either $x \succ y$ or $y \succ x$. Therefore $\succ$ is complete. To show $\succ$ is transitive, consider $x, y, z \in D'$ and suppose $x \succ y$ and $y \succ z$. We consider two cases:

(a) $k = 2$. We know that there exists $M, M' \in \mathscr{MS}_{D'}$ such that $y \in M$, $d(M) = x$ and $z \in M'$, $d(M') = y$. Consider $M'' \in \mathscr{MS}_{D'}$ such that $x, z \in M''$. Since $|M''| = 2$, by Sequential-NBC, we know that $d(M'') \neq z$, implying $d(M'') = x$ giving us $x \succ z$.

(b) $k > 2$. Consider $M \in \mathscr{MS}_{D'}$ such that $x, y, z \in M$ and $w \notin M$ for all $w \neq x, y, z$. By Sequential-α, we know that $d(M) \neq z$ and $d(M) \neq y$. Therefore $d(M) = x$ implying $x \succ z$.

Step 4: We have shown that $\succ$ is a preference order over D'. Now, we show that D is non-empty. Assume for contradiction that D is empty i.e. $X = D'$. By step 1, all

[16] $\min(X, \succ) := \{x \in X : y \succ x \ \forall y \in X\}$

minimal sufficient segments are of the same length. Since $\succ$ is a preference order over X, we have a unique maximal element. W.L.O.G let it be x. Consider an arbitrary minimal sufficient segment M such that $x \in M$. Since $\succ$ is antisymmetric, we know that $d(M) = x$. Therefore, by definition of D, we must have $x \in D$, a contradiction. Now, consider an arbitrary preference order $\bar{\succ}$ on X such that $\succ \subset \bar{\succ}$ and $x \mathrel{\bar{\succ}} y$ for all $x \in D$ and $y \in D'$. Let $a^* = \min(D, \bar{\succ})$. Now, we will show that $(k, \bar{\succ}, a^*)$ rationalize d. Consider an arbitrary $S \in X^{\mathbb{N}}$ with $d(S) = x$. There are two possible cases: (i) The segment of S_k does not contain any alternative from D. That is $S|_k \in \mathscr{MS}_{D'}$. Suppose there exists $y \in S|_k$ with $y \neq x$ such that $y \succ x$. Then we have a contradiction to the antisymmetry of $\succ$. Therefore, by completeness of $\succ$ we have $x \succ y$ for all $y \in S|_k$ (ii) The segment $S|_k$ contains at least one alternative from D. That is $M \in \mathscr{MS}_D$ for some $S|_k \triangleright M$. By step 2, we must have $x \in D$ and x is the first alternative from D to feature in $S|_k$ and we are done. $\square$

Proof of Theorem 6

Suppose d is a configuration dependent rule. Then there exists an f such that $d(S) = x \iff f(B(S)) = x(S)$. Consider any S, S' such that $d(S) = x$ and for all $i \in \mathbb{N}$, $S'(i) = \sigma(S(i))$. We know that $B(S) = B(S')$ and $\sigma(x)(S') = x(S)$. Therefore $f(B(S')) = \sigma(x)(S')$ which implies $d(S') = \sigma(x) = \sigma(d(S))$.

To show the other direction, consider a choice rule d that satisfies Neutrality. Define the relation $\sim_\sigma$ as follows: $S \sim_\sigma S'$ if and only if there exists a bijection $\sigma : X \to X$ such that $S'(i) = \sigma(S(i)) \;\; \forall i \in \mathbb{N}$. Note that $\sim_\sigma$ is an equivalence relation and hence partitions $X^{\mathbb{N}}$. Now, consider any arbitrary $S \in X^{\mathbb{N}}$ such that $d(S) = y$ for some $y \in X$. Define f as $f(B(S)) = y(S)$. Consider any S' such that $S \sim_\sigma S'$, for some bijection $\sigma : X \to X$. By Neutrality, we know that $d(S') = \sigma(d(S))$. Also, since $B(S) = B(S')$ and $y(S) = \sigma(y)(S')$, we have $f(B(S')) = \sigma(y)(S')$, by construction. For $B \in \mathscr{B}$ such that $B \neq B(S)$ for any $S \in X^{\mathbb{N}}$, define f arbitrarily. Hence we have defined an f such that $d(S) = x$ if and only if $f(B(S)) = x(S)$ for all $S \in X^{\mathbb{N}}$. Therefore, d is a configuration dependent rule. $\square$

Proof of Theorem 7

Necessity is straightforward, so we prove the sufficiency. Define the following "revealed" relation over configurations $\succ^c$ as follows: For any $a, b \in \{0,1\}^{\mathbb{N}}$, $a \succ^c b$ iff there exists $S \in X^{\mathbb{N}}$ with $x(S) = a$ and $y(S) = b$ for some $x, y, \in X$ and $d(S) = x$. Reflixivity of $\succ^c$ is immediate from its definition. Now, we show that $\succ^c$ is antisymmetric. Suppose not, then there exist distinct $a, b \in \{0,1\}^{\mathbb{N}}$ such that $a \succ^c b$ and $b \succ^c a$, i.e. there exist $S, S' \in X^{\mathbb{N}}$ and $x, y, w, z \in X$ with $x(S) = w(S') = a$, $d(S) = x$ and $y(S) = z(S') = b$, $d(S') = z$. There are four possible cases:

(i) $x = w$ and $y = z$. Note that since $S \sim_y S'$ and $S' \sim_x S$, by Acylicity, we have $d(S') \neq y$, a contradiction.

(ii) $x \neq w$ and $y = z$. Define a bijection $\sigma : X \to X$ to be such that $\sigma(x) = w$ and $\sigma(x') = x'$ for all $x' \neq x$. Let S'' be such that $S''(i) = \sigma(S(i))$ for all $i \in \mathbb{N}$. By Neutrality, we get $d(S'') = w$. Since $S'' \sim_y S'$ and $S' \sim_w S''$. By Acylicity, we have $d(S') \neq y = z$, a contradiction.

(iii) $x = w$ and $y \neq z$. Define a bijection $\sigma : X \to X$ to be such that $\sigma(z) = y$ and $\sigma(z') = z'$ for all $z' \neq z$. Let S'' be such that $S''(i) = \sigma(S'(i))$ for all $i \in \mathbb{N}$. By Neutrality, we have $d(S'') = y$ and that $S'' \sim_y S$ and $S \sim_x S''$, by Acylicity, we have $d(S) \neq x$, a contradiction.

(iv) $x \neq w$ and $y \neq z$. Let $\sigma : X \to X$ be such that $\sigma(z) = y$, $\sigma(w) = x$ and $\sigma(x') = x'$ for all $x' \neq z, w$. Let S'' be such that $S''(i) = \sigma(S'(i))$ for all $i \in \mathbb{N}$. By Neutrality, we have $d(S'') = y$ and since $S \sim_y S''$ and $S'' \sim_x S$,by Acylicity, we have $d(S) \neq x$, a contradiction.

Similarly, using Acyclicity and Neurality, we can show that the $\succ^c$ is also acyclic. Now, using $\succ^c$, define an indirect "revealed" relation $\succ^i$ over $\{0,1\}^{\mathbb{N}}$ as follows: For any $a, b \in \{0,1\}^{\mathbb{N}}$, $a \succ^i b$ iff there exists a chain of alternatives $a_1, \ldots a_n \in \{0,1\}^{\mathbb{N}}$ with $a = a_1$ and $a_n = b$ such that $a_1 \succ^c \ldots \succ^c a_n$. Note that since $\succ^c$ is acyclic, the indirect relation $\succ^i$ is antisymmetric and transitive i.e. a partial order. By Szpilrajn's lemma [17], there exists a preference order over $\{0,1\}^{\mathbb{N}}$ such that $\succ^i \subseteq \succ$. Consider an arbitrary such extension and denote it as $\succ$. Define $\tilde{d} : X^{\mathbb{N}} \to X$ as follows:

$$\tilde{d}(S) = \{x : x(S) \succ y(S) \ \ \forall y \ \ \text{with } y = S(i) \text{ for some } i \in \mathbb{N}\}$$

Now, consider an arbitrary S and let $d(S) = x$. We know that $x(S) \succ^c y(S)$ for all y such that $S(i) = y$ for some $i \in \mathbb{N}$. Since we know $\succ^c \subseteq \succ^i \subseteq \succ$, we have that $d(S) = \tilde{d}(S)$. Since S was chosen arbitrarily, we have shown that d is a rational configuration dependent rule. $\square$

[17] Szpilrajn's lemma states that a partial order over a set can be "extended" to a preference order.

1.9.3 *A topological approach to the reduction lemma*

In this section, we discuss a topological approach to our main result. Having already introduced the product topology on the set of all inputs $X^{\mathbb{N}}$, we now explicitly describe convergence with respect to the product topology. The following definition is standard.

Definition 3: *Suppose, $(Z, \mathscr{T})$ is a topological space. A Z–valued sequence (z_n)* converges *in $\mathscr{T}$ to z_* if, for every $U \in \mathscr{T}$ with $z_* \in U$, there exists $n_U \in \mathbb{N}$ such that $z_n \in U$ for all $n \geq n_U$.*

The phrase "(z_n) converges in $\mathscr{T}$ to z_*" shall often be abbreviated as "$z_n \longrightarrow z_*$ in $\mathscr{T}$". In particular, the meaning of any X–valued sequence (S) converges in the product topology $\Pi_{X^{\mathbb{N}}}$ to some $S_* \in X^{\mathbb{N}}$ stands specified. Now, consider the following proposition.

Proposition **2.** *Let (S_n) be an X–valued sequence and $S_* \in X^{\mathbb{N}}$. Then, $S_n \longrightarrow S_*$ in $\Pi_{X^{\mathbb{N}}}$, if and only if, for every $k \in \mathbb{N}$, there exists $n_k \in \mathbb{N}$ such that $S_n|_k = S_*|_k$ if $n \geq n_k$.*

Proof. First, we assume: $S_n \longrightarrow S_*$ in $\Pi_{X^{\mathbb{N}}}$. Fix an arbitrary $k \in \mathbb{N}$ and define $M : \{1, \ldots, k\} \to X$ as follows: $M(i) \coloneqq S_*(i)$ for all $i \in \{1, \ldots, k\}$; that is, $M = S_*|_k$. Now, $\Pi_{X^{\mathbb{N}}} = \mathscr{T}_{X^{\mathbb{N}}}$ by proposition 1 and recall $\mathcal{B}_{X^{\mathbb{N}}} \subseteq \mathscr{T}_{X^{\mathbb{N}}}$ by definition of $\mathscr{T}_{X^{\mathbb{N}}}$. Thus, $B(M) \in \Pi_{X^{\mathbb{N}}}$. Then, by definition 3, $S_n \longrightarrow S_*$ in $\Pi_{X^{\mathbb{N}}}$ implies: there exists $n_{B(M)} \in \mathbb{N}$ such that $S_n \in B(M)$ if $n \geq n_{B(M)}$. Let $n_k \coloneqq n_{B(M)}$. Finally, note that $S_n \in B(M)$ and $M = S_*|_k$ implies: $S_n|_k = S_*|_k$.

Now, we prove the converse. For this, consider an arbitrary $U \in \Pi_{X^{\mathbb{N}}}$ with $S_* \in U$. By proposition 1, $U \in \mathscr{T}_{X^{\mathbb{N}}}$. By definition of $\mathscr{T}_{X^{\mathbb{N}}}$, there exists $k \in \mathbb{N}$ and a map $M : \{1, \ldots, k\} \to X$ such that $S_* \in B(M) \subseteq U$. Also, there exists $n_k \in \mathbb{N}$ such that $S_n|_k = S_*|_k$ if $n \geq n_k$. However, $S_* \in B(M)$ implies $S_*|_k = M$. Thus, $S_n \in S_*|_k$ implies $S_n \in B(M)$. Therefore, $S_n \in B(M)$ for all $n \geq n_k$. Since $B(M) \subseteq U$, we have: $S_n \in U$ if $n \geq n_k$. Thus, let $n_U \coloneqq n_k$ to complete the proof. $\square$

In addition to the topology over the domain $X^{\mathbb{N}}$, we must formalize a certain continuity property associated with any stopping rule. For this, we associate to each decision rule $d : X^{\mathbb{N}} \to Y$ a natural map $k_d : X^{\mathbb{N}} \to \mathbb{N} \cup \{0, \infty\}$ which was defined as:

$$k_d(S) \coloneqq \inf\{k \in \mathbb{N} : (\forall T \in X^{\mathbb{N}})[d(S|_k \cdot T) = d(S)]\}$$

for all $S \in X^{\mathbb{N}}$. Assuming that $d : X^{\mathbb{N}} \to Y$ is not a constant function, the fact that it is a stopping rule is equivalent to asserting that k_d is $\mathbb{N}$–valued. Also, we consider the set $\mathbb{N}$ to be endowed with the discrete topology $2^{\mathbb{N}}$. Note, this topology is the restriction to $\mathbb{N}$ of the standard topology on $\mathbb{R}$.

With this background in place, we are ready to prove Theorem 1 via topological methods. The first of these proofs is as follows.

Proof 2. Since X is finite, the topology 2^X makes X compact. Thus, by Tychonoff's theorem, the product topology $\Pi_{X^{\mathbb{N}}}$ makes $X^{\mathbb{N}}$ compact. Since the continuous image of a compact set is compact, it follows that $k_d(X^{\mathbb{N}}) := \{k_d(S) : S \in X^{\mathbb{N}}\}$ is a compact subset of $\mathbb{N}$ endowed with the topology $2^{\mathbb{N}}$. Now, $2^{\mathbb{N}}$ is the restriction of the standard topology of $\mathbb{R}$ to $\mathbb{N}$. Thus, $k_d(X^{\mathbb{N}})$ is a compact subset of $\mathbb{R}$. Since compact subsets of $\mathbb{R}$ must be bounded, there exists $k_* \in \mathbb{N}$ such that $k_d(S) \leq k_*$ for all $S \in X^{\mathbb{N}}$. Thus, the decision rule $d : X^{\mathbb{N}} \to Y$ has a uniform stopping time. $\square$

While the above argument is short, it rests in large measure on the abstract result that continuous maps over compact set have a compact range. The following argument, however, removes the role of this result in furnishing a proof of Theorem 1.

Proof 3. Let $d : X^{\mathbb{N}} \to X$ be a non–constant stopping rule. Thus, the map k_d is $\mathbb{N}$–valued. For each $k \in \mathbb{N}$, let $\mathcal{A}_k := \{S \in X^{\mathbb{N}} : k_d(S) = k\}$. Since k_d is $\mathbb{N}$–valued, we have: $X^{\mathbb{N}} = \cup_{k \in \mathbb{N}} \mathcal{A}_k$. Then, $\{\mathcal{A}_k : k \in \mathbb{N}\}$ is an open cover of $X^{\mathbb{N}}$ if we can argue: $\mathcal{A}_k \in \Pi_{X^{\mathbb{N}}}$ for every $k \in \mathbb{N}$. For this, fix an arbitrary $k \in \mathbb{N}$ and observe:

$$\Big\{S \in X^{\mathbb{N}} : k_d(S) = k\Big\} = \bigcup \Big\{B(S|_k) : k_d(S) = k\Big\}.$$

Thus, $\mathcal{A}_k \in \mathscr{T}_{X^{\mathbb{N}}}$ by definition of $\mathscr{T}_{X^{\mathbb{N}}}$. Then, proposition 1 implies that $\mathcal{A}_k \in \Pi_{X^{\mathbb{N}}}$. That is, $\{\mathcal{A}_k : k \in \mathbb{N}\}$ is an open cover of $X^{\mathbb{N}}$ in the topology $\Pi_{X^{\mathbb{N}}}$. However, $\Pi_{X^{\mathbb{N}}}$ makes $X^{\mathbb{N}}$ compact by Tychonoff's theorem (see Munkres (1974)). Thus, there exists $L \in \mathbb{N}$ and $k_1 < \ldots < k_L$ such that:

$$X^{\mathbb{N}} = \bigcup_{l=1}^{L} \Big\{S \in X^{\mathbb{N}} : k_d(S) = k_l\Big\}.$$

Define $k_* := \max\{k_l : l = 1, \ldots, L\}$. Thus, $k_d(S) \leq k_*$ for all $S \in X^{\mathbb{N}}$. $\square$

A direct comparison of proofs 2 and 3 suggests that the essential ingredient for Theorem 1 to hold is the compactness of $X^{\mathbb{N}}$ under the product topology. This was the point of step 2 in the elementary proof (*i.e.*, proof 1). However, only the sequential compactness of $X^{\mathbb{N}}$ was established there which is weaker than compactness. Further, this was done by a direct argument rather than appealing to the theorem of Tychonoff. The next proof essentially casts the elementary proof via the compactness of various subsets of $X^{\mathbb{N}}$ under $\Pi_{X^{\mathbb{N}}}$.

Proof 4. Suppose, $d : X^{\mathbb{N}} \to Y$ is a stopping rule that does not have a uniform stopping time. Thus, for each $j \in \mathbb{N}$, there exists $k_j \in \mathbb{N}$ such that (1) $k_j < k_{j+1}$ for all $j \in \mathbb{N}$, and (2) the image of the map k_d, which is the set $k_d(X^{\mathbb{N}}) := \{k_d(S) : S \in X^{\mathbb{N}}\}$, is $\{k_j : j \in \mathbb{N}\}$. Then, for each $j \in \mathbb{N}$, define $\mathcal{A}_j := \{S \in X^{\mathbb{N}} : k_d(S) \geq k_j\}$. From (1) and (2), we have: $\mathcal{A}_j \neq \varnothing$ for every $j \in \mathbb{N}$. Further, $\mathcal{A}_{j+1} \subseteq \mathcal{A}_j$ for every $j \in \mathbb{N}$. Thus, $\{\mathcal{A}_j : j \in \mathbb{N}\}$ satisfies the finite–intersection property.

We now argue that, if $\mathcal{A}_j$ is compact for every $j \in \mathbb{N}$, then the claim of Theorem 1 holds. This is because the collection $\{\mathcal{A}_j : j \in \mathbb{N}\}$ satisfies the finite–intersection property. Thus, by Cantor's theorem:

$$\bigcap \Big\{\mathcal{A}_j : j \in \mathbb{N}\Big\} \neq \varnothing.$$

Thus, there exists $S_* \in X^{\mathbb{N}}$ such that $S_* \in \mathcal{A}_j$ for all $j \in \mathbb{N}$. Then, by definition of $\mathcal{A}_j$, we have: $k_d(S_*) \geq k_j$ for every $j \in \mathbb{N}$. Since $k_j < k_{j+1}$ and $k_j \in \mathbb{N}$ for all $j \in \mathbb{N}$, we have: $k_d(S_*) = \infty$. This contradicts the fact that d is a stopping rule. Thus, our supposition that the stopping rule d does not have a uniform stopping time must be wrong.

Therefore, it only remains to argue: $\mathcal{A}_j$ is compact for each $j \in \mathbb{N}$. For this, fix an arbitrary $j \in \mathbb{N}$. Notice, $\mathcal{A}_1 = X^{\mathbb{N}}$ which is compact under $\Pi_{X^{\mathbb{N}}}$ by Tychonoff's theorem. Hence, we assume $j \geq 2$. For any map $f : \{1, \dots, L\} \to X$ and $L_* \leq L$, we shall denote the map $l \in \{1, \dots, L_*\} \mapsto M(l)$ by $M|_{L_*}$. That is, $M|_{L_*}$ is the *truncation* of the map M at L_*. Now, define:

$$\mathcal{M}_j := \Big\{M \in X^{\{1,\dots,k_j\}} : \neg(\exists S \in X^{\mathbb{N}})[k_d(S) < k_j \,;\, S|_{k_d(S)} = M|_{k_d(S)}]\Big\}.$$

Thus, $\mathcal{M}_j$ is precisely the collection of maps $M : \{1, \dots, k_j\} \to X$ such that[18] $k_d(M \cdot T) \geq k_j$ for any $T \in X^{\mathbb{N}}$. That is:

$$\mathcal{A}_j = \bigcup \Big\{B(M) : M \in \mathcal{M}_j\Big\}.$$

Since X is finite, it follows that $\mathcal{M}_j$ is finite. Thus, to show that $\mathcal{A}_j$ is compact it is enough to argue that $B(M_*)$ is compact for any $M_* : \{1, \dots, K_*\} \to X$.[19] However, the map $f_{M_*} : X^{\mathbb{N}} \to B(M_*)$, defined by $f_{M_*}(T) := M_* \cdot T$ for all $T \in X^{\mathbb{N}}$, is a

[18]Recall, $M \cdot T$ is the concatenation of the map T to the map M.
[19]Here, $B(M_*)$ is endowed with the topology $\Pi_{X^{\mathbb{N}}} \cap B(M_*)$.

homeomorphism. To see why, it is enough to argue that the following holds:

$$A \in \Pi_{X^{\mathbb{N}}} \iff \{M_* \cdot T : T \in A\} \in \Pi_{X^{\mathbb{N}}} \cap B(M_*).$$

This follows from the fact that[20] $\{M_* \cdot T : T \in B(M)\} = B(M_* \cdot M)$ for any map $M : \{1, \ldots, K\} \to X$ and Lemma 1. Since $X^{\mathbb{N}}$ is compact and f_{M_*} is a homeomorphism from $X^{\mathbb{N}}$ to $B(M_*)$, it follows that $B(M_*)$ is compact. Thus, the set $\mathcal{A}_j$ is compact for each $j \in \mathbb{N}$. $\square$

[20]For $M_* : \{1, \ldots, K_*\} \to X$ and $M : \{1, \ldots, K\} \to X$, the map $M_* \cdot M : \{1, \ldots, K_* + K\} \to X$ is defined by: $[M_* \cdot M](k) \coloneqq M_*(k)$ if $k \in \{1, \ldots, K_*\}$; otherwise, $[M_* \cdot M](k) \coloneqq M(k - K_*)$.

Chapter 2

An Axiomatic Analysis of Rejection Behavior

2.1 Introduction

2.1.1 *An overview*

It is well-known that framing of a decision problem can play a role in the decision-making process of an individual. While rational choice theory fails to accommodate these effects, there is a growing literature on bounded rationality that tries to analyze the effects of framing. Some of the "frames" that are observed to affect behavior are the order in which alternatives are presented, default alternatives, repeated presentation etc. Another experimentally well-established instance of differing behavior resulting due to the framing of a decision problem is when a decision maker (DM) is asked to reject vis-á-vis when she is asked to choose. While psychologists have suggested several mechanisms to explain these differences in behavior, this paper provides a choice theoretic foundation to this framing effect using a standard tool of choice theory: binary relations.

As is standard in much of economic theory, decision problems are formulated as "choice" problems —individual or collective. In such problems, the DM is often assumed to maximize some binary relation and select the "best" alternatives according to it—the maximal set. We posit that in the case of "rejection" problems, the DM eliminates the "worst" alternatives. The notion of worst alternatives in a decision problem according to an underlying binary relation is captured by what we term as the *minimal* set. In our formulation, the minimal set is *not* equal to the set of non-maximal alternatives and this results in observed behavior that is different from the one generated via maximization. Such behavior is consistent with some of the experimental findings in the psychology literature. In particular, this formulation of rejection behavior results in (weakly) larger selections than that generated by maximization.

It is often the case that the DM has to make unique selections in decision problems. A combination of rejection followed by maximization seems plausible in many such settings. Consider the following example that illustrates one such sequential procedure:

Example 1. A committee wishes to hire an applicant from a fixed set of applicants. Each member of the committee has a ranking over the set of applicants based on the submitted applications. The Pareto relation[1] generated by this collection of rankings is used to shortlist the set of applicants for the interview round in the following way: any applicant that is beaten by some applicant using the Pareto relation and does not beat any applicant is rejected. From the shortlisted set of applicants, the committee selects the best according to a common ranking generated by the performances of applicants in the interview.

The above example shows that rejection can form the basis of shortlisting or consideration before a final choice is made. This consideration set is what we term a *rejection filter*. While having discussed the rejection behavior procedurally, we also provide an axiomatic foundation of it. Our axiomatization of the rejection filter is related to the axiomatic characterization of the maximal set correspondences by Sen (1971). The choice of maximal set is characterized by a contraction condition (α) and an expansion condition (γ). We find that a weakening of the contraction condition and a strengthening of the expansion condition together with a binary consistency condition characterizes rejection behavior as defined above. We then contrast it with the *attention filter* and the *competition filter* of Masatlioglu et al. (2012) and Lleras et al. (2017) respectively that are defined axiomatically.

We formalize the idea of the above example and introduce a two-stage choice procedure—*Choice by Rejection* (CBR). In this procedure the DM uses a rejection filter to shortlist alternatives in the first stage, followed by maximization using a complete rationale (defined below) in the second stage leading to a unique choice. In order to study the empirical content of this procedure, we follow a revealed preference approach. In case of the unobservable first stage, often instances of violation of standard rational-

[1]The Pareto or the unanimity relation is a partial order i.e. a transitive binary relation. It can be generated by a collection of rankings or linear orders by taking their intersection.

ity conditions provide some insight into the first stage behavior. For instance, in the case of the Choice with Limited Attention (CLA) model of Masatlioglu et al. (2012), a "reversal" in choice, say from an alternative x to z upon removal of an another alternative y reveals attention of the DM for y. Similarly, Horan (2016), in his analysis of the Rational Shortlist Method (RSM) and Transitive Shortlist Method (TSM) defines certain reversals in choice that reveal information about the first stage shortlisting and put restrictions on the underlying rationales. We follow a similar approach and define two types of choice reversals between a pair of alternatives —*weak* and *strong* reversals. These reversals form the basis of our characterization of the CBR model.

We require four conditions on the observable choices to characterize the CBR choice function. The first condition, termed *Never Chosen*, requires an alternative that is chosen in a menu to be chosen in *at least* one pairwise comparison with other alternatives in that menu. Equivalently, it prohibits the choice of an alternative in a menu if it is not chosen in any pairwise comparison. The second condition, termed *Weakened Contraction Consistency*, ensures that for any strong or weak reversal between a pair of alternatives, we can associate a third alternative that "causes" that reversal. The third and the fourth conditions are on a revealed preference relation defined using strong and weak reversals. This revealed preference relation captures all the first stage information revealed by choices to the outside observer. The third condition is an acylicity condition—called R-acyclicity —that requires the revealed relation to be acyclic and a subset of the revealed pairwise relation. The fourth condition is a congruence condition à la Richter (1966) and Tyson (2013) which essentially requires that if an alternative is revealed to be not rejected *and* not chosen in a menu, then it cannot be chosen in another menu where the chosen alternative of the previous menu is revealed to be not rejected.

We introduce a variant of this procedure where in addition to the completeness of the second rationale, we impose the condition of transitivity. We characterize this variant using the above mentioned conditions and an adaptation of the congruence condition to incorporate transitivity of the second rationale. The conditions that we provide may lack normative appeal. However, we argue that the plausibility of the procedure and the normative appeal of axiomatic foundation of the rejection filter are sufficient grounds to study the *testability* of this procedure.

We then address the question of identification of the underlying rationales. If the first stage is observable, then the underlying rationales are fully identifiable. However, as is common with most of the two-stage procedures with an unobservable first stage, the identification of the underlying rationales is partial in this model. We identify the common parts of the revealed rationales and show that these are identified using the reversals. Further, we show that the CBR model enjoys a small menu property: any two CBR-representable choice functions that "agree" on small menus —of size 2 and 3—must coincide.

There are many two-stage models in the literature that are able to provide explanations for the empirically observed phenomenon such as *compromise* effect and *decoy* effects. The natural extensions of these effects are the *two-compromise* and *two-decoy* effects. CBR provides a novel explanation for these observed phenomenon which many two-stage models are unable to[2]. We relate our model to the RSM model of Manzini and Mariotti (2007) and TSM of Horan (2016). We observe that the class of CBR and RSM choice functions are not nested. However, there is a non-empty intersection and we examine what extra conditions pin down RSM-representable and TSM-representable choice functions within the class of CBR choice functions.

Using the idea of rejection filters, we introduce a simple model of stochastic choice along the lines of Echenique and Saito (2019). In this model, the DM first rejects the worst alternatives using a binary relation and then follows a Luce procedure to assign probabilities to the non-rejected alternatives. The characterization of this procedure is done by a straightforward adaption of the axioms characterizing the rejection filter to the stochastic choice setup and the *Cyclical Independence* condition of Echenique and Saito (2019).

2.1.2 *Related literature and outline*

The idea that rejection frames affect decision-making is well established in the psychology literature. Huber et al. (1987) found experimentally that rejection behavior produces larger consideration sets as compared to choice behavior whereas Sokolova and Krishna (2016) posit that a DM resorts to a more delibrative process in the "rejection-type" tasks vis-à-vis "choice-type" tasks. This, they show, is reflected in the attenuation of decision biases across the two types of decision problems. Our model provides a choice theoretic foundation for this procedural difference between choice and rejection behavior. In economic theory, framing effects have been studied by Salant and Rubinstein (2008). In addition to framing effects resulting in a procedural difference in decision-making, we

[2]While the Choice with Limited Attention (CLA) model of Masatlioglu et al. (2012) can explain these observed choice patterns, models like Rational Shortlist Method (RSM) of Manzini and Mariotti (2007) and TSM of Horan (2016) are unable to explain them.

believe that there are certain situations that are "weaker" than choice and the idea of rejection in itself suggests natural way to make decisions. These situations include can for instance, observable online behavior like wishlisting, adding items to cart, favouriting, liking etc. that are termed as "approval" in Manzini et al. (2019) and Wang (2022).

The idea that decision-making involves some procedure of elimination goes back to at least Tversky (1972) who developed a probabilistic theory of choice based on a process of elimination. More recently, Masatlioglu and Nakajima (2007) introduced a deterministic theory of choice that is based on elimination of alternatives. In their model, an alternative is eliminated only if it is dominated by another alternative in its "comparable" set and hence differs from our model. Apesteguia and Ballester (2013) proposed a procedure that involves sequential pairwise elimination of "disliked" alternatives until only one alternative remains.

The classical notion of rationality has been synonymous with preference order maximization and the early literature in choice theory has a thorough treatment of a decision-making via preference maximization (see Houthakker (1950), Chernoff (1954) Arrow (1959), Sen (1971)). However, with mounting evidence against it in the form of observed phenomenon like cyclic choices, choice reversals via the *decoy effect, compromise effect* etc. (see Huber et al. (1982), Simonson (1989)), the literature has seen emergence of models of bounded rationality have tried to explain this observed choice behavior (see Manzini and Mariotti (2007), Cherepanov et al. (2013), Masatlioglu et al. (2012) etc). Many of these procedures are characterized by some weakening of the *Weak axiom of revealed preference*(WARP). One of the most well-known weakening of WARP is Weak-WARP (WWARP), first introduced in Manzini and Mariotti (2007) [3]. While these models are able to accomodate the above mentioned effects, many of them cannot accommodate their natural extensions: *two-compromise effect* and *two-decoy effect*. Such behavior has been observed in different experimental settings (see Tserenjigmid (2019), Manzini and Mariotti (2010), Teppan and Felfernig (2009)). This is because such effects are observed as "double" reversals: choices of the form $C(xy) = x$, $C(A) = y$ and $C(A') = x$ for some $\{x, y\} \subset A \subset A'$ (where C is a choice function that will be defined formally below). Our two-stage model provides a novel explanation for such effects.

The outline of the paper is as follows: in the next section we introduce the model and discuss rejection filters. In Section 2.3, we introduce the two-stage procedure and provide

[3]Some of the models which directly use WWARP to characterize their models are Manzini and Mariotti (2007), Manzini and Mariotti (2012), Lombardi (2009b), Cherepanov et al. (2013), Ehlers and Sprumont (2008)

its behavioral characterization. In Section 2.4, we examine the problem of identification of the parameters of our model. Section 2.5 provides a further discussion of the model and its relation to the literature. Section 2.6 introduces a two-stage stochastic choice model and provides its behavioral characterization. Section 2.7 concludes. The proofs omitted in the main body of the paper are relegated to the Appendix.

2.2 Rejection Filters

2.2.1 *Preliminaries*

Throughout the paper, we denote by X a finite non-empty set of alternatives. Let $\mathcal{P}(X)$ denote the set of all non-empty subsets of alternatives. A menu A is an element of $\mathcal{P}(X)$. A binary relation R over X is a subset of $X \times X$. For any $x, y \in X$, $(x, y) \in R$ is can also be written as xRy. A binary relation is asymmetric if for any $x, y \in X$, xRy implies $\neg yRx$. We call an asymmetric binary relation a rationale. A binary relation R is complete if for any distinct $x, y \in X$, either xRy or yRx. It is transitive if for any distinct $x, y, z \in X$, xRy and yRz implies xRz. We say that x is R-unrelated to y if $\neg xRy$ and $\neg yRx$. An alternative is R-unrelated to a set S if it is R-unrelated to all $y \in S$. In some places, whenever no confusion arises, we abuse notation and write a set, say $\{x, y, z\}$, as xyz.

For any $A \in \mathcal{P}(X)$ and a rationale R, the maximal set is denoted by $\max(A, R) \coloneqq \{x \in A : \neg yRx \; \forall \; y \in A\}$. We denote its minimal set by

$$\min(A, R) \coloneqq \{x \in A : \neg xRy \; \forall y \in A \text{ and } \exists \; y \in A \text{ such that } yRx\}$$

It is important to note that the although for a given set, the minimal set is disjoint from the maximal set, it is *not* the relative complement of the maximal set. For an alternative to be in the minimal set, it must be beaten by some alternative and it should not beat any alternative. While on the other hand, an alternative is in the maximal set if it is not beaten by any alternative. Therefore, for a given rationale R and a set A, the relative complement of the minimal set can contain the maximal set strictly (see Example 2 below).

A consideration set mapping is a map $\Gamma : \mathcal{P}(X) \to \mathcal{P}(X)$ such that $\Gamma(A) \subset A$ for all $A \in \mathcal{P}(X)$.[4]

[4]Such maps are also called choice correspondences in the literature. The term "consideration set" comes from the marketing literature and was first coined by Wright and Barbour (1977).

2.2.2 *Rejection Filters*

Now, we formalize the idea of a rejection procedure using minimal sets. The DM is endowed with a transitive rationale and for any menu, she rejects its minimal set and considers only its non-minimal alternatives. The consideration set mapping thus generated is what we term as a *rejection filter*.[5] Formally, it is defined as follows

Definition 1. *A consideration set mapping* Γ *is a rejection filter if there exists a transitive rationale* R *such that* $\Gamma(A) = A \setminus \min(A, R)$ *for all* A.

Two prominent consideration set mappings that have been introduced in the literature on bounded rationality are the *attention* and the *competition* filters (see Masatlioglu et al. (2012) and Lleras et al. (2017)). A consideration map is an attention filter if it satisfies the following property: For all A and $x \in A$ such that $x \notin \Gamma(A)$,

$$\Gamma(A \setminus \{x\}) = \Gamma(A)$$

The defining property of attention filters is interpreted as: if an alternative that is not considered (not paid to attention to) is made unavailable, then the set of alternatives considered should not change. On the other hand a consideration map is a competition filter if it satisfies condition the well-known α (see Sen (1971)). That is, for any $A, A' \in \mathcal{P}(X)$ with $\{x\} \subset A' \subset A$,

$$[x \in \Gamma(A)] \implies [x \in \Gamma(A')]$$

This property is interpreted as: if an alternative is considered in a larger menu, then it must be considered in a smaller menu as well. Using a simple example, we show that a rejection filter is different than both attention and competition filters as it violates both their defining properties.

Example 2: Let $X = \{w, x, y, z\}$ and $R = \{(x, y), (y, z), (x, z)\}$. Then the corresponding rejection filter Γ is given in Table 1.

While different than the attention and competition filters, the rejection filter generated by a binary relation relates to the maximal set of the same binary relation in the following way: it always contains the maximal set. Further, upon "iterative" application

[5]A more general rejection filter can be defined by dropping the transitivity assumption. We will study a two-stage choice procedure using such filters in Section 6. However, for the present section and the next section on a two-stage choice procedure, we assume that a rejection filter is generated by a transitive rationale.

A	$\Gamma(A)$	A	$\Gamma(A)$	A	$\Gamma(A)$
$\{x,y\}$	$\{x\}$	$\{x,y,z\}$	$\{x,y\}$	$\{x,y,z,w\}$	$\{w,x,y\}$
$\{x,z\}$	$\{x\}$	$\{x,y,w\}$	$\{x,w\}$		
$\{x,w\}$	$\{x,w\}$	$\{x,z,w\}$	$\{x,w\}$		
$\{y,z\}$	$\{y\}$	$\{y,z,w\}$	$\{y,w\}$		
$\{y,w\}$	$\{y,w\}$				
$\{z,w\}$	$\{z,w\}$				

Table 1: A Rejection Filter

—one can arrive at the maximal set. That is, successive removal of the minimal sets from a menu $(\Gamma(\Gamma\ldots(\Gamma(A)))$ results in its maximal set, thus giving a procedural description of selecting the maximal set. We analyze such a rejection procedure of iterative elimination in a companion paper.

2.2.3 *Axiomatic Foundations*

In addition to a procedural description of the rejection filter, we can formulate it axiomatically as well. We use three axioms to characterize the rejection filter. The first two are variants of the axioms characterizing maximal-element rationalizability using acyclic binary relations. Sen (1971) showed that choice of the maximal set is pinned down by two conditions —α and γ. The first condition introduced above in the definition of competition filters, is also known as contraction consistency. It says that if an alternative is selected in a given menu, then it must be selected in a sub-menu as well if it is feasible. We propose the following weakening of it:

Weak Contraction: A consideration set mapping Γ satisfies Weak Contraction if for any $A \in \mathcal{P}(X)$ and $x, y \in A$ such that $y \in \Gamma(xy)$,

$$[x \in \Gamma(A)] \implies [x \in \Gamma(A \setminus \{y\})]$$

This axiom has the interpretation that if an alternative is not rejected from a menu, then it should continue to be not rejected upon the removal of a (weakly) better alternative. There are two possible cases: (i) $\Gamma(\{x,y\}) = y$, where we know that x is rejected in binary comparison with y. Then, our axiom requires if x is not rejected in the presence of an alternative that "dominates" it, then it should still be not rejected upon the removal of this "dominating" alternative. (ii) $\Gamma(\{x,y\}) = \{x,y\}$. Here, we know that neither alternative rejects the other. Then, our axiom requires that if x is not rejected in a menu, it should still be not rejected upon the removal of this "incomparable" alternative. It can be easily observed that this is a weakening of condition α. The second axiom that Sen (1971) used is also called expansion consistency and it requires that if an alternative is selected in the two decision problems, then it must be selected in their union also. We propose the following strengthening of it:

Strict Expansion: A consideration set mapping Γ satisfies Strict Expansion if the following statements hold

(i) For any $A, A' \in \mathcal{P}(X)$, if $x \in \Gamma(A) \cap \Gamma(A')$, then $x \in \Gamma(A \cup A')$; and

(ii) For any $A \in \mathcal{P}(X)$ (with $|A| \geq 2$), if $\{x\} = \Gamma(A)$, then $x \in \Gamma(A')$ for all $A' \in \mathcal{P}(X)$ such that $A \subset A'$.

This axiom is consists of two parts. The first one is says that if an alternative is not rejected in two menus, then it must not be rejected in their union as well. The second part says that if in some menu an alternative is the unique selection i.e. it rejects every other alternative, then it cannot be rejected in any larger menu (in terms of set inclusion). This axiom requires the same conclusion as that of condition γ from a weaker premise and therefore is a strengthening of it. The final axiom that we require is a binary consistency condition that effectively requires rejections to be transitive.

Binary Rejection Consistency: A consideration set mapping Γ satisfies Binary Rejection Consistency if for any distinct $x, y, z \in X$, if $x = \Gamma(xy)$ and $y = \Gamma(yz)$, then $x = \Gamma(xz)$.

Using the above three axioms, we provide a characterization of the rejection filter.[6]

Theorem 1. *A consideration set mapping is a rejection filter if and only if it satisfies Weak Contraction, Strict Expansion and Binary Rejection Consistency*

Proof. To show the necessity of the axioms, consider a transitive rationale R and let Γ_R be defined as $\Gamma_R(A) \coloneqq A \setminus \min(A, R)$ for all $A \in \mathcal{P}(X)$.

(i) Weak Contraction: Consider an arbitrary $A \in \mathcal{P}(X)$, $x \in \Gamma_R(A)$ and $y \in A$ such that $y \in \Gamma_R(\{x, y\})$. If yRx, then since $x \in \Gamma_R(A)$, we have xRz for some $z \in A \setminus \{y\}$, implying $x \in \Gamma_R(A \setminus \{y\})$. If y is R-unrelated to x, then we have $x \notin \min(A \setminus \{y\}, R)$ implying $x \in \Gamma_R(A \setminus \{y\})$ and therefore Γ_R satisfies Weak Contraction.

[6] A more general characterization follows by dropping the condition of Binary Rejection Consistency. However, since our two stage model assumes transitivity of the first stage rationale, we provide a characterization with it.

(ii) Strict Expansion: Consider two arbitrary $A, A' \in \mathcal{P}(X)$ and $x \in X$ such that $x \in \Gamma_R(A) \cap \Gamma_R(A')$. If xRy for some $y \in A \cup A'$ then we know that $x \notin \min(A \cup A', R)$ and hence $x \in \Gamma_R(A \cup A')$. If $\neg yRx$ for all $y \in A \cup A'$, then $x \notin \min(A \cup A', R)$ implying $x \in \Gamma_R(A \cup A')$. Now suppose $x = \Gamma_R(A)$. Then we know xRy for some $y \in A$ and hence $x \in \Gamma_R(A \cup A')$. Therefore, Γ_R satisfies Strict Expansion.

(iii) Binary Rejection Consistency: this follows from the definition of Γ_R and the transitivity of R.

Now, we prove sufficiency of the axioms. Consider a Γ that satisfies the three axioms. Define R as follows: xRy if and only if $\Gamma(\{x, y\}) = x$. The relation is R is asymmetric by definition and transitive by Binary Rejection Consistency. Let Γ_R be defined as above. We will show $\Gamma = \Gamma_R$ using strong induction on the size of the menus. For the base case i.e. for all $A \in \mathcal{P}(X)$ such that $|A| = 2$, we have $\Gamma(A) = \Gamma_R(A)$ by definition. Now, fix $k \geq 2$ and assume that $\Gamma(A) = \Gamma_R(A)$ for all $A \in \mathcal{P}(X)$ such that $|A| \geq k$. Consider an arbitrary A such that $|A| = k + 1$.

First, we show that $\Gamma(A) \subseteq \Gamma_R(A)$. Consider some $x \in \Gamma(A)$. If $\neg yRx$ for all $y \in A$, we get $x \notin \min(A, R)$ and we are done. So, suppose yRx for some $y \in A$. Now, we show that there exists some $z \in A$ such that $x \in \Gamma(A \setminus \{z\})$. Suppose not i.e. $x \notin \Gamma(A \setminus \{z\})$ for all $z \in A \setminus \{x\}$ .Then, by Weak Contraction, we must have xRz for all $z \in A \setminus \{x\}$. By Strict Expansion we must have $x \in \Gamma(A \setminus \{z\})$ for any z, a contradiction. Therefore, there exists $z \in A$ such that $x \in \Gamma(A \setminus \{z\})$. There are two possible cases: (i) $y \neq z$. Then by our inductive hypothesis, since $x \notin \min(A \setminus \{z\}, R)$ and yRx, there exists $w \in A \setminus \{z\}$ such that xRw and we are done. (ii) $y = z$. Suppose $\neg xRz$ for all $z \in A \setminus \{y\}$. Pick and arbitrary $z \in A \setminus \{y\}$. Since $z \in \Gamma(\{x, z\})$, by Weak Contraction, we have $x \in \Gamma(A \setminus \{z\})$. By our inductive hypothesis and our supposition that yRx, there must exist some $w \in A \setminus \{z\}$ such that xRw, a contradiction. Therefore, there exists some $z \in A \setminus \{y\}$ such that xRz. So, we get $\Gamma(A) \subseteq \Gamma_R(A)$.

To show $\Gamma_R(A) \subseteq \Gamma(A)$, consider some $x \in \Gamma_R(A)$. There are two cases: (i) xRy for some $y \in A$ implying $x = \Gamma(\{x, y\})$ and by the second part of strict expansion, we get $x \in \Gamma(A)$. (ii) $\neg yRx$ for all $y \in A$ and by the repeated application of the first part of Strict Expansion, we get $x \in \Gamma(A)$. Therefore, we have $\Gamma_R(A) \subseteq \Gamma(A)$, which completes the proof. $\square$

2.3 A Two-stage Procedure

2.3.1 *CBR and reversals*

It is often the case that the DM has to make a unique choice in a decision problem and the analyst only observes the final choices i.e. a single-valued choice funtion. A choice function is any map $C : \mathcal{P}(X) \to X$ such that $C(A) \in A$ for all $A \in \mathcal{P}(X)$. We examine what testable implications does it have if the DM follows a two stage procedure where in the the first stage she uses a rejection filter to create a shortlist and then maximize a complete rationale on the shortlisted set to arrive at a unique choice. We call this procedure Choice by Rejection (CBR) and it is defined formally as follows.

Definition 2. *A choice function C is a Choice by Rejection (CBR) if there exists a rejection filter Γ and a complete rationale P such that for all $A \in \mathcal{P}(X)$,*

$$C(A) = \max(\Gamma(A), P)$$

Observable reversals in choice provide a succinct framework for analysis of boundedly rational models of choice. We borrow insights from the characterizations of the RSM and the TSM models [7] by Horan (2016). His characterizations involve consistency conditions which are expressed using different types of choice reversals. In a similar manner, we categorize inconsistencies in choices in terms of choice reversals. We define two mutually exclusive reversals that help analyze our model and provide basis for our characterization.

Consider three distinct alternatives x, y, z and a menu A such that $\{x, y\} \subseteq A$ and $z \notin A$. We say that the choice function C displays an (xy) reversal *due* to z on A if we observe the following choices:

$$C(xy) = C(A) = x, \quad C(A \cup \{z\}) = y$$

Note that A can be $\{x, y\}$ as well. Due to the addition of a third alternative z, the choice shifts from x to y. We categorize such (xy) reversals as *weak* or *strong* depending on whether reversal is due to an alternative which is either pairwise *dominated* or *dominates* x. Let $\succ_c$ denote the pairwise relation revealed by the choice function. That is, $x \succ_c y$ if and only if $C(xy) = x$. An (xy) reversal is a **weak** (xy) reversal due to z if $x \succ_c z$. This reversal is a weak reversal (due to z) in the sense that the introduction of an apparently "weak" alternative (z) shifts the choice from x to y. The second type

[7] TSM is a special case of the RSM model where both the rationales are transitive

of reversal is called a **strong** (xy) reversal due to z if $z \succ_c x$. This reversal is a strong reversal (due to z) as the introduction of an apparently "strong" alternative shifts the choice from x to y. By definition, if (xy) has a weak(strong) reversal due to z, then (xy) cannot have a strong(weak) reversal due to z.[8] We say that there is a reversal *in the presence of* x if it is already present in a menu on which a reversal happens. That is, for some $x \neq y, w$ and $A \ni x$, we have $C(A) = C(yw) = y$ and $C(A \cup \{z\}) = w$.

As it turns out, these reversals can provide us information about the first stage rationale. Intuitively, there are two ways a reversal can occur in CBR. The first way is when an alternative that is chosen in a menu is "pushed" into the minimal set upon the addition of an alternative to that menu. Consider a set A such that $x, y \in A$ and $z \notin A$. Suppose x is R-unrelated to A and xPw for all $w \in \Gamma(A)$ with respect to R. Then x is chosen in A. Further, we also have xPy. Now, if we have zRx, then the addition of z to A pushes x into the minimal set of $A \cup \{z\}$. Now, if yPw for all $w \in \Gamma(A \cup \{z\})$, it gets chosen. This is the underlying mechanism of a *strong* reversal.

Second, an alternative that is not chosen in a menu and is in the minimal set is "pulled" out of it by the addition of an alternative and is chosen. To see this, consider a set A such that $x, y \in A$ and $z \notin A$. Suppose xRy and $\neg yRw$ for all $w \in A \setminus \{x\}$. Here y is in the minimal set of A and is not shortlisted. If yRz, then addition of z to A leads to y being shortlisted in $A \cup \{z\}$. Further, if yPw for all $w \in \Gamma(A \cup \{z\})$, then y is chosen. In particular, we have yPx. This is the underlying mechanism of a *weak* reversal. Since R is assumed to be transitive, we get xRz (since xRy and yRz). Whereas, in the case of a strong reversal, we have zRx. Therefore, transitivity of R helps us differentiate between strong and weak reversal.

2.3.2 *A characterization*

In order to capture the first stage information revealed by reversals, we define a revealed relation $\triangleright_c$ on X such that $x \triangleright_c y$ if and only if there is a:

- **weak** (xy) reversal due to w for some $w \in X$ or;
- **weak** (wx) reversal due to y for some $w \in X$ or;
- **strong** (yw) reversal due to x for some $w \in X$

[8] Horan (2016) introduces Weak and Direct reversals in a similar spirit. A choice function C displays a Weak (xy) reversal on $B \supset \{x, y\}$ if $C(xy) = x$ and $C(B) \neq C(B \setminus \{y\})$. C displays a direct (xy) reversal on $B \subseteq X \setminus \{x\}$ if $C(B) = y$ and $C(B \cup \{x\}) \notin \{x, y\}$.

The following example illustrates instances of weak and strong reversals and the resulting $\triangleright_c$.

EXAMPLE 3: Let $X = \{x, y, z, w\}$ and the choice function C is depicted below:

A	C(A)	A	C(A)	A	C(A)
$\{x,y\}$	x	$\{x,y,z\}$	y	$\{x,y,z,w\}$	x
$\{x,z\}$	z	$\{x,y,w\}$	x		
$\{x,w\}$	x	$\{x,z,w\}$	x		
$\{y,z\}$	y	$\{y,z,w\}$	y		
$\{y,w\}$	y				
$\{z,w\}$	z				

It can be seen that in the choice function above, we have a (i) **strong** (xy) reversal due to z, and (ii) **weak** (zx) reversal due to w. The revealed relations $\succ_c$ and $\triangleright_c$ are represented below.

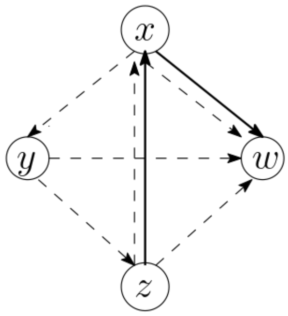


Figure 1: Dashed arrow indicates $\succ_c$ and solid arrow indicates $\triangleright_c$

An implication of CBR-representability is that these reversals imply reversals on "small" menus—menus of size 2 and 3 —as well, a result which we will prove later. This permits us to define $\triangleright_c$ solely based on choices from small menus. This is discussed in detail in Section 2.4. Now, we turn to a behavioral characterization of CBR.

We require four consistency conditions on choice functions to characterize CBR. The first condition requires that if an alternative in a collection is not chosen in the pairwise comparison with each of the other alternatives in that collection, then it cannot be chosen in the collection as well. This corresponds to the *condorcet loser* principle for choice correspondences (see Ehlers and Sprumont (2008)).

NEVER CHOSEN (NC): A choice function C satisfies NC if for any $A \in \mathcal{P}(X)$ and any $x \in A$,

$$[C(xy) \neq x \ \forall \ y \in A \setminus \{x\}] \implies [C(A) \neq x]$$

The second condition is a contraction consistency condition. It requires that if the choice from a set belongs to some pair of alternatives, then the choice from *some* set of one cardinality less should also belong to that pair of alternatives. This relates to the condition of Weak Contraction Consistency (see Ehlers and Sprumont (2008)) which requires any chosen alternative in a set to be chosen in *some* set of one cardinality less. Our condition can be seen as a weakening of it in the setting of single-valued choice functions.

Weakened Contraction Consistency (WCC): A choice function C satisfies WCC if for any $A \in \mathcal{P}(X)$ and $x, y \in A$,

$$[C(A) \in \{x, y\}] \implies [C(A \setminus \{z\}) \in \{x, y\} \text{ for some } z \in A \setminus \{x, y\}]$$

This condition is important in ensuring that when choice shifts from an alternative to another that was not previously chosen, it can be associated to the addition of an alternative. In other words, there are no "jump" shifts in choice such as the following: $C(xy) = x, C(xyz) = z, C(xyw) = w$ and $C(xyzw) = y$. Therefore, WCC ensures that if $C(A) \in \{x, y\}$ for some x, y, then there is a "path" from the set $\{x, y\}$ to the set A of x's and y's.

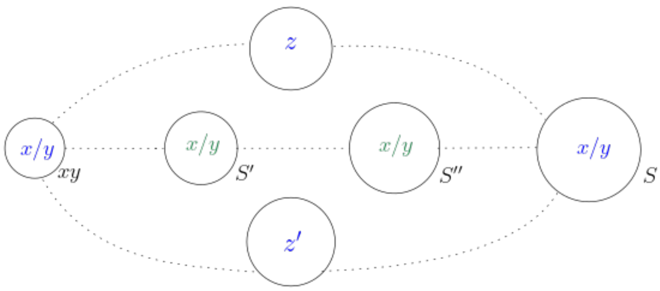


Figure 2: WCC ensures a "path" of x's and y's with $\{x, y\} \subset S' \subset S" \subset S$

The third condition that we require ensures that alternatives that are related via $\vartriangleright_c$ do not display cycles. It also ensures that $\vartriangleright_c$ is a subset of $\succ_c$.

Reject-Acyclicity (R-Acyclicity): A choice function C satisfies R-acyclicity if for any $x_1, \ldots x_n \in X$,

$$[x_1 \vartriangleright_c \ldots \vartriangleright_c x_n] \implies [\neg\, x_n \succ_c x_1]$$

The final and main condition is a "congruence" condition à la Richter (1966) and Tyson (2013). It requires that if an alternative x is chosen in a set in the presence of another alternative y where y is revealed to be not rejected, then y cannot be chosen in presence of x whenever x is revealed to be not rejected. The first stage domination is captured using the transitive closure of $\triangleright_c$ relation which is denoted by $\bar{\triangleright}_c$. An element x is revealed to be rejected in a set S if $x \in \min(\bar{\triangleright}_c, A)$. To define revealed shortlisted set of alternatives, let $U(A, \bar{\triangleright}_c) = A \setminus \min(A, \bar{\triangleright}_c) \subseteq A$. The condition below is a weakening of WARP which requires that $U(A, \bar{\triangleright}_c) = A$ for all $A \in \mathcal{P}(X)$.

Reject-WARP (R-WARP): For any x and y and $A\ A' \in \mathcal{P}(X)$ with $\{x, y\} \subseteq A, A'$ and $y \in U(A, \bar{\triangleright}_c)$,

$$[C(A) = x \text{ and } x \in U(A', \bar{\triangleright}_c)] \implies [C(A') \neq y]$$

If $y \in U(\bar{\triangleright}_c, A)$ and $C(A) = x$, then this reveals that x is preferred over y according to the second rationale. R-WARP ensures that when x is shortlisted in any other menu, y cannot be chosen. Now, we are ready to state our result, a proof of which is relegated to the Appendix.

Theorem 2. *A choice function C is a CBR if and only if it satisfies NC, WCC, R-Acyclicity and R-WARP.*

2.3.3 *Transitive CBR*

Now, we consider a special case where the second stage rationale is a strict linear order i.e. complete and transitive. We call it the Transitive-CBR (T-CBR) choice function. In order to incorporate the transitivity of the second rationale, we need to strengthen R-WARP to arbitrary "chains". Consider any chain of alternatives $x_1, \ldots x_n$ such that if for all $i \in \{2, \ldots n\}$, x_i is revealed to be shortlisted in some menu where x_{i-1} is chosen, then the condition requires that whenever x_1 is revealed to be shortlisted in a menu, then x_n cannot be chosen in that menu. Formally,

Reject-SARP(R-SARP): For all $A_1, \ldots, A_n \in \mathcal{P}(X)$ and distinct $x_1, \ldots, x_n \in X$, if $x_{i+1} \in U(A_i, \bar{\triangleright}_c)$, $C(A_i) = x_i$ for $i = 1, \ldots, n-1$, then

$$[x_1 \in U(A_n, \bar{\triangleright}_c)] \implies [C(A_n) \neq x_n]$$

It turns out that a characterization of *Transitive-CBR* requires no more than this generalization of R-WARP to any arbitrary chain of alternatives. The characterization is then given by the following result, a proof of which is relegated to the Appendix.

Theorem 3. *A choice function C is a Transitive-CBR if and only it satisfies NC, WCC, R-Acyclicity and R-SARP.*

2.4 Identification

When given with a choice data (choices from all menus) that satisfies the above discussed conditions that characterize CBR, the next step is to examine to what extent the underlying rationales can be recovered. This enables one to conduct any meaningful welfare analysis. However, if the first stage shortlisting is unobservable, it is often difficult to uniquely pin down the rationales generating the choice function. This is the case with many popular two stage models such as the RSM of Manzini and Mariotti (2007) and CLA of Masatlioglu et al. (2012). In the case of CBR as well, there can be multiple representations of behavior. Consider the following example of choice function C given in the figure below where the arrows indicate the pairwise choices.

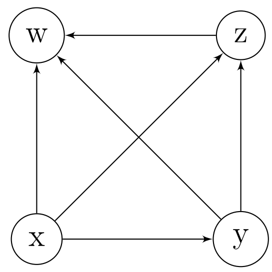


Figure 3: $C(xyz) = C(yzw) = C(xyzw) = y$ and $C(xyw) = C(xzw) = x$

Consider the following pair of rationales: $R_1 = \{(x,y),(x,z),(x,w),(y,z),(\mathbf{z},\mathbf{w})\}$ and $P_1 = \{(\mathbf{x},\mathbf{z}),(x,w),(y,x),(\mathbf{y},\mathbf{z}),(y,w),(\mathbf{w},\mathbf{z})\}$. This is a CBR-representation of the above described choice function (note that R_1 is transitive and P_1 is complete). Consider another pair of rationales: $R_2 = \{(x,y),(x,z),(x,w),(y,z),\}$ and $P_2 = \{(\mathbf{z},\mathbf{x}),(x,w),(y,x),(\mathbf{z},\mathbf{y}),(y,w),(\mathbf{z},\mathbf{w})\}$. This is also a CBR-representation of the above described choice function. Notice that in the second representation, z *never* gets shortlisted in the presence of x or y. Therefore, the choice function is not affected if we have (z,x) (z,y) in the representation or (x,z) (y,z) or any combination of these. However, in the first representation, z gets shortlisted in the presence of w, Since x is chosen in $\{x,z,w\}$ and y is chosen in $\{y,z,w\}$, we must have xP_1z and yP_1z. Another important thing to note is that since w is never chosen in the presence of z *and* w is not involved in any reversal, this gives us freedom to put (z, w) in any of the two rationales. Same is the case with (y, w) and (x, w).

In case of multiplicity of representations, as highlighted in Dutta and Horan (2015), a question of interest is whether we can find what, if any, are the common parts to *every* representation? Suppose for a given choice function C that is CBR representable, let $\mathcal{R}(C) := \{(R_i, P_i)\}_{i=1}^n$ be the collection of all its representations. we define the common parts for every representation denoted by $(\hat{R}, \hat{P})$ as follows:

(i) $\hat{R} := \{(x, y) : xR_iy \ \forall \ (R_i, P_i) \in \mathcal{R}(C)\}$

(i) $\hat{P} := \{(x, y) : xP_iy \ \forall \ (R_i, P_i) \in \mathcal{R}(C)\}$

We now show that $\trianglerighteq_c$ captures all the revealed first stage information that is true for every representation. For the second stage, there complications as one can see from the example above. Since (z, w) is a "free" pair, it can be assigned to any rationale —first or second. However, if we assign to the first rationale, then we are constrained to have (x, z) and (y, z) in the second rationale to ensure that z is not chosen in (xzw) and (xyz) respectively. Therefore, we can only specify the common parts of the second rationale for every rationale to certain pairs. These are the alternatives pairs, the information of which is revealed by reversals. We summarize this reasoning in the following result.

Theorem 4. *For any CBR representable choice function C, (i) $x\hat{R}y$ if and only if $x \trianglerighteq_c y$ (ii) $x\hat{P}y$ if there exists $A \in \mathcal{P}(X)$ such that $y \trianglerighteq_c w$ for some $w \in A$ and $C(A) = x$.*

Proof. To show that if part of (i), suppose $x \trianglerighteq_c y$ for some x, y. Now, consider any arbitrary representation (R, P) of C. That is $(R, P) \in \mathcal{R}(C)$. We will use the following lemma to establish the result [9]

Lemma **1.** *If there is a weak (xy) reversal due to z, then xRy, yPx and yRz and if there is a strong (xy) reversal due to z , then $\neg xRy$, xPy, zRx.*

Proof. Suppose there is a weak (xy) reversal due to some z. Since $x \succ_c y$, we have $\neg yRx$ and either xRy or xPy. Suppose xPy holds. Since $C(A) = x$ and $C(A \cup \{z\}) = y$, it must be that $x \in \min(A \cup \{z\}, R)$ and $x \notin \min(A, R)$. Therefore, we must have zRx, contradicting $x \succ_c z$. Thus xRy holds and $x \notin \min(A \cup \{z\}, R)$ implying yPx. For $C(A) = x$, it must be that $y \in \min(A, R)$ and for $C(A \cup \{z\}) = y$, yRz must be true. Suppose there is a strong (xy) reversal due to z. If xRy holds, then by the argument

[9]This Lemma appears as Claim 1 in the proof of Theorem 1.

above, yPx and yRz holds. By transitivity of R, xRz holds which contradicts $z \succ_c x$. Therefore xPy holds and x and y are R-unrelated. For $C(A \cup \{z\}) = y$, it must be that $x \in \min(A \cup \{z\}, R)$ and therefore for $C(A) = x$, it must be that x is R-unrelated to A. By an analogous argument in the case above, zRx and yPz hold. □

Since $x \bar{\triangleright}_c y$ implies $x \succ_R x_1 \succ_R \ldots \succ_R x_n \succ_R y$ and by the above lemma we have $xRx_1R \ldots Rx_nRy$ and by the transitivity of R, we have xRy. The "only if" part follows from the proof of Theorem 1, where we construct a representation with R^c as a first rationale.

To prove part (ii), let P^c be defined as xP^cy if there exists $A \in \mathcal{P}(X)$ such that $y \bar{\triangleright}_c w$ for some $w \in A$ and and $C(A) = x$. Then, by the argument above we know yRw and since $C(A) = x$, we must have xPy.

□

2.4.1 *A Small Menu Property*

Choice functions that are CBR rationalizable satisfy what we term as a "small menu" property. It says that if two choice functions from the class of CBR rationalizable choice functions coincide on menus of size 2 and 3, then they must coincide everywhere. Such a small menu property is also satisfied by the rational choice, TSM and RSM models. To establish this property, we will first show that strong and weak reversals on arbitrary menus are reflected in small menus as well. In particular,a weak reversal on any set implies a weak reversal on a pair of alternatives. Further, a strong reversal on any set implies either a strong reversal on a pair or triple of alternatives. The proof of this result is relegated to the Appendix.

Theorem 5. *If C and $\bar{C}$ are CBR-representable, then $C(\cdot) = \bar{C}(\cdot)$ if and only if $C(A) = \bar{C}(A)$ for all $A \subseteq X$ such that $|A| \leq 3$.*

2.5 A Discussion

2.5.1 *Weak WARP and Exclusivity*

Rational choice theory does not allow for reversals i.e. the choice of an alternative x when y is available in a menu and the choice of y when x is available in a different menu. The weak axiom of revealed preference (WARP) captures precisely this requirement. However, there is strong empirical evidence of choices displaying such reversals. Two prominent behavioral explanations of such reversals have been the *compromise effect* and the *attraction effect* which is also popularly known as the *decoy effect*. The compromise effect first discussed in Simonson (1989) says that individuals avoid "extreme" alternatives and "compromise" for non-extreme alternatives. The idea is that addition of an alternative to a menu makes the previously chosen alternative appear "extreme". Hence the choice shifts to an alternative which was not previously chosen, causing a reversal. The attraction effect —first discussed in Huber et al. (1982) —on the other hand says that the addition of an alternative to a menu acts as a "decoy" for an alternative that was previously not chosen, hence causing a reversal. For alternatives x, y and z, both the effects would be reflected behaviorally as

$$C(x, y) = x \quad \text{and} \quad C(xyz) = y$$

with z acting as alternative that makes x appear "extreme" in the compromise effect and z acting as a "decoy" for y in the decoy effect. We extend the idea above to what we term as a *single reversal* which is formally defined as

DEFINITION 3. *A choice function C displays a single (xy) reversal if $x \succ_c y$ and*

(i) There exists $\{x, y\} \subset A$ such that $C(A) = y$; and

(ii) For all $\{x, y\} \subset A \subset A'$, such that $C(A) = y$, we have $C(A') \neq x$.

The above definition permits for at most one reversal with respect to a pair (xy) in terms of set inclusion. It is easy to see that if a choice function satisfies WARP, then for a pair of alternatives (xy) , $x \succ_c y$ would imply that y can never be chosen from any menu that contains x. Expressed in terms of reversals, WARP allows for no reversal in choices between x and y along any sequence of sets (containing x and y) ordered by set inclusion. Whereas the Weak WARP condition of Manzini and Mariotti (2007) allows for single reversals in choices.

A natural implication of the compromise effect and the decoy effect is what Tserenjigmid (2019) calls the *two-compromise effect* and the *two-decoy effect*. In the case of the two-compromise effect, the argument is that an addition of the fourth alternative w to a menu would make x no longer appear an "extreme" alternative and the choice would revert to x. In case of the two-decoy effect, w would act as a "decoy" for x, nullifying the decoy effect of z for y. Again, both the effects would be reflected behaviorally as

$$C(xy) = x \quad \text{and} \quad C(xyz) = y \quad \text{and} \quad C(xyzw) = x$$

In a similar manner as a *single reversal*, we extend the above idea to what we term as a *double reversal* that is defined as follows.

DEFINITION 4. *A choice function C displays a double (xy) reversal if $x \succ_c y$ and*

(i) There exists $\{x, y\} \subset A \subset A'$ such that $C(A) = y$, $C(A') = x$; and

(ii) For all $\{x, y\} \subset A \subset A' \subset A''$, such that $C(A) = y$ and $C(A') = x$, we have $C(A'') \neq x$.

There is experimental evidence of double reversals (see Tserenjigmid (2019), Manzini and Mariotti (2010) , Teppan and Felfernig (2009)). We can see in Example 3 that CBR allows for a double reversal and this is what differentiates CBR from many shortlisting models in the literature.

An implication of R-WARP and WCC is that for any pair (xy), there can be no more than two reversals. So for a (xy) reversal from S to S', we can identify a menu T and alternative z, such that $S \subseteq T \subset S'$, $C(T) = x$ and $C(T \cup \{z\}) = y$, and choice is x for all sets in a "path" between S and T, and choice is y in a "path" between $T \cup \{z\}$ and S'. Similarly, for a *double reversal*, we can identify two menus where addition of an alternative leads to a reversal in the "path". Thus, an (xy) double reversal in the choice is associated with two alternatives z_1 and z_2 due to which the reversal takes place. The above axioms imply a weaker version of WWARP which we call R-WARP*. This condition restricts the number of reversals in any pair to at most two.

DEFINITION 5. *A choice function C satisfies R-WARP* if for all menus A, A', A'' such that $\{x, y\} \subset A' \subset A \subset A''$*

$$[C\{x, y\} = x,\ C(A') = y, C(A) = x] \implies [C(A'') \neq y]$$

The above discussed restriction can be summarized by the following result, a proof of which is relegated to the Appendix.

PROPOSITION **1.** *If C satisfies R-WARP and WCC, then it satisfies R-WARP*. Hence, CBR satisfied R-WARP**

Another implication of the axioms above is a condition which imposes clear limitations on the possibility of certain simultaneous weak and strong reversals. For a given weak reversal it precludes certain strong reversals and vice-versa. This is captured in a property which we call *Exclusivity*.[10] It allows for only one type of reversal between a pair due to any alternative. Formally, it is defined as follows:

[10] This is closely related to the Exclusivity condition of Horan (2016)

Definition 6. *Exclusivity: For any pair of alternatives* (xy)*, either:*

- *C displays no* ***weak*** (xy) *reversal; or*
- *C displays no* ***strong*** (xy) *reversal*

For any pair of alternatives, this condition precludes choice behavior which exhibits both types of reversals, strong and weak. Put differently, the possibility of strong reversals for a given pair of alternatives is ruled out by observing a single weak reversal for that pair (and vice versa). As we show in Appendix, the following result is an implication of Lemma 4.

Proposition **2.** *If C is CBR-representable, then it satisfies Exclusivity*

5.2 *RSM and TSM*

Manzini and Mariotti (2007) showed that the violation of rationality (WARP) is attributed to violation of either of the following two consistency conditions: *Always Chosen*[11] or *No Binary Cycles*[12]. Various boundedly rational models explain violation of rationality using violation of either of these conditions. Manzini and Mariotti (2007) show that RSM is able to accomodate the violation of *No Binary Cycles.* However, a violation of *Always Chosen* cannot be explained by RSM. CBR, however, is able to explain both the violations.

We now compare two related models with CBR and show that what extra conditions do we require to pin down those subclasses within the class of CBR choice functions.

Rational Shortlist method: RSM is not a special case of our model since it is characterized it by two axioms that may be violated by CBR: Expansion (EXP) [13] and WWARP. However, as shown earlier, CBR satisfies a weaker version of this axiom (R-WARP*) which allows for at most two reversals. Also, CBR may violate EXP as a weak (xy) reversal due to z implies $C(xyz) = y$, $C(xy) = x = C(xz)$, $C(yz) = y$ which violates *Always Chosen.* The reversals discussed in this paper establish a relation between our model and RSM. This is expressed in the following result, a proof of which can be found in the Appendix .

Proposition **3.** *If Choice function C is CBR, then C is RSM if and only if C displays no weak reversals*

[11]If x is chosen in pairs, then it must been chosen union of those pairs
[12]Relation derived from pairwise choices cannot have a cycle
[13]For all $S, S' \supset \{x, y\}$, $C(S) = C(S') = x$ implies $C(S \cup S') = x$

Transitive Shortlist method: The TSM is a variant of the RSM where both the rationales are transitive (but possibly incomplete). Horan (2016) analyzes this choice procedure in terms of two choice reversals: direct and weak* [14] reversal. A direct $\langle x, y\rangle$ reversal on $B \subset X \setminus \{x\}$ is defined as

$$C(B) = y \text{ and } C(B \cup \{x\}) = z \notin \{x, y\}$$

A weak* $\langle x, y\rangle$ reversal on $B \supset \{x, y\}$ is defined as

$$C(xy) = x \text{ and } C(B \setminus \{y\}) \neq C(B)$$

TSM satisfies an Exclusivity condition which says that for a pair x, y, either there is no direct $\langle x, y\rangle$ reversal on $B \subset X \setminus \{x\}$ or, there is no weak* $\langle x, y\rangle$ reversal. CBR violates this axiom when there is a double reversal. It can be seen in Example 3. There is a direct $\langle z, x\rangle$ reversal on $\{x, z\}$ and a weak* $\langle z, x\rangle$ reversal on $\{y, z, w\}$. Another property satisfied by TSM is Expansion (and by implication *always chosen*), which CBR need not satisfy. Thus, TSM is also not a special case of CBR.

Note that since TSM satisfies WWARP, in the case of a direct $\langle x, y\rangle$ reversal it must be be that $C(yz) = y$. Hence, whenever there is a strong or a weak reversal, we have a direct reversal. Conversely, as TSM also satisfies *Always Chosen*, a choice function that is TSM-representable cannot display a weak reversal. Therefore, this would be a strong reversal. As in the case of RSM, our model relates to TSM in the following way. A proof of the following result is analagous to the proof the previous proposition,

Proposition **4.** *If Choice function C is T-CBR, then C is TSM if and only if C displays no weak reversals*

2.6 Stochastic Choice

In this section, we consider a DM who chooses stochastically. A stochastic choice function is any map $p : X \times \mathcal{P}(X) \to [0, 1]$ such that for all $A \in \mathcal{P}(X)$, it satisfies the following two conditions: (i) $\sum_{a \in A} p(a, A) = 1$ and (ii) $p(a, A) = 0$ for all $a \notin A$. We study a DM who uses assigns zero probabilities to the worst alternatives in every menu. We follow the technique introduced by Echenique and Saito (2019) who provide a general Luce model that allows for zero probabilities. A special case of the general Luce model is

[14]Weak reversal defined in Horan (2016). $\star$ added to avoid confusion with weak reversal of this paper

their two-stage Luce model where only the maximal alternatives of an underlying binary relation survive the first stage shortlisting i.e. get positive probabilities.

We propose another special case of the general Luce model where any non-minimal alternative also positive probability. Formally, we define it as follows.

Definition 7. *A stochastic choice function p is a Reject-Luce (R-Luce) function if there exist a $u : X \to \mathbb{R}_{++}$ and a rejection filter*[15] Γ *such that for all* $A \in \mathcal{P}(X)$

$$p(x, A) = \begin{cases} \frac{u(x)}{\sum_{y \in \Gamma(A)} u(y)} & \text{if } x \in \Gamma(A) \\ 0 & \text{if } x \notin \Gamma(A) \end{cases}$$

We show that by adapting the characterizing of the axioms the rejection filter to a stochastic setting and using a generalization of the Luce-IIA condition given by Echenique and Saito (2019), we get a characterization of the R-Luce model. The axiom of Weak contraction is adapted as follows.

S-Weak Contraction: For any $A \in \mathcal{P}(X)$ and $x, y \in A$, such that $p(x, xy) < 1$,

$$[p(x, A) > 0] \implies [p(x, A \setminus \{y\}) > 0]$$

Similarly the axiom of Strict Expansion is adapted as follows.

S-Strict Expansion : For all $A, A' \in \mathcal{P}(X)$ and for all $x \in A \cap A'$,

$$[p(x, A) > 0, p(x, A') > 0] \textbf{ or } [p(x, A) = 1] \implies [p(x, A \cup A') > 0]$$

Since we do not assume transitivity of the underlying rationale, we do not require Binary Rejection Consistency. Theorem 1 of Echenique and Saito (2019) characterizes the general Luce model using the above generalization of Luce IIA which they term the cyclical independence condition and is stated as follows.

Cyclical independence: For any sequence $x_1, x_2, \ldots, x_n \in X$, if there exists a sequence $S_1 \ldots, S_n \in \mathcal{X}$ s.t. $p(x_i, S_i) > 0$ and $p(x_{i+1}, S_i) > 0$, for $i = 1, \ldots, n-1$, and

[15] Here, we do not assume the transitivity of the underlying rationale generating the rejection filter. However, if we require transitivity of the underlying rationale, with an additional axiom that is the stochastic analogue of Binary Rejection Consistency, we can be obtain a characterization of the R-Luce choice function with a transitive rationale.

$p(x_n, S_n) > 0$ and $p(x_1, S_n) > 0$, then

$$\frac{p(x_1, S_n)}{p(x_n, S_n)} = \frac{p(x_1, S_1)}{p(x_2, S_1)} \cdot \frac{p(x_2, S_2)}{p(x_3, S_2)} \cdot \ldots \cdot \frac{p(x_{n-1}, S_{n-1})}{p(x_n, S_{n-1})}$$

Since R-Luce model is a special case of the general Luce model, we require cyclical independence condition along with the above two conditions. We conclude this section by providing a characterization of the R-Luce model using these three conditions.

THEOREM 6. *A stochastic choice function is a R-Luce function if and only if it satisfies S-Weak Contraction, S-Weak Expansion and Cyclical Independence*

Proof. To show the sufficiency part, define R as xRy if and only if $p(x,(xy)) = 1$ and $\Gamma_R : \mathcal{P}(X) \to \mathcal{P}(X)$ as $\Gamma_R(A) := \{x \in A | xRy \text{ for some } y \in A \text{ or } \neg yRx\ \forall y \in A\}$. Note that R is asymmetric by definition. Let $c(A) =$supp A. Now, we first show that $\Gamma_R = c$.

Consider an arbitrary $A \in \mathcal{P}(X)$. Suppose $x \in \Gamma_R(A)$. Then there are two possible cases: (i) $p(x, xy) = 1$ and by S-Strict Expansion, we get $p(x, A) > 0$ implying $x \in c(A)$. (ii) $p(x, xy) \in (0, 1)$ for all $y \in A \setminus \{x\}$. Again by S-Strict Expansion, we get $p(x, A) > 0$ implying $x \in c(A)$. Therefore, $\Gamma_R(A) \subset c(A)$. To show $c(A) \subset \Gamma_R(A)$, we use strong induction on the size of the menus. For the base case —menus of size 2, by the definition of Γ_R, we have $\Gamma_R(A) = c(A)$. Now, for the inductive step, consider any $k \geq 2$ and assume that $\Gamma_R(A) = c(A)$ for all menus A with $|A| \leq k$. Consider any menu A' such that $|A'| = k + 1$ and $x \in A'$ such that $x \in c(A')$ i.e. $p(x, A') > 0$. Assume for contradiction that $x \notin \Gamma_R(A')$. Then there exists $y \in A'$ such that $p(y, xy) = 1$. Suppose for some $z \neq y$, $z \in A'$, we have $p(x, xz) \in [0, 1)$. Then by S-Weak Contraction, we have $p(x, A' \setminus \{z\}) > 0$. Since $y \in A' \setminus \{z\}$ and $p(y, xy) = 1$, by our induction hypothesis, we must have $p(x, xw) = 1$ for some $w \in A' \setminus \{z\}$ implying $x \in \Gamma_R(A')$, a contradiction. Therefore, for all $z \neq y$, $z \in A'$, we have $p(x, xz) = 1$ and again, we have $x \in \Gamma_R(A')$, a contradiction. So, $x \in \Gamma_R$ and we have shown that $c(A') \subset \Gamma_R(A')$.

Now, using Cyclical Independence, we can use the arguments in the proof of Theorem 1 of Echenique and Saito (2019) to establish that p is a R-Luce function. □

2.7 Conclusion

In this paper, we introduced a new model of decision-making that formalized the idea of rejection behavior using binary relations. We introduced a procedure of rejection that generates a consideration set mapping which we term a rejection filter. We provided an axiomatic foundation in line with axiomatic definitions of other popular consideration set mappings in the literature: the attention filter and the competition filter. Using the rejection filter, we introduced a two-stage procedure that combines rejection and maximization to produce a unique choice that we term Choice by Rejection (CBR). We used observable reversals in choice to conduct a revealed preference analysis of this procedure and provided a behavioral characterization. Further, using reversals in choice we provided results on partial identification of the underlying rationales. We compared our procedure with another prominent two-stage model, the Rational Shortlist Method (RSM) and showed that within the class of CBR choice functions, RSM choice functions are precisely those that do not display a certain choice reversal. Finally, as another application of the rejection filter, we introduced and characterized a two-stage stochastic choice procedure.

2.8 Appendix

2.8.1 Proof of Theorem 2

To show the necessity of the conditions, Let C be a CBR and R (transitive) and P (complete) be the first and second stage rationales. First we prove the following claim:

Claim 1. *If there is a weak (xy) reversal due to z, then xRy, yPx and yRz and if there is a strong (xy) reversal due to z , then $\neg xRy$, xPy, zRx.*

Proof. Suppose there is a weak (xy) reversal due to some z. Since $x \succ_c y$, we have $\neg yRx$ and either xRy or xPy. Suppose xPy holds. Since $C(A) = x$ and $C(A \cup \{z\}) = y$, it must be that $x \in \min(A \cup \{z\}, R)$ and $x \notin \min(A, R)$. Therefore, we must have zRx, contradicting $x \succ_c z$. Thus xRy holds and $x \notin \min(A \cup \{z\}, R)$ implying yPx. For $C(A) = x$, it must be that $y \in \min(A, R)$ and for $C(A \cup \{z\}) = y$, yRz must be true. Suppose there is a strong (xy) reversal due to z. If xRy holds, then by the argument above, yPx and yRz holds. By transitivity of R, xRz holds which contradicts $z \succ_c x$. Therefore xPy holds and x and y are R-unrelated. For $C(A \cup \{z\}) = y$, it must be that $x \in \min(A \cup \{z\}, R)$ and therefore for $C(A) = x$, it must be that x is R-unrelated to A. By an analogous argument in the case above, zRx and yPz hold. □

From the above claim it immediately follows that if C is RAC, then $x \rhd_c y$ implies xRy. Now, we show the necessity of each of the axioms.

NC: For any A with $C(A) = x$, it must be that $x \notin \min(S, R)$. Therefore, either xRz holds for some $z \in S$ or x is R-unrelated to A . If xRz holds, then we know that $C(xz) = x$. If x is R-unrelated to A, then we must have at least one $z \in A$ such that $z \notin \min(A, R)$. Therefore we must have xPz and we get $C(xz) = x$.

WCC: Let $A = \{x, y, x_1, x_2, ..., x_n\}$ and $C(xy) = x$ and $C(A) \in \{x, y\}$. Denote by A_i the subset of S which is $A_i := A \setminus \{x_i\}$. Assume for contradiction that $C(A_i) \notin \{x, y\}$ for all $i \in \{1, 2, ..., n\}$. Hence, $C(A_i) = x_j$ for some $i \neq j$. We denote the choice in A_i by c_i. Consider the first case where $C(A) = x$. If xRy then $x \notin \min(A_i, R)$ for all i. For c_i to be chosen in A_i, c_iPx must hold for all i. Note that for $C(A) = x$, it must be that $c_i \in \min(A, R)$ for all i, which is possible when c_i is R-unrelated to A_i and x_iRc_i for all i. But, for every i, there exists a $j \neq i$ such that $c_i = x_j$, implying that there exists at least one $c_i \notin \min(A, R)$ which is a contradiction.

Now, let $\neg xRy$ and thus xPy hold. As $x \notin \min(A, R)$, and $A_i \cup A_j = A$ for any $i \neq j$, we have $x \in \min(A_i, R)$ in at most one A_i. [16] If $x \notin \min(A_i, R)$ for all i, then argument

[16] For any $A, B \in \mathcal{P}(X)$ and R, $x \in \min(A, R) \cap \min(B, R)$ implies $x \in \min(A \cup B, R)$.

becomes similar to the case above where xRy holds. Assume $x \in \min(A_n, R)$ ($i = n$ W.L.O.G). For c_i to be chosen in A_i ($i \neq n$), $c_i P x$ holds and for x to be chosen in S, $c_i \in \min(S, R)$ for all $i \neq n$, for which c_i R-unrelated to S_i and $x_i R c_i$ for all i. Therefore we must have $c_i = x_n$ for all $i \neq n$. Since, $x \in \min(S_n, R)$ and $x \notin \min(A, R)$, we have xRx_n, a contradiction since we require c_i to be R-unrelated to A_i.

Let us now consider the second case where $C(A) = y$. Now we have (xy) reversal. If xRy holds, then we have yPx as $x \notin \min(A, R)$ and there exists a $x_k \in A$ such that yRx_k holds. As $y \notin \min(A_i, R)$ for all $i \neq k$, choice of c_i in A_i requires $c_i P y$. Further, as y is chosen in A, $c_i \in \min(A, R)$ for all $i \neq k$. By the arguments in the previous case, we require $x_i R c_i$, c_i R-unrelated to A_i and $c_i = x_k$ for all $i \neq k$. Since yRx_k holds, this is a contradiction since we require c_i to be R-unrelated to A_i. Suppose $\neg xRy$ and xPy hold. The choice of y in A requires $x \in \min(A, R)$ i.e. there exists a x_k (say x_n) such that $x_n R x$ holds and for no alternative z, xRz is true. It must be that $y \in \min(A_i, R)$ for at most one A_i. If $y \notin \min(A_i, R)$ for all i then we get $c_i P y$ for all i. Therefore we must have c_i R-isolated in A_i and $x_i R c_i$ and by the argument in the previous case, we have a contradiction. Suppose $y \in \min(A_k, R)$ for some k, using arguments mentioned above, we have $c_i P y$, $x_i R c_i$ and c_i is R-unrelated to A_i for all $i \neq k$ which restricts $c_i = x_k$ for all $i \neq k$. Since, $y \in \min(A_k, R)$ and $y \notin \min(A, R)$, we have yRx_k, a contradiction since we require c_i to be R-unrelated to A_i.

R-acyclicity: By claim 1, we know that $x \ \triangleright_c \ y$ implies xRy. Since R is transitive, and xRy implies $x \succ_c y$, R-acyclicity follows.

R-WARP: Consider $\{x, y\} \subseteq A, A' \in \mathcal{P}(X)$ such that $y \notin \min(A, \triangleright_c)$, $C(A) = x$, and $x \notin \min(A', \triangleright_c)$. Consider the case when $y \triangleright_c z$ for some $z \in A$. By claim 1 we have yRz and $y \notin \min(A, R)$. As $C(A) = x$, we must have xPy. Now, if $x \triangleright_c w$ for some $w \in A'$, then $x \notin \min(A', R)$. This implies $C(A') \neq y$. Now suppose, $\neg x \triangleright_c w$ for all $w \in A'$. For $C(A') = y$, we need $x \in \min(A', R)$. Suppose that $C(xy) = x$. Since C satisfies WCC, there exists a $z \in A'$, such that there is a strong (xy) reversal due to z as weak reversal implies xRy. By definition $z \triangleright_c x$ holds, which is a contradiction as for $x \notin \min(A', \triangleright_c)$, we need $x \triangleright_c w$ for some $w \in A'$. If $C(xy) = y$, then by similar argument, there exists a $w \in A$ such that there is a strong/weak (yx) reversal due to some $w \in A$. If the reversal is weak, then $y \triangleright_c x$ holds. For $x \notin \min(A', \triangleright_c)$, there is a $w' \in A'$ such that $x \triangleright_c w'$ holds, a contradiction. If the reversal is strong, then yPx holds, again a contradiction.

Now consider the case when y is $\triangleright_c$-unrelated to A. If $C(xy) = y$, we have a (yx) reversal due to some $z \in A$. If it is a weak reversal, then $y \triangleright_c x$ and if it is a strong reversal, then $z \triangleright_c y$, a contradiction. Therefore we have $C(xy) = x$. Now, suppose $C(A') = y$, then we have a (xy) reversal due to some $z \in A'$. If it is a weak reversal then we get $x \triangleright_c y$, a contradiction. Therefore, it must be a strong (xy) reversal implying xPy. Since $z \triangleright_c x$ and we have $x \notin \min(A', \overline{\triangleright}_c)$, there exists $w \in A'$ such that $x \overline{\triangleright}_c w$ implying xRw. Therefore $x \notin \min(A', R)$ and we cannot have $C(A') = y$.

Now, we show the sufficiency of the axioms. Before, constructing the revealed rationales, we prove some useful lemmas. Suppose C satisfies NC, WCC , R-Acyclicity and R-WARP.

Lemma **2.** *A strong (xy) reversal implies $\neg x \overline{\triangleright}_c y$.*

Proof. Given there is a strong (xy) reversal on A due to some z i.e. $C(A) = x$ and $C(A \cup \{z\}) = y$. By the definition of $\triangleright_c$, we have $z \triangleright_c x$. Suppose $x \overline{\triangleright}_c y$ holds. This implies $z \overline{\triangleright}_c y$ by transitivity of $\overline{\triangleright}_c$. Suppose $y \notin \min(A, \overline{\triangleright}_c)$. Then since $C(A) = x$ and $x \overline{\triangleright}_c y$ implying $x \notin \min(A \cup \{z\}, \overline{\triangleright}_c)$, by R-WARP we must have $C(A \cup \{z\}) \neq y$. Therefore, we have $y \in \min(A, \overline{\triangleright}_c)$. Since $z \triangleright_c y$, by R-Acyclicity, we have $C(yz) = z$ and there is a (zy) reversal from $\{y, z\}$ to $A \cup \{z\}$. By WCC, there exists a $x_1 \in A$ such that the reversal is due to x_1. If it is a weak (zy) reversal due to x_1, then by definition of $\triangleright_c$, we have $y \triangleright_c x_1$, a contradiction. Therefore, it must be that we have strong (zy) reversal due to x_1. By definition of $\triangleright_c$, we have $x_1 \triangleright_c z$ implying $x_1 \neq x$ by R-Acyclicity. This further implies that $x_1 \overline{\triangleright}_c y$ and by R-acyclicity, we have $C(x_1 y) = y$.

Now, we have a $(x_1 y)$ reversal due to some $x_2 \in A$. By similar argument as above, this must be a strong reversal, implying $x_2 \triangleright_c x_1$. Also, by transitivity of $\overline{\triangleright}_c$, we have $x_2 \overline{\triangleright}_c z$ and $x_2 \overline{\triangleright}_c x$ and therefore $x_2 \neq z, x$. Further, by R-Acyclicity we have $x_2 \overline{\triangleright}_c y$. This leads to $(x_2 y)$ reversal due to some $x_3 \in A$. Since A is finite, proceeding inductively, this leads to $x_i \overline{\triangleright}_c y$ for all $x_i \in (A \cup \{z\}) \setminus \{x\}$ as in each step, $x_{i+1} \neq x_k$, $k \leq i$ by R-Ayclicity. Since $x_i \succ_c y$ for all $x_i \in A \cup \{z\}$, NC implies that $C(A \cup \{z\}) \neq y$, a contradiction. Therefore, our initial supposition that $x \overline{\triangleright}_c y$ holds is incorrect and we have $\neg x \overline{\triangleright}_c y$. ☐

Lemma **3.** *A strong (xy) reversal on set A implies x is $\overline{\triangleright}_c$-unrelated to S.*

Proof. Suppose a strong (xy) reversal is observed on set A due to some z. That is, $C(A) = C(xy) = x$ and $C(A \cup \{z\}) = y$. By Lemma 2, we have $\neg x \overline{\triangleright}_c \ y$ and by R-acylicity, we have $\neg y \overline{\triangleright}_c \ x$. Hence, $y \notin \min(\{x, y\}, \overline{\triangleright}_c)$. Now, since $y = C(A \cup \{z\})$, we have $x \in \min(A \cup \{z\}, \overline{\triangleright}_c)$ by R-WARP. This implies $\neg x \overline{\triangleright}_c w$ for all $w \in S$. Suppose $x \in \min(A, \overline{\triangleright}_c)$. Then $w \overline{\triangleright}_c x$ for some $w \in S$. Then by R-Acyclicity, we have $C(xw) = w$

and there is a (wx) reversal. By WCC and Lemma 2 this is a weak (wx) reversal due to some $w' \in A$. By definition, $x \rhd_c w'$, contradicting $x \in \min(A \cup \{z\}, \bar{\rhd}_c)$. Therefore, $x \notin \min(A, \bar{\rhd}_c)$ and is $\bar{\rhd}_c$-unrelated to A. $\square$

Lemma 4. *If there is a weak (xy) reversal on some set A, then there does not exist $y' \in X$ and a menu A' with $y \in A'$ such that there is a strong (xy') reversal on; or there is a strong (yy') reversal on A' with $x \in A'$.*

Proof. Let there be a weak (xy) reversal due to some z. By definition of $\rhd_c$, we have $x \rhd_c y$. Suppose there is strong (xy') reversal on A' with $y \in A'$. Then by Lemma 3, we must have x $\bar{\rhd}_c$-unrelated to A', a contradiction. Suppose, there is a strong (yy') reversal on A' with $x \in A'$, by Lemma 3, we must have x $\bar{\rhd}_c$-unrelated to A', a contradicton. $\square$

Now, we will define a revealed rationales pair (R^c, P^c) with R^c transitive and P^c complete that rationalizes the choice. Define R^c as xR^cy if and only if $x\bar{\rhd}_c y$. We will define P^c to be a union of two binary relations P_1 and P_2. We define P_1 as xP_1y if and only if there exists a A such that $C(A) = x$ and $y \notin \min(A, \bar{\rhd}_c)$ and P_2 is defined as $P_2 \coloneqq R^c \setminus P_1 \cup P_1^{-1}$, where P_1^{-1} is defined as $xP_1^{-1}y$ if and only if yP_1x. Therefore, $P^c \coloneqq P_1 \cup P_2$.

Note that $\rhd_c$ is acyclic by R-Acyclicity. Therefore, since R^c is the transitive closure of $\rhd_c$, it is asymmetric and transitive.

Now, we will show that P^c is asymmetric. Assume for contradiction that it is not. That is, xP^cy and yP^cx for some $x, y \in X$. There are four possible cases: (i)xP_1y and yP_1x. This violates R-WARP. (ii)xP_2 and yP_2x. Since $P_2 \subset R^c$ and we have shown that R^c is asymmetric, this violates the asymmetry of R^c. (iii) xP_1y and yP_2x. Since xP_1y implies $yP_1^{-1}x$, by definition of P_2 we have $\neg yP_2x$. (iv) xP_2y and yP_1x. This case is the same as case (iii). So, we have established the asymmetry of P^c.

To show the completeness of P^c, consider any arbitrary $x, y \in X$. Assume W.L.O.G that $C(xy) = x$. Suppose x is R^c-unrelated to y. Then we have $y \notin \min(\{x, y\}, \bar{\rhd}_c)$ and by the definition of P_1, we have xP_1y. Clearly, we cannot have $y\bar{\rhd}_c x$ as that would imply $C(xy) = y$ by R-Acyclicity. Suppose $x\bar{\rhd}_c y$. Then if $C(A) = y$ for some A such that $\{x, y\} \subset A$, then we have yP_1x and we are done. If $C(A) = x$ for some A such that $y \notin \min(A, \bar{\rhd}_c)$, then we have xP_1y. If $C(A) \neq y$ for all A such that $\{x, y\} \subset A$, then we have $\neg yP_1x$ implying $\neg xP_1^{-1}y$. Since xR^cy and $P_2 \subset R^c$, we have xP_2y. So, we have xP^cy and since x and y were chosen arbitrarily, we have established the completeness of P^c.

Finally, we will show that the above defined R^c and P^c rationalize the choice function C. That is $C(A) = \max(\Gamma(A), P^c)$ where $\Gamma(A) = A \setminus (A, R^c)$. Consider an arbitrary set A and suppose $C(A) = x$. Now, suppose that $x \in \min(A, R^c)$. Then there exists a $y \in A \setminus \{x\}$ such that $yR^c x$ holds and $\neg x R^c y'$ for all $y' \in A \setminus \{x\}$. By R-Acyclicity, we have $C(xy) = y$. Note that by WCC, there exists a sequence of sets ordered by set inclusion from $\{x, y\}$ to S with choices belonging to $\{x, y\}$ and there exists a $z \in S$ such that there is a (yx) reversal due to z. Since $y \bar{\triangleright}_c x$, this reversal cannot be a strong reversal by Lemma 3 as it implies $\neg y \triangleright_c x$, violating R-Acyclicity .Therefore, it is a weak reversal. By Lemma 5, we have $y \succ_c x \succ_c z$ and $C(xyz) = x$. Thus, we must have $xR^c z$, leading to a contradiction.
Now we show that $x = \max(S \setminus \min(S, R^c), P^c)$. Consider any y such that $y \notin \min(S, R^c)$ and $yP^c x$ holds. We know that by construction of P^c, we have $xP_1 y$ ($\implies xP^c y$) which contradicts the asymmetry of P^c. Therefore $xP^c y$ for all $y \notin \min(S, R^c)$. Since P^c is asymmetric and complete, x is the unique maximal element.

2.8.2 Proof of Theorem 3

Consider a choice function C that is T-CBR representable and consider a representation (R, P) of it where R is transitive and P is complete and transitive. Since T-CBR is a subclass of CBR, the necessity of WCC, NC and R-Acyclicity follows from the proof of Theorem 1. Let us now prove the necessity of R-SARP. Suppose for some $S_1, \ldots, S_n \in \mathcal{P}(X)$ and distinct $x_1, \ldots, x_n \in X$, we have $x_{i+1} \notin \min(S_i, \bar{\triangleright}_c)$, $C(S_i) = x_i$ for $i = 1, \ldots, n-1$, and $x_1 \notin \min(S_n, \bar{\triangleright}_c)$. Using the argument in proving the necessity of R-WARP, we must have $x_i P x_{i+1}$ for all i. Since P is transitive, we must have $x_1 P x_n$. If $C(S_n) = x_n$, by a similar argument, it would imply $x_n P x_1$ a contradiction to the asymmetry of P.

To show the sufficiency of the axioms, note that since T-CBR is a subclass of CBR, Lemmas 2, 3 and 4 hold. So now, we will define a revealed rationale pair (R^c, P^c): Let $R^c \equiv \bar{\triangleright}_c$ and $P^c \equiv \bar{P}_1 \cup \hat{P}_2$ where xP_1y if and only if there exists a S such that $C(S) = x$ and $y \notin \min(S, \bar{\triangleright}_c)$ and $\bar{P}_1 = tc(P_1)$ i.e. $\bar{P}_1$ is the transitive closure of P_1. And, $\hat{P}_2 \equiv R^c \setminus (\bar{P}_1 \cup \bar{P}_1^{-1})$

R^c is asymmetric and transitive as discussed in the proof of Theorem 1. Next we show that P^c is a linear order. First, we show that P^c is asymmetric. If possible, say for some x, y, both xP^cy and yP^cx is true. Either $x\bar{P}_1y$ or $y\bar{P}_1x$ is true (else both will be derived through R^c which is a contradiction to the asymmetry of R^c). W.L.O.G suppose $x\bar{P}_1y$ holds. That is, there exists $A \in \mathcal{P}(X)$ such that $y \notin \min(A, \triangleright_c)$ and $C(A) = x$. Then $y\bar{P}_1x$ cannot hold by R-SARP. Now since yP^cx, it must be $y\hat{P}_2y$. Since $\hat{P}_2 = R^c \setminus (\bar{P}_1 \cup \bar{P}_1^{-1})$, we get that x is $\bar{P}_1$-unrelated to y, a contradiction. Therefore, P^c is asymmetric.

To show the completeness of P, consider any arbitrary $x, y \in X$ and let us assume that $C(xy) = x$. If x and y are not related with respect to R^c, this implies xP_1y since $y \notin \min(\{x, y\}, \triangleright_c)$. Now, if $(xy) \notin \bar{P}_1 \cup \bar{P}_1^{-1}$, then x and y are related with respect to R^c. Therefore, (xy) is related with respect to $\hat{P}_2$.

To show transitivity of P^c, note that since it is complete, we only need to show that it is acyclic. Assume for contradiction that P^c is cyclic. Then there exists $x_1 \ldots x_n$ such that $x_1P^c \ldots P^cx_nP^cx_1$. Since P^c is complete, this must imply a 3-cycle. Therefore, we only need rule out a 3-cycle. Suppose we have a xyz 3-cycle i.e. for some x, y and z, $xP^cyP^czP^cx$. Since $C(xyz) \neq \emptyset$, we must have a reversal. Assume W.L.O.G that there is a (xy) reversal. If it is a strong reversal then we have xP_1y. If it is a weak reversal, then have yP_1x. Therefore at least one of the pair must be related in $\bar{P}_1$. The following cases are possible: (i) $x\bar{P}_1y$ and $y\bar{P}_1z$: This would imply $x\bar{P}_1z$ by the transitivity of $\bar{P}_1$ and we get xP^cz. (ii) $x\bar{P}_1y$ and $y\hat{P}_2z$: Since zP^cx is true, following two cases are true:

- $z\bar{P}_1x$: This would imply $z\bar{P}_1y$ and $\neg y\bar{P}_1z$ therefore zP^cy, a contradiction to the asymmetry of P_c.
- $z\hat{P}_2x$: Since $\hat{P}_2 \subset R^c$ and by the transitivity of R^c, we have yR^cx. By NC and R-Acyclicity, $C(xyz) \in \{y, z\}$. If $C(xyz) = y$, then we have $y\bar{P}_1z$ and by the transitivity of $\bar{P}_1$, we have $x\bar{P}z$, a contradiction. If $C(xyz) = y$, then we have $y\bar{P}_1z$ and since $\bar{P}_1$ and $\hat{P}_2$ are disjoint by definition, we have a contradiction to $y\hat{P}_2z$.

Now, we show that the above defined (R^c, P^c) rationalize the C. Consider a set S and $C(S) = x$. Suppose $x \in \min(S, R^c)$. Then there exists $y \in S \setminus \{x\}$ such that yR^cx holds and $\neg xR^cy'$ for all $y' \in S \setminus \{x\}$. Note that by R-Acylicity, there exists a sequence of sets ordered in set inclusion from $\{x, y\}$ to S with choices belonging to $\{x, y\}$. R-SARP ensures that there exists a $z \in S$ that causes the (yx) reversal. As argued above, Lemma 2 and 3 imply that the reversal is weak. Therefore, we have $y \succ_c x \succ_c z$ and $C(xyz) = x$ and by definition, $x \bar{\rhd}_c z$, leading to a contradiction.

Finally, we show that $x = \max(S \setminus \min(S, R^c), P^c)$. Consider any such y that $y \notin \min(S, \bar{\rhd}_c)$ and yP^cx holds. We know that by construction of P^c, we have $xP_1y (\implies xP^cy)$ holds which contradicts the asymmetry of P^c. Therefore xP^cy for all $y \notin \min(S, \bar{\rhd}_c)$.

2.8.3 Proof of Theorem 5

The "only if" part is true by definition. To show the "if" part, we first prove the following two useful lemmas:

Lemma **5.** *If there is a weak (xy) reversal due to z, then $x \succ_c y \succ_c z$ and $C(xyz) = y$.*

Proof. Suppose for some x, y we have a weak (xy) reversal due to some z on set $A \neq \{x, y, z\}$. That is, $C(xy) = C(A) = x$ and $C(A \cup \{z\}) = y$. By definition of $\triangleright_c$, we have $x \triangleright_c y$, $y \triangleright_c z$ and hence $x \bar{\triangleright}_c z$. By R-Acyclicity, $x \succ_c y \succ_c z$ and $C(xyz) \neq z$ due to NC (since $x \succ_c z$). Note that $x \notin \min(A \cup \{z\}, \bar{\triangleright}_c)$ and $y = C(A \cup \{z\})$. Since $y \notin \min(\{x, y, z\}, \bar{\triangleright}_c)$, R-WARP implies that $C(xyz) \neq x$. Therefore we have $C(xyz) = y$. Now, suppose $A = \{x, y, z\}$. If $C(A) = x$, then by R-WARP, we cannot have a (xy) reversal on any set A' with $\{x, y, z\} \subseteq A'$, a contradiction. Therefore, we have $C(A) = y$. □

Remark: This lemma can also be proved from Lemma 1.

Lemma **6.** *If there is a strong (xy) reversal due to z, then either (i) $C(xyz) = y$ and $x \succ_c y \succ_c z \succ_c x$; or (ii) $C(xyz) = z$ and $z \succ_c x \succ_c y$ and for some w, $C(xyw) = x$ and $C(xyzw) = y$.*

Proof. Suppose for some x, y we have a strong (xy) reversal due to some z on some set A. That is, $C(xy) = C(A) = x$ and $C(A \cup \{z\}) = y$. By definition of $\triangleright_c$, we have $z \triangleright_c x$. We know that $\neg y \bar{\triangleright}_c z$ (as $y \bar{\triangleright}_c z$ would imply $y \bar{\triangleright}_c x$, violating R-Acyclicity). If $C(xyz) = x$, then we have a (zx) reversal due to y. It cannot be a strong reversal as it would imply $y \bar{\triangleright}_c z$. Therefore, it must be a weak (zx) reversal due to y which implies $x \bar{\triangleright}_c y$, a contradiction to Lemma 2. If $C(xyz) = y$, then by NC, we must have $C(yz) = y$ and we get $x \succ_c y \succ_c z \succ_c x$.

Now, suppose we have $C(xyz) = z$. By definition of $\triangleright_c$, we have $z \triangleright_c x$ implying $z \notin \min(S \cup \{z\}, \bar{\triangleright}_c)$. Now since $C(A \cup \{z\}) = y$ and $C(xyz) = z$, R-WARP implies that $y \in \min(\{x, y, z\}, \bar{\triangleright}_c)$. This implies $z \bar{\triangleright}_c y$ as $\neg x \bar{\triangleright}_c y$ by Lemma 2. By R-acyclicity, we have $C(yz) = z$. Since $C(S \cup \{z\}) = y$, by WCC there exists a $w \in S$ such that there is a (zy) reversal is due to w. This cannot be a strong reversal by Lemma 2 and therefore is a weak reversal. Therefore, by definition of $\triangleright_c$, we have $y \triangleright_c w$. This implies $z \bar{\triangleright}_c w$ by the transitivity of $\bar{\triangleright}_c$ and $C(zw) = z$ by R-Acyclicity.

Now, we will show that $C(xyw) = x$. By Lemma 3, we know that x is $\bar{\triangleright}_c$-unrelated to S and therefore $x \notin \min(\{x, y, w\}, \bar{\triangleright}_c)$. By R-WARP, we must have $C(xyw) \neq y$. Now, suppose $C(xyw) = w$. Then we have a (yw) reversal and since $y \rhd_c w$, by Lemma 2, it is a weak (yw) reversal due to x. By the transitivity of $\bar{\rhd}_c$, we have $y \bar{\rhd}_c x$, a contradiction. Therefore, we have $C(xyw) = x$.

Finally, we will show that $C(xyzw) = y$. Suppose we have $C(xyzw) = z$. Now, since $y \rhd_c w$, we have $y \notin \min(xyzw, \bar{\rhd}_c)$ and since $z \rhd_c y$, we have $z \notin \min(A \cup \{z\}, \bar{\rhd}_c)$. By R-WARP, $C(A \cup \{z\}) \neq y$, a contradiction. Now, suppose we have $C(xyzw) = x$. Then since $z \rhd_c x$, by Lemma 2, we have a weak (zx) reversal due to w. By definition, $x \rhd_c w$, violating Lemma 3 since x is $\bar{\rhd}_c$-unrelated to S. Now, suppose we have $C(xyzw) = w$. Then by WCC, the (yw) reversal is either due to x or z. Since $y \rhd_c w$, the (yw) reversal is a weak reversal and is not due to x by Lemma 3 and since $C(xy) = x$. If the reversal is due to z, then we have $y \bar{\rhd}_c z$ and since $C(zy) = z$, we have a violation of R-acyclicity. Therefore, we have $C(xyzw) = y$. $\square$

Using the above Lemma, we can say that if we have either a weak or a strong reversal, then it will be reflected in a small menu reversal i.e. on a set A such that $|A| \leq 3$. Now consider two CBR-representable choice functions C and $\bar{C}$ with same choices in small menus, but different choice in at least one set A where $|A| > 3$. W.L.O.G., let the choice in that set be $C(A) = x$ and $\bar{C}(A) = y$ where $x \neq y$. Let $C(xy) = x$. Hence, we have a (xy) reversal in the choice function $\bar{C}$. Since $\bar{C}$ is CBR representable, there exists a $z \in A$ such that there is a (xy) reversal due to z. If the reversal is weak, then we have giving yR^*z and yP^*x where (R^*, P^*) is a representation of C. Since $y \notin \min(A, R)$ and yP^*x, x cannot be chosen in a set containing y and z which is contradiction as $C(A) = x$. If the reversal is strong, then there are two possible cases:

(i) $x \succ_c y \succ_c z \succ_c x$ and $C(xyz) = y$. Here the (xy) reversal is seen in a small menu giving $x\bar{P}y$ and $\neg x\bar{P}y$, $y\bar{P}z$, $z\bar{R}x$, where $(\bar{R}, \bar{P})$ is a representation of $\bar{C}$. Note that $C(xyz) = y$ and $C(A) = x$ where $\{xyz\} \subset A$. As $C(xz) = z$, we have a (zx) reversal due to some $w \in S$. Knowing that $z\bar{R}x$ is true, this is a weak reversal, which is seen in small menus implying $x\bar{R}w$ holds. This means $x \notin \min(S, \bar{R})$ contradicting $\bar{C}(A) = y$.

(ii) $z \succ_c x \succ_c y$, $z \succ_c y \succ_c w$, $C(xyz) = z$ and $C(yzw) = y$ for some $w \in X$. Note that since $C(A) = x$ and $C(xz) = z$, again we have a weak (zx) reversal due to some $k \in S$, leading to a contradiction to $\bar{C}(A) = y$ as above.

Therefore, we must have $\bar{C}(A) = x$. Since A was arbitrary, we have shown that $C = \bar{C}$.

2.8.4 Proofs from Section 2.5

Proof of Proposition 1

Suppose for some A, A' and $A'' \in \mathcal{P}(X)$ with $A \subset A' \subset A''$, we have $C(A) = C\{x,y\} = x$ and $C(A') = y$. Note that we have a (xy) reversal. By WCC, there exists a $z \in S$ that causes this reversal. If $C(xz) = x$ then it is a weak (xy) reversal. By definition, $x \not\triangleright_c y$ and $y \not\triangleright_c z$, implying $C(A'') \neq y$ by R-WARP. If $C(xz) = z$, this implies a strong (xy) reversal. Since $z \not\triangleright_c x$ there is a weak (zx) reversal due to some $z' \in S$ (by WCC). Therefore, we have $x \notin \min(S', \not\triangleright_c)$, which implies that $C(S') \neq y$ by R-WARP.

Proof of Proposition 3 and 4

To prove the result, we first prove two intermediate results

Definition 8. *Negative Expansion (NE): For all $S, S' \supset \{x,y\}$,*

$$C(S) = C(S') = x \textit{ implies } C(S \cup S') \neq y$$

Lemma **7.** *If C is CBR-representable, then it satisfies Negative Expansion*

Proof. If possible, suppose choice function C violates NE. Then there exists $A, A' \supset \{x,y\}$ such that $C(A) = C(A') = x$ and $C(A \cup A') = y$. If $C(xy) = x$, then we have an (xy) reversal. By WCC, it is either weak or strong. Let (R,P) be a representation of the choice function. A weak (xy) reversal due to $z \in A$ and $z' \in A'$ implies yPx, yRz and yRz', therefore it must be that $y \notin \min(A,R)$. This implies $C(A) \neq x$. Also, if there is a strong (xy) reversal due to some $z \in A \cup A'$, then we have zRx. W.L.O.G suppose $z \in A$. Given $C(A) = x$, we know that $x \notin \min(A,R)$ and there is a $w \in A$ such that xRw holds. As $x \notin \min(A \cup A', R)$ and xPy holds, we have a contradiction to $C(A \cup A') = y$.
Now, if $C(xy) = y$, we have an (yx) double reversal. A weak (yx) reversal (due to some $z \in A$ and $z' \in A'$) implies xRz and xRz' . Given xPy and $x \notin \min(A \cup A', R)$, we cannot have $C(A \cup A') = y$. If the (yx) reversal is strong, then we have yPx. Since $C(A) = C(A') = x$, we must have $y \in \min(A,R)$ and $y \in \min(A',R)$ (thus $y \in \min(A \cup A')$) contradicting $C(A \cup A') = y$ $\square$

LEMMA **8.** *If C is CBR-representable, then a (xy) double reversal due to z_1, z_2 implies a strong (xy) reversal due to z_1 and a weak (z_1x) reversal due to z_2*

Proof. Let C be a CBR-representable choice function. An (xy) double reversal due to z_1, z_2 implies $x \succ_c y$ and there exist A, A' with $\{x,y\} \subset A' \subset A$ such that $C(A') = x, C(A' \cup z_1) = y$ $C(A) = y, C(A \cup z_2) = x$ and for all B, B', B'' with $\{x,y\} \subset B' \subset B \subset B''$, if $C(B) = C\{x,y\} = x$ and $C(B') = y$, then $C(B'') \neq y$. As C satisfies WCC and Exclusivity, each reversal is either weak or strong. If the first (xy) reversal is weak, then we have xRy, yRz_1 and yPx. As $x, y \notin \min(B, R)$ for all $B \supset A$, there can be no double reversal due to R-WARP. Thus, the first reversal is strong, implying z_1Rx, xPy and yPz_1. For x to be chosen again in $A \cup \{z_2\}$, it must be that xRz_2 and xPz_1 hold. This implies that $z_1 \succ_c x \succ_c z_2$ and $C(xz_1z_2) = x$ and hence a weak (z_1x) reversal due to z_2 □

Let us prove the if part. Consider a CBR-representable choice function C which is also RSM. If possible, for some x, y we have a weak (xy) reversal due to some z. We have $x \succ_c y \succ_c z$ and $C(xyz) = y$. However, this violates Expansion as $C(xy) = C(xz) = x$, but $C(xyz) = y$. As RSM satisfies Expansion, this is a contradiction.

Now consider C which is CBR, with no weak reversals. By WCC, if there is any reversal, it has to be strong. If possible, let C violate Expansion, i.e. there exists A, A' such that $C(A) = C(A') = x$, but $C(A \cup A') = y \neq x$. If $\{x,y\} \subset A \cap A'$, this violates NE leading to a contradiction (by Lemma 7). WLOG, let $y \in A \setminus A'$. If $C(xy) = y$, we have double reversal which is a contradiction by Lemma 8. Thus, $C(xy) = x$ implying a (xy) strong reversal due to some $z \in A'$. We know that xPy and zRx hold and since $C(A') = x$, there exists a $w \in A'$ such that xRw is true. Note that this implies $x \notin \min(A \cup A', R)$ which implies $C(A \cup A') \neq y$. Now, if C violates WWARP, given it satisfies R-WARP, there is a double reversal. But, that is equivalent to a strong and a weak reversal which is a contradiction. Thus, C satisfies Expansion and WWARP, implying RSM representation.

The argument for the proof of Proposition 4 is analogous to that of Proposition 3, and is therefore omitted.

2.8.5 Independence of Axioms in Theorem 1

Example 1. The choice function below satisfies R-acyclicity, WCC and R-WARP but violates NC: $X = \{x, y, z\}$

S	**C(S)**	**S**	**C(S)**
$\{x, y\}$	y	$\{x, y, z\}$	x
$\{x, z\}$	z		
$\{y, z\}$	y		

Example 2. The choice function below satisfies R-acyclicity, NC and R-WARP but violates WCC: $X = \{x, y, z, w\}$

S	**C(S)**	**S**	**C(S)**	**S**	**C(S)**
$\{x, y\}$	x	$\{x, y, z\}$	z	$\{x, y, z, w\}$	y
$\{x, z\}$	z	$\{x, y, w\}$	w		
$\{x, w\}$	w	$\{x, z, w\}$	w		
$\{y, z\}$	z	$\{y, z, w\}$	y		
$\{y, w\}$	y				
$\{z, w\}$	w				

Example 3. The choice function below satisfies NC, WCC and R-WARP but violates R-acyclicity: $X = \{x, y, z, w\}$

S	**C(S)**	**S**	**C(S)**	**S**	**C(S)**
$\{x, y\}$	x	$\{x, y, z\}$	y	$\{x, y, z, w\}$	y
$\{x, z\}$	z	$\{x, y, w\}$	y		
$\{x, w\}$	x	$\{x, z, w\}$	x		
$\{y, z\}$	y	$\{y, z, w\}$	z		
$\{y, w\}$	y				
$\{z, w\}$	z				

Example 4. The choice function below satisfies NC, WCC and R-acyclicity but violates R-WARP: $X = \{x, y, z, w\}$

S	**C(S)**	**S**	**C(S)**	**S**	**C(S)**
$\{x, y\}$	x	$\{x, y, z\}$	y	$\{x, y, z, w\}$	x
$\{x, z\}$	x	$\{x, y, w\}$	x		
$\{x, w\}$	x	$\{x, z, w\}$	x		
$\{y, z\}$	y	$\{y, z, w\}$	y		
$\{y, w\}$	y				
$\{z, w\}$	z				

Chapter 3

Copeland Choice Rules

3.1 Introduction

Tournaments, binary relations that are *complete* and *asymmetric*, arise in various settings in economic theory. In the context of social choice, they represent the pairwise majority relation of an underlying population. In the context of individual choice, they may represent the pairwise relation generated by a multi-self or a multi criteria evaluation of a decision maker (DM). As is well known, in the presence of pairwise cycles, the maximal set may turn out to be empty. In order to address this problem, various tournament "solutions" have been proposed in the literature that provide some form of approximation of the maximal set. The most natural and immediate solution is to select the set of alternatives that have the highest number of pairwise "wins". That is, an alternative is selected if it "beats" the maximum number of alternatives. The set of all such alternatives is popularly referred to as the *Copeland Set*.[17] Further, Copeland scores —number of pairwise wins of alternatives —also provide a natural way to rank alternatives for a given tournament. This paper provides examines the Copeland set and Copeland scores from a revealed preference perspective.

We first study a DM who is endowed with a *fixed* tournament and from every menu of alternatives, chooses the Copeland set of the tournament restricted to that menu. We

[17] Although attributed to Copeland (1951), versions of this method have also been suggested by Ramon Llull in the thirteenth century (see Colomer (2013)) and Ernst Zermelo (see Zermelo (1929) and Moon (2015)).

term such correspondences *Copeland rules*. In contrast with other characterizations of the Copeland set in the literature that vary the tournament and study the property of the "solution" as a choice correspondence (see Henriet (1985)), we fix the tournament and only vary the set of feasible alternatives, the menus. This is in line with the revealed preference approach where the only observable data is that of the choices made by the DM. Such revealed preference analysis has been conducted for some other tournament solutions such as the top-cycle set, the uncovered and the minimal covering set (see Ehlers and Sprumont (2008) , Lombardi (2008) and Lombardi (2009a)).

In addition to its simplicity, the Copeland set also possesses some desirable properties. A basic criteria for any tournament solution is *Condorcet consistency*. This requires that if an alternative beats every other alternative in a pairwise comparison, then it must be the unique choice. In addition to Condorcet consistency, the Copeland set has been characterized by the "minisum" property: it is the set of alternatives that beat all the alternatives in the smallest total number of steps (see Sanver and Selçuk (2010)), thus making it "closest" to a Condorcet winner if it does not exist. From the computational point of view, it is simple as it can be computed in linear time. In terms of decisiveness, it is a subset of the uncovered set which is itself a subset of the top cycle set (see Chapter 3 of Brandt et al. (2016)). That is, for an arbitary tournament on a set, the Copeland set is always contained within the uncovered set which is contained within the top-cycle set.

Unlike the characterization of the Copeland set, we use axioms that model how observed data may look if a DM used Copeland set to make choices. In our characterization of the Copeland rules, we introduce two new axioms : *Symmetry* and *Responsiveness*. These axioms are based on an observation that the Copeland score of any alternative in a menu can change by at most 1 upon addition or removal of another alternative to a menu. Symmetry requires that the chosen alternatives in a menu are treated symmetrically when an alternative that beats the chosen alternatives or is beaten by the chosen alternatives in pairwise comparisons is added to the menu. On the other hand, Responsiveness requires that if for a given menu and two chosen alternatives in it, a third alternative that is beaten by one chosen alternative and beats the other chosen alternative is added to that menu, then the latter alternative cannot be chosen in the enlarged menu. That is, the choice is "responsive" to the added alternative. In addition to these two axioms, we require weakenings of three standard consistency conditions used in choice theory to characterize Copeland rules.

Next, we study a DM who chooses in a probability distribution over all alternatives for every menu. She is endowed with a tournament *and* a scoring function. Using the scoring function, she attaches a real number, the score, to every alternative based on its Copeland score in a given menu. She assigns every alternative a probability that is equal to the relative score of that alternative in the menu. This way of choosing is closely related to the famous *Luce* model (Luce (1956)). However, it is distinct from the Luce model since the "Luce weights" of alternatives change across menus, leading to menu-dependent behavior. We term this procedure Copeland Stochastic Choice Rule (CSCR).

The "Luce" way of choosing the alternatives is reflected in the behavioral characterization of this procedure. An adaptation of Luce's Independence of Irrelevant Alternatives (IIA) condition is the main axiom in the characterization of this procedure. We discuss two variants of CSCR : with and without abstention. That is, in the model with abstention, DM has the option to abstain from making choice in every menu (see Manzini and Mariotti (2014)). The adaptation of the IIA condition along with a stochastic version of the above mentioned responsiveness axiom and two mild axioms, characterizes the CSCR with abstention. To characterize the model without abstention, we modify the IIA condition along the lines of Echenique and Saito (2019) and introduce a cyclical independence axiom. This, together with the previous set of axioms characterizes the CSCR model without abstention.

The outline of the paper is as follows. In the next section, we formally introduce Copeland rules and provide their behavioral characterization. In Section 2.3 we introduce and behaviorally characterize CSCR. In Section 2.4, we provide discussion of the relation between our characterization of Copeland rules and that of top-cycle and uncovered set rules.

3.2 Copeland Rules

Throughout the paper, we denote by X a finite non-empty set of alternatives. Let $\mathcal{P}(X)$ denote the set of all non-empty subsets of alternatives. A menu A is an element of $\mathcal{P}(X)$. A binary relation R over X is a subset of $X \times X$. For any $x, y \in X$, $(x, y) \in R$ is written as xRy. A binary relation is asymmetric if for any $x, y \in X$, xRy implies $\neg yRx$. It is complete if for any distinct $x, y \in X$, either xRy or yRx. It is transitive if for any distinct $x, y, z \in X$, xRy and yRz implies xRz. A complete and asymmetric binary relation is called a *tournament* and we will denote it by T. For a given tournament T

and a menu $A \in \mathcal{P}(X)$, we denote by $T|_A$, the restriction of T to A. It is easy to verify that $T|_A$ is a tournament on A. In some places, whenever no confusion arises, we abuse notation and write a set, say $\{x, y, z\}$, as xyz. For a given tournament T and menu A, the Copeland score of an alternative $a \in A$ is given by the following

$$\text{Cop}_T(a, A) \coloneqq |\{x \in A : aTx\}|$$

The Copeland set of A, $\text{Cop}_T(A)$ is the set of all alternatives that have the highest Copeland scores in it. That is, $\text{Cop}_T(A) \coloneqq \{x \in A : \text{Cop}_T(x, A) \geq \text{Cop}_T(y, A) \; \forall y \in A\}$. For a given tournament T and a menu A, an alternative $x \in A$ is a *Condorcet* winner in A if it beats every alternative of $A \setminus \{x\}$ in pairwise comparison. That is, xTy for all $y \in A \setminus \{x\}$. It is easy to see that if a Condorcet winner exists in $T|_A$, then it is the only alternative in the Copeland set of A.

A choice rule is a map $C : \mathcal{P}(X) \to \mathcal{P}(X)$ such that $C(A) \subseteq A$ for all $A \in \mathcal{P}(X)$. A choice rule is *resolute* if $|C(A)| = 1$ for all A such that $|A| = 2$.[18] We say that an alternative x *dominates* a set A if $\{x\} = C(xy)$ for all $y \in A$ and is *dominated* by a set A if $x \notin C(xy)$ for all $y \in A$. Now, we define Copeland rules formally as the choice correspondences that select from every menu, the Copeland set according to a fixed underlying tournament restricted to that menu.

Definition 1. *A choice rule C is a Copeland rule if there exists a tournament T such that $C(A) = Cop_T(A)$ for all $A \in \mathcal{P}(X)$.*

We require five simple axioms to characterize Copeland rules. The first axiom is a standard "expansion" axiom that also appears in the characterization of top cycle rules and uncovered set rules by Ehlers and Sprumont (2008) and Lombardi (2008), respectively. This axiom ensures that the Condorcet winner —if it exists —is the unique choice in a menu.

Binary Dominance Consistency(BDC): A choice rule C satisfies BDC if for all $A \in \mathcal{P}(X)$ and $x \in A$, if $\{x\} = C(xy)$ for all $y \in A$, then $C(A) = \{x\}$.

The second axiom we need is a weakening of the well known condition α introduced in Sen (1971). Condition α requires a chosen alternative from a menu to be chosen in any sub-menu if it is feasible. We weaken it by requiring a chosen alternative from a menu to be chosen after removal of a "dominating" alternative —one that beats it in a pairwise comparison. The idea behind this condition is that if an alternative is chosen

[18]The condition of resoluteness appears in Ehlers and Sprumont (2008) and Lombardi (2008).

a menu and an alternative that pairwise beats it is removed from the menu, then this "boosts" that alternative and therefore, it must still be chosen.

Independence of Dominating Alternatives (IDA): A choice rule C satisfies IDA if for all $A \in \mathcal{P}(X)$ and $x, y \in A$ such that $C(xy) = \{y\}$, if $x \in C(A)$, then $x \in C(A \setminus \{y\})$.

The third axiom we require is a weakening of Weakened WARP (WWARP) condition that appears in Ehlers and Sprumont (2008) in their characterization of top-cycle and upper class rules. Weakened WARP itself is a weakening of the Weak Axiom of Revealed Preference (WARP) which requires that if an alternative x is chosen and another alternative y is rejected in some menu, then in whenever x is present in another menu, then y cannot be chosen. Weakened WARP weakens WARP in the following way: if an alternative x is chosen and another alternative y is rejected in some menu, then whenever x is present in another menu *and* if y is chosen in that menu, then x must also be chosen. Alternatively put, if x is chosen and y is rejected in a menu, then y cannot be chosen *and* x rejected in another menu.

Our weakening of WWARP requires WWARP to be applicable on not arbitrary menus but for only certain menus. We require that if x is chosen and y is rejected in a menu, then upon addition of any alternative, if x is chosen, then y must also be chosen.

Weak WWARP: A choice rule C satisfies Weak WWARP if for all $A \in \mathcal{P}(X)$ and $x, y \in A$, if $x \in C(A)$ and $y \in A \setminus C(A)$, then there does not exist $z \in X \setminus A$ such that $y \in C(A \cup \{z\})$ and $x \in A \cup \{z\} \setminus C(A \cup \{z\})$.

The fourth axiom we require is the choice theoretic adaptation of the "positive responsiveness" axiom that Rubinstein (1980) uses to axiomatically characterize the method that ranks alternatives based on their Copeland scores. The axiom by Rubinstein requires that if an alternative is ranked weakly better than another alternative for a *given* tournament and if the tournament is modified by only increasing the Copeland score of the weakly preferred alternative, then it must be ranked strictly better than the other alternative in this modified new tournament. We adapt this to our setting in the following way: if two alternatives are chosen in a menu and an alternative added to the menu is beaten by one and beats the other in pairwise comparison, then the latter alternative cannot be chosen in the new menu. That is, the choice is "responsive" to the added alternative.

Responsiveness: A choice rule C satisfies Responsiveness if for all $A \in \mathcal{P}(X)$ and

$x, y \in C(A)$ and $z \in X \setminus A$ such that $\{x\} = C(xz)$, if $y \notin C(yz)$ then $y \notin C(A \cup \{z\})$.

The last axiom requires the chosen alternatives to be treated "symmetrically" upon addition of an alternative that is symmetric to them in the following sense: it either beats both of them or is beaten by both of them in pairwise comparisons. It says that if one chosen alternative is chosen (not chosen) in the expanded menu, then the other alternative should also be chosen (not chosen).

Symmetry: A choice rule C satisfies if for all $A \in \mathcal{P}(X)$ and $x, y \in C(A)$ and $z \in X \setminus A$ such that z dominates or is dominated by $\{x, y\}$, if $x \in C(A \cup \{z\})$ then $y \in C(A \cup \{z\})$.

Now, we are ready to characterize the Copeland choice rule using the above stated axioms.

Theorem 1. *A resolute choice rule is a Copeland rule if and only if it satisfies BDC, IDA, Weak WWARP, Symmetry and Responsiveness.*

Proof. First, we show the necessity of the axioms. Consider a Copeland rule C generated by a tournament T. The asymmetry of T ensures that C is resolute. Since C is Condorcet consistent, BDC is immediate. To show a C satisfies IDA, consider an arbitrary $A \in \mathcal{P}(X)$ and $x \in C(A)$. That is, $x \in \mathrm{Cop}_T(A)$. Let yTx implying $\{y\} = C(xy)$ for some $y \in A \setminus \{x\}$. Since the Copeland score of an alternative cannot increase by shrinking the set, we have $x \in \mathrm{Cop}_T(A \setminus \{y\}) = C(A \setminus \{y\})$. To show Weak WWARP, suppose for some menu A and $x, y \in A$, we have $x \in C(A)$ and $y \in A \setminus C(A)$. Then we know that $\mathrm{Cop}_T(x, A) > \mathrm{Cop}_T(y, A)$ and therefore for any $z \in X \setminus A$, we have $\mathrm{Cop}_T(x, A \cup \{z\}) \geq \mathrm{Cop}_T(y, A \cup \{z\})$ (since the Copeland score can increase by at most 1) and hence we cannot have $y \in \mathrm{Cop}_T(A \cup \{z\}) = C(A \cup \{z\})$ and $x \notin \mathrm{Cop}_T(A \cup \{z\}) = C(A \cup \{z\})$. To show Symmetry consider an arbitrary A and $x, y \in C(A)$. We know that $\mathrm{Cop}_T(x, A) = \mathrm{Cop}_T(y, A)$. Suppose some z dominates or is dominated by $\{x, y\}$, then we have $\mathrm{Cop}_T(x, A \cup \{z\}) = \mathrm{Cop}_T(y, A \cup \{z\})$ and therefore if $x \in C(A \cup \{z\})$, then $y \in C(A \cup \{z\})$. Finally, to show Responsiveness, consider an arbitrary A and $x, y \in C(A)$. We know that $\mathrm{Cop}_T(x, A) = \mathrm{Cop}_T(y, A)$. Suppose some $\{x\} = C(xz)$ and $y \notin C(yz)$ for some $z \in X \setminus A$. Then, we have $\mathrm{Cop}_T(x, A \cup \{z\}) > \mathrm{Cop}_T(y, A \cup \{z\})$ implying y cannot be in the Copeland set of $A \cup \{z\}$.

Now, we prove the sufficiency of the axioms. Construct the revealed tournament T as xTy if and only if $C(xy) = \{x\}$ for any distinct $x, y \in X$. The binary relation T is complete by resoluteness and asymmetric by definition. Therefore, it is a tournament.

Also, note that by definition, $\text{Cop}_T(A) = C(A)$ for all A such that $|A| = 2$. Consider an arbitrary $A \in \mathcal{P}(X)$ such that $|A| = 3$. Let $A = \{x_1, x_2, x_3\}$. There are two possible cases. First, $x_i T x_j$ for some $i \in \{1, 2, 3\}$ and for all $j \neq i, j \in \{1, 2, 3\}$. In this case by BDC, we have $\text{Cop}_T(A) = C(A)$. Second, there is a cycle i.e. $x_i T x_j T x_k T x_i$ for $i, j, k \in \{1, 2, 3\}$. Since $C(A) \neq \emptyset$, W.L.O.G let $x_i \in C(A)$. Since $x_k T x_i$, we have $x_k \in C(A)$ by Weak WWARP. Since $x_j T x_k$, we have $x_j \in C(A)$, again by Weak WWARP. Therefore, we have $C(A) = A = \text{Cop}_T(A)$.

Now, we will proceed by strong induction on the cardinality of the menus. Consider an arbitrary $k \geq 3$ and assume that $C(A) = \text{Cop}_T(A)$ for all A such that $|A| \leq k$. Consider an arbitrary A such that $|A| = k + 1$. First we show that $C(A) \subseteq \text{Cop}_T(A)$. Consider $x \in C(A)$. Assume for contradiction that $x \notin \text{Cop}_T(A)$. We know that $\text{Cop}_T(A) \neq \emptyset$. So consider an arbitrary $y \in \text{Cop}_T(A)$. We will consider two possible cases:

(i) $\text{Cop}_T(x, A) + 1 < \text{Cop}_T(y, A)$. Since $|A| \geq 4$, there exists at least one alternative $z \neq y$ such that yTz and zTx. Since $\text{Cop}_T(x, A \setminus \{z\}) < \text{Cop}_T(y, A \setminus \{z\})$, we have $x \notin \text{Cop}_T(A \setminus \{z\})$ and by our inductive hypothesis, $x \notin C(A \setminus \{z\})$. Since, $\{z\} = C(xz)$, we must have $x \notin C(A)$ by IDA, a contradiction.

(ii) $\text{Cop}_T(x, A) + 1 = \text{Cop}_T(y, A)$. Here, we have two further subcases: (a) y covers x i.e. yTx and $xTz \implies yTz$. Note that if y is the Condorcet winner (yTz for all $z \in A \setminus \{y\}$), then we have $y = C(A)$ implying $x \notin C(A)$. Therefore, there exists $z \neq x, y$ such that zTy. Since $\text{Cop}_T(x, A) = \text{Cop}_T(y, A) + 1$ and y covers x, we have zTx as well. Then it follows that $\text{Cop}_T(x, A \setminus \{z\}) + 1 = \text{Cop}_T(y, A \setminus \{z\})$ and by our inductive hypothesis $x \notin C(A \setminus \{z\})$. Again, by IDA, we must have $x \notin C(A)$, a contradiction. (b) y does not cover x. Then there exists $z \neq x, y$ such that yTz and zTx. So $\text{Cop}_T(x, A \setminus \{z\}) = \text{Cop}_T(y, A \setminus \{z\})$. If $x \notin C(A \setminus \{z\})$, then by IDA , we have $x \notin C(A)$, a contradiction. If $x \in C(A \setminus \{z\})$, then by our inductive hypothesis, we have $y \in C(A \setminus \{z\})$. Since $\{y\} = C(yz)$ and $x \notin C(xz)$, by Responsiveness, we must have $x \notin C(A)$, a contradiction. Therefore, we have $x \in \text{Cop}_T(A)$ and we have shown that $C(A) \subseteq \text{Cop}_T(A)$.

To show $\text{Cop}_T(A) \subseteq C(A)$, consider an arbitrary $x \in \text{Cop}_T(A)$. If x is the Condorcet winner, then by BDC, we have $\{x\} = C(A)$. If x is the only alternative in the Copeland set i.e. $\{x\} = \text{Cop}_T(A)$, then since $C(A) \neq \emptyset$ and $C(A) \subseteq \text{Cop}_T(A)$, we must have $\{x\} = C(A)$. Suppose $|\text{Cop}_T(A)| \geq 2$ and there exists $y \in \text{Cop}_T(A)$ and $y \in C(A)$ but $x \notin C(A)$. Since x is not the Condorcet winner, there exists z such that zTx. Again, we will consider two possible cases:

(i) $z \neq y$. If zTy , then we have $x, y \in \text{Cop}_T(A \backslash \{z\})$ and by our inductive hypothesis, we have $x, y \in C(A \setminus \{z\})$. Since $y \in C(A)$, we must have $x \in C(A)$ by Symmetry. Suppose yTz. Then, we know $x \in \text{Cop}_T(A \setminus \{z\})$ but $y \notin \text{Cop}_T(A \setminus \{z\})$ and by our inductive hypothesis, we have $x \in C(A \setminus \{z\})$ and $y \notin C(A \setminus \{z\})$. Since $y \in C(A)$, we must have $x \in C(A)$ by Weak WWARP.

(ii) y is the only alternative in A such that yTx. Then we have xTz for all $z \in A \backslash \{x, y\}$. Since $\text{Cop}_T(x, A) = \text{Cop}_T(y, A)$, there exists $z \in A \backslash \{x, y\}$ such that zTy and xTz. Note that since xTz for all $z \in A\{x, y\}$, we have $\text{Cop}_T(x, A) = \text{Cop}_T(y, A) = |A| - 2$. Therefore there exists exactly one such z such that zTy. Since $|A| \geq 4$, there exists $w \neq z$ such that yTw and xTw implying $\text{Cop}_T(y, A \setminus \{w\}) = \text{Cop}_T(x, A \setminus \{w\}) = |A \setminus \{w\}| - 2$. We claim that $x \in C(A \backslash \{w\})$ (and by our inductive hypothesis $y \in C(A \backslash \{w\})$). To show this, it is sufficient to establish that $x \in \text{Cop}_T(A \setminus \{w\})$ and the claim follows by our inductive hypothesis. If $x \notin \text{Cop}_T(A \setminus \{w\})$, then there exists some $w' \in A \setminus \{w\}$ such that $\text{Cop}_T(x, A \setminus \{w\}) < \text{Cop}_T(w', A \setminus \{w\})$ i.e. $\text{Cop}_T(w, A \setminus \{w\}) = |A \setminus \{w\}| - 1$ implying w' is the Condorcet winner. Since xTz, this alternative cannot be z. Further, since we have yTw' and xTw' for all $w' \in A \setminus \{w, z\}$, there cannot be a Condorcet winner in the set $A \setminus \{w\}$. Therefore $x \in \text{Cop}_T(A \setminus \{w\})$ and $y \in \text{Cop}_T(A \setminus \{w\})$ implying $x, y \in C(A \setminus \{w\})$ and hence by Symmetry, we must have $x \in C(A)$. Therefore we have shown that $C(A) = \text{Cop}_T(A)$ which completes the proof. $\square$

3.3 Copeland Stochastic Choice

We now introduce a stochastic choice model that uses the Copeland scores to generate the choice probabilities in any menu. An intuitive way to choose stochastically is to assign a probability to each alternative equal to its relative Copeland score in a menu. That is, the probability of an alternative x in a menu A is

$$\frac{\text{Cop}_T(a, A)}{\sum_{a' \in A} \text{Cop}_T(a', A)}$$

This is similar to the Luce model of stochastic choice in which the decision maker has positive weights given by a function $u : X \to \mathbb{R}_{++}$ that are *fixed* across menus and the probability of each alternative is its relative weight in each menu. However, in our case, since the Copeland scores of alternatives vary across menus, this leads to a menu-dependence which differentiates it from the Luce model. In what follows, we generalize this idea of assigning the probabilities to alternatives in a menu based on

Copeland scores and endow the decision maker with a scoring function that uses the Copeland scores to generate the "Luce weights" in every menu.

In addition to the setup in the previous section, here we allow the DM to abstain from making a choice as in Manzini and Mariotti (2014). This is done by adding a "default" alternative to every menu $A \in \mathcal{P}(X)$. This enables us to characterize the model in terms of a Luce IIA type condition. In the next section, we drop this assumption and characterize a more general model. Denote by X^* the "augmented" set of alternatives i.e. $X^* := X \cup \{a^*\}$ and for any $A \in \mathcal{P}(X)$, let $A^* := A \cup \{a^*\}$. We will follow the convention as in Manzini and Mariotti (2014) and will suppress the presence of the default alternative in every menu. With the addition of the option to abstain, the Copeland score can be analogously defined as

$$\mathrm{Cop}_T(a, A) := |\{x \in A^* : aTx\}|$$

A *stochastic choice rule* (scr) is a map $p : X^* \times \mathcal{P}(X) \to [0, 1]$ such that $\sum_{a \in A^*} p(a, A) = 1$ for all $A \in \mathcal{P}(X)$ and $p(b, A) = 0$ when $b \notin A^*$. A stochastic choice rule is *positive* if for any $A \in \mathcal{P}(X)$ and $a \in A^*$, we have $p(a, A) > 0$. We will focus on positive stochastic choice rules in this paper and unless specified otherwise, a stochastic choice rule is considered to be positive.

For a given tournament T, let n_T denote the maximum Copeland score that any alternative has in X and let $N_T := \{0, 1, \ldots, n_T\}$. A scoring function is a map $S : N_T \to R_{++}$ that assigns a real number to every Copeland score. Given a tournament T, a scoring function is monotone with respect to T if for any $A \in \mathcal{P}(X)$ and $a, b \in A$, $\mathrm{Cop}_T(a, A) > \mathrm{Cop}_T(b, A)$ implies $S(\mathrm{Cop}_T(a, A)) > S(\mathrm{Cop}_T(b, A))$. The DM is endowed with a tournament T and scoring function that is monotone with respect to T and assigns a probability equal to the relative score of an alternative in a menu. The monotonicity of the scoring function captures the idea that the DM "prefers" the alternatives that have a higher Copeland score than the ones having lower Copeland scores. Since the only requirement that the scoring function respect the ranking generated by Copeland score, this can also acccomodate "intensity" of such "preferences". Further, simply assigning the probability of each alternative as its relative Copeland score is a special case of the scoring function- the identity map $S(i) = i$ for all $i \in N_T$. Now, we define the choice procedure formally.

Definition 2. *An scr is Copeland Stochastic Choice Rule (CSCR) if there exists a tournament T and scoring function S that is monotone with respect to T such that for* *all* $A \in \mathcal{P}(X)$ *and* $a \in A$,

$$p(a, A) = \frac{S(Cop_T(a, A))}{\sum_{b \in A^*} S(Cop_T(b, A))}$$

To state our axioms, we need some notation. Define the revealed relation $\succ_t$ as follows: $a \succ_t b$ if and only if $p(a, \{a, b\}) > p(a, \{a, b\})$. For any set $A \in \mathcal{P}(X)$ and $a \in A$, denote by $s(a, A)$ the number of alternatives a "beats" in the set A according to the revealed relation $\succ_t$, i.e. $s(a, A) = |\{x \in A : a \succ_t x\}|$. So, $s(a, A)$ can be understood as the revealed Copeland score of an alternative a in the menu A. Our first axiom is a version of Luce's IIA adapted to the setting of tournaments. It states that for any two menus A and A', if there exist a two pairs of alternatives a, a' and b, b' —where $a, b \in A$ and $a', b' \in A'$—that have the same revealed Copeland score, then the ratio of their probabilities will be the same. An axiom similar to ours appears in Tserenjigmid (2021) in the setting of stochastic choice over lists. Our axiom is formally stated as follows.

T-IIA: An scr p satisfies T-IIA if for any $a, b, a', b' \in X^*$ and $A, A' \in \mathcal{P}(X)$ if $s(a, A) = s(a', A')$ and $s(b, A) = s(b', A')$, then

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

Our second axiom restricts the possibility of probability "reversals" between a pair of alternatives upon addition of an alternative. This can be seen as a stochastic analogue of the Responsiveness axiom discussed in Section 2. It states that if for a pair of alternatives $a, b \in A$ if a has a higher probability of being chosen, then the addition of an alternative c in A that is relatively "favorable" towards a than b cannot lead to a probability reversal between a and b. Here, the term favorable means that in pairwise comparisons with the added alternative, a is at least as good as b. It is formally stated as follows.

No Probability Reversal(NPR): An scr p satisfies NPR if for any A, $a, b \in A$ and $c \notin A$, if $p(a, A) \geq p(b, A)$ and $p(a, \{a, c\}) \geq p(b, \{b, c\})$ with atleast one inequality strict, then

$$p(a, A \cup \{c\}) > p(b, A \cup \{c\})$$

Note that the axiom requires at least one inequality to be strict. In the case where both are equalities, we can show —in conjunction with T-IIA—that the probabilities of both alternatives remain the same in the expanded set i.e. $p(a, A \cup \{c\}) = p(b, A \cup \{c\})$.

We need another mild axiom that ensures the asymmetry of the underlying tournament. This axiom prohibits ties in binary comparisons between two alternatives. It states whenever a menu consists of two alternatives (excluding the default alternative), both of them cannot be chosen with equal probabilities.

No Binary Ties (NBT): An scr p satisfies NBT if for any $a, b \in X$, $p(a, \{a, b\}) \neq p(b, \{a, b\})$.

Our final axiom is a mild restriction which says that abstention is unlikely when the menus are singletons i.e. it is unlikely that the DM will not choose to abstain when only one alternative is offered in a menu. A stronger version of this axiom appears in Gerasimou (2020) as an axiom called "Desirability" that requires singleton probabilities to be 1.

No Binary Abstention (NBA): An scr p satisfies NBA if for any $a \in X$, $p(a, \{a\}) > p(a^*, \{a\})$.

We show below that these four axioms characterize Copeland Stochastic Choice Rules.

Theorem 2. *An scr is a CSCR if and only if it satisfies T-IIA, NPR, NBA and NBT. Further, the identified tournament is unique and the scoring function is unique upto multiplication by a positive constant.*

Proof. Define the revealed tournament T as follows: aTb if and only if $p(a, \{a, b\}) > p(b, \{a, b\})$. Completeness follows from from NBT and asymmetry follows from construction . Note that for the revealed tournament, the Copeland score of any alternative a in a menu A corresponds to $s(a, A)$. Also, NBA implies that a^* is the Condorcet loser[19] in any set A and hence $s(a^*, A) = 0$ for any $A \in \mathcal{P}(X)$. Before defining the scoring function, we first prove the following useful lemma

Lemma **1.** *For any $a, b \in X$ and $A \in \mathcal{P}(X)$,*

$$s(a, A) > s(b, A) \iff p(a, A) > p(b, A)$$

Proof. Consider any arbitrary a, b and $A \in \mathcal{P}(X)$ with $a, b \in A$. Suppose $s(a, A) =$

[19] A condorcet loser is an alternative that is beaten by every other alternative in a pariwise comparison

$s(b, A)$. Letting $a' = b$,$b' = a$ and $A = A'$, by T-IIA, we have

$$\frac{p(a,A)}{p(b,A)} = \frac{p(b,A)}{p(a,A)}$$
$$\implies p(a,A)^2 = p(b,A)^2$$
$$\implies p(a,A) = p(b,A)$$

To show $s(a, A) > s(b, A)$ implies $p(a, A) > p(b, A)$, we will consider the following two cases:

(i) $s(a, A) > s(b, A) + 1$. Let $A_a = \{x \in A | aTx\}$ and $A_b = \{x \in A | bTx\}$. We know that $|A_a| > |A_b| + 1$. Pick an arbitrary $\bar{A}_a \subset A_a$ such that $A_a \cap A_b \subseteq \bar{A}_a$ and $|\bar{A}_a| = |A_b|$. Let $\hat{A}_a = A_a \setminus \bar{A}_a$. Note that since $|A_a| > |A_b| + 1$, the set $\hat{A}_a$ is non-empty and all alternatives in $\hat{A}_a$ beat b and are beaten by a. Now, consider the set $A \setminus \hat{A}_a$. By construction, we have $s(a, A \setminus \hat{A}_a) = s(b, A \setminus \hat{A}_a)$ and therefore T-IIA implies that $p(a, A \setminus \hat{A}_a) = p(b, A \setminus \hat{A}_a)$. Note that $p(a, \{a, a'\}) > p(a', \{a, a'\})$ and $p(a', \{b, a'\}) > p(b, \{b, a'\})$ for all $a' \in \hat{A}_a$. Observe that $p(a, \{a, a'\}) = p(a', \{b, a'\})$ and $p(a', \{a, a'\}) = p(b, \{b, a'\})$ by T-IIA and the fact that $\sum_{a \in A^*} p(a, A) = 1$ for all A. Thus, we have $p(a, \{a, a'\}) > p(b, \{b, a'\})$ for all $a' \in \hat{A}_a$ by T-IIA. Now consider any arbitrary $a' \in \hat{A}_a$. Since $p(a, A \setminus \hat{A}_a) = p(b, A \setminus \hat{A}_a)$, by NPR we have $p(a, \{a'\} \cup A \setminus \hat{A}_a) > p(b, \{a'\} \cup A \setminus \hat{A}_a)$. Now, by repeated application of NPR, we get $p(a, A) > p(b, A)$.

(ii) $s(a, A) = s(b, A) + 1$. Suppose there exists $c \in A \setminus \{a, b\}$ such that aTc and cTb. Then consider the set $A \setminus \{c\}$. We know that $s(a, A \setminus \{c\}) = s(a, B \setminus \{c\})$ and by the argument in part (i) and NPR, we have $p(a, A) > p(b, A)$. Suppose there does not exist $c \in A \setminus \{a, b\}$ such that aTc and cTb. Therefore $aTc \implies bTc$ for all $c \in A \setminus \{b, c\}$. Since $s(a, A) = s(b, A) + 1$, we must have aTb and therefore we know that $p(a, \{a, b\}) > p(b, \{a, b\})$. Also, for any $c \in A \setminus \{a, b\}$, we know that $aTc \iff bTc$, implying $p(a, \{a, c\}) = p(b, \{b, c\})$ by IIA and a repeated application of NPR gives us $p(a, A) > p(b, A)$.

□

Now, we construct the scoring function S. Let x be the alternative with the highest Copeland score of the revealed tournament T i.e. $s(x, X) \geq s(y, X)$ for all $y \neq x$. Let $s(a, X) = k$ and $A_x = \{a_1, \ldots a_k\}$ be the set of all alternatives that x beats in a pairwise comparison. Define $A_{x_i} \coloneqq x \cup \{a_1, \ldots a_i\}$. Note that $i = s(x, A_{x_i})$ and define S as

$S(0) := 1$, $S(1) := \dfrac{p(x,\{x\})}{p(a^*,\{x\})}$ and $S(i) := \dfrac{p(x, A_{x_{i-1}})}{p(a^*, A_{x_{i-1}})}$ for all $i \in \{2, \dots, n_T\}$. Clearly, S is well defined. Now, we need to show that (i) For any $A \in \mathcal{P}(X)$ and $a \in A$,

$$p(a, A) = \frac{S(s(a, A))}{\sum_{b \in A^*} S(s(b, A))}$$

and (ii) S is monotone with respect to T.

Consider an arbitrary $A \in \mathcal{P}(X)$ and $a \in A$. Since $\sum_{a \in A^*} p(a, A) = 1$, to establish (i), it is sufficient to show that $\dfrac{p(a, A)}{p(b, A)} = \dfrac{S(s(a, A))}{S(s(b, A))}$ for all $b \in A^*$. Consider any $b \in A^*$ and observe that

$$\frac{p(a, A)}{p(b, A)} = \frac{p(a, A)}{p(a^*, A)} . \frac{p(a^*, A)}{p(b, A)} \tag{2}$$

$$= \frac{p(x, A_{x_{s(a,A)}})}{p(a^*, A_{x_{s(a,A)}})} . \frac{p(a^*, A_{x_{s(b,A)}})}{p(x, A_{x_{s(b,A)}})} \tag{3}$$

$$= \frac{S(s(a, A))}{S(s(b, A))} \tag{4}$$

where the (2) follows from T-IIA and (3) follows from the construction of the scoring function. To show (ii) consider any $i > j \in N_T$ such that $s(a, A) = i$ and $s(b, A) = j$ for some $a, b \in A$. Lemma 1 implies that $p(a, A) > p(b, A)$ and by part (i), we have $S(i) = S(s(a, A)) > S(s(b, A)) = S(j)$.

It is clear that the revealed tournament is unique. To show that the scoring function is unique upto multiplication by a positive constant, consider two scoring functions S and S' that "rationalize" p. Then, consider an arbitrary A and $a, b \in A$. We know that

$$\frac{p(a, A)}{p(b, A)} = \frac{S(s(a, A))}{S(s(b, A))}$$

and

$$\frac{p(a, A)}{p(b, A)} = \frac{S'(s(a, A))}{S'(s(b, A))}$$

which implies $S' = kS$ where $k > 0$.

□

3.3.1 *Stochastic Choice without Abstention*

To characterize the case where there is no abstention, we require a strengthening of T-IIA which relates the probability ratio of two alternatives in a set to a product of probability ratios. This condition is similar to *Cylical Independence* axiom introduced in Echenique and Saito (2019) and Ahumada and Ülkü (2018). However, their condition is a weakening of Luce IIA while our condition is strengthening of T-IIA.

Strong T-IIA: Consider any a, b, a', b' such that $s(a, A) = s(a', A')$ and $s(b, A) = s(b', B')$ for some $A, A', B' \in \mathcal{P}(X)$. If there exists a sequence of menus $A' = A_0, \ldots, A_n = B'$ and alternatives $a_1, \ldots, a_{2n}$ such that $s(a_{2i}, A_i) = s(a_{2i-1}, A_{i-1})$ for all $i \in \{1, \ldots n\}$, then

$$\frac{p(a, A)}{p(b, A)} = \frac{p(a', A')}{p(a_1, A')} \cdot \frac{p(a_2, A_1)}{p(a_3, A_1)} \cdots \cdot \frac{p(a_{2n}, B')}{p(b', B')}$$

To see why this is a strengthening of T-IIA, note that if B' is the same A', then statement of Strong T-IIA is the same as that of T-IIA. We show below that this strengthening of T-IIA, along with the other three axioms characterize CSCR without abstention.

Theorem 3. *An scr is a CSCR without abstention if and only if it satisfies Strong T-IIA, NPR, NBA and NBT. Further, the scoring function is unique upto multiplication by a positive constant and the tournament is unique.*

Proof. Define the revealed tournament T as follows: aTb if and only if $p(a, \{a, b\}) > p(b, \{a, b\})$. Completeness of T is ensured by NBT and asymmetry of T follows from construction . Since Strong T-IIA implies T-IIA, Lemma 1 holds and therefore, for any a, b and A, we have

$$s(a, A) > s(b, A) \iff p(a, A) > p(b, A)$$

Now, we will construct the scoring function S. As in the proof of Theorem 2, let x be the alternative with the highest Copeland score of the revealed tournament T i.e. $s(x, X) \geq s(y, X)$ for all $y \neq x$. Let $s(a, X) = k$ and $A_x = \{a_1, \ldots a_k\}$ be the set of all alternatives that x beats in a pairwise comparison. Define $A_{x_i} := x \cup \{a_1, \ldots a_i\}$. For any $j \in \{1, \ldots k\}$, relabel the set A_{x_j} as $\{x, a_1, \ldots, a_j\}$ in such that $s(x, A_{x_j}) > s(a_j, A_{x_j}) \geq s(a_{j-1}, A_{x_j}) \ldots \geq s(a_1, A_{x_j})$. That is, the indices are in an increasing order with respect to the revealed Copeland scores.

Now, for every $j \in \{1, \ldots, k\}$, define a corresponding sequence of sets $A^1_{x_j} \ldots A^{j_m}_{x_j}$ as follows $A^1_{x_j} := \{a_1\} \cup \{a_i \in A_{x_j} : a_1 T a_i\}$. That is, $A^1_{x_j}$ is the set of all alternatives in A_{x_j} that a_1 beats in a pairwise comparison. Note that a_1 is the Condorcet winner in the set $A^1_{x_j}$. Relabel the elements of $A^1_{x_j}$ as $\{a_1, a_2, \ldots, a_{k'}\}$ where $k' < k$ such that $s(a_1, A^1_{x_j}) > s(a_2, A^1_{x_j}) \geq \ldots \geq s(a'_k, A^1_{x_j})$. Now, define $A^2_{x_j} \subset A^1_{x_j}$ as $A^2_{x_j} := \{a_2\} \cup \{a_i \in A^1_{x_j} : a_2 T a_i\}$ and relabel $A^2_{x_j}$ as $\{a_2, a_3 \ldots a_{k''}\}$ where $k'' < k'$ such that $s(a_2, A^2_{x_j}) > s(a_3, A^1_{x_j}) \geq \ldots \geq s(a''_k, A^2_{x_j})$. So, the sequence can be recursively defined as $A^l_{x_j} = \{a_l\} \cup \{a_i \in A^{l-1}_{x_j} : a_l T a_i\}$. Let j_m be the smallest index such that there exists $b \in A^{j_m}_{x_j}$ where $s(b, A^{j_m}_{x_j}) = 0$. Since for every j, the sequence $A^1_{x_j} \ldots$, is a strictly decreasing sequence of sets, such an index exists. Let $b = a_{j_{m^*}}$. That is, $s(a_{j_{m^*}}, A^{j_m}_{x_j}) = 0$. Now, we define the scoring function S as follows: $S(0) := 1$ and for all $j \in N_T \setminus \{0\}$

$$S(j) := \frac{p(x, A_{x_j})}{p(a_1, A_{x_j})} . \frac{p(a_1, A^1_{x_j})}{p(a_2, A^1_{x_j})} \cdots \frac{p(a_{j_m}, A^{j_m}_{x_j})}{p(a_{j_{m^*}}, A^{j_m}_{x_j})}$$

Consider an arbitrary $A \in \mathcal{P}(X)$ and $a \in A$. Since $\sum_{b \in A} p(b, A) = 1$, to show that

$$p(a, A) = \frac{S(s(a, A))}{\sum_{b \in A} S(s(b, A))}$$

it is sufficient to show that $\dfrac{p(a, A)}{p(b, A)} = \dfrac{S(s(a, A))}{S(s(b, A))}$ for all $b \in A$. Consider any $b \in A$. Suppose $s(a, A) = i$ and $s(b, A) = i'$. Then $s(a, A) = s(x, A_{x_i})$ and $s(b, A) = s(x, A_{x_{i'}})$. Since $s(a_{i_{m^*}}, A^{i_m}_{x_i}) = s(a_{i'_{m^*}}, A^{i'_m}_{x_{i'}}) (= 0)$, by Strong T-IIA and the constructed S, we get

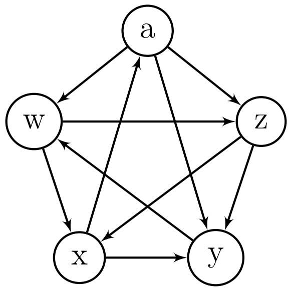
$$\begin{aligned} \frac{p(a, A)}{p(b, A)} &= \left[\frac{p(x, A_{x_i})}{p(a_1, A_{x_i})} . \frac{p(a_1, A^1_{x_i})}{p(a_2, A^1_{x_i})} \cdots \frac{p(a_{i_m}, A^{i_m}_{x_i})}{p(a_{i_{m^*}}, A^{i_m}_{x_i})}\right] \cdot \left[\frac{p(a_{i'_{m^*}}, A^{i'_m}_{x_{i'}})}{p(a_{i'_m}, A^{i'_m}_{x_{i'}})} \cdots \frac{p(a_1, A_{x'_i})}{p(x, A_{x'_i})}\right] \\ &= \frac{S(i)}{S(i')} \\ &= \frac{S(s(a, A))}{S(s(b, A))} \end{aligned}$$

Monotonicity of S follows from Lemma 1. Therefore, we have defined a tournament T and a scoring function S that is monotone with respect to T that "rationalizes" p. The arguments for uniqueness of the representation are the same as in Theorem 1. □

3.4 Top Cycle and Uncovered Set Rules

Two prominent solutions in the literature on tournaments are the top-cycle set and the uncovered set. For a tournament T on a set X, the top cycle set, denoted by $TC_T(X)$ is the smallest set (in terms of set inclusion) that has the property that everything inside the set beats everything outside that set. A refinement of the top-cycle set is the uncovered set denoted by $UC_T(X)$. An alternative is in its uncovered set if it pairwise beats every other alternative in at most two "steps". That is, for a given tournament T on X, an alternative x is in the uncovered set if it for any $y \in X \setminus \{x\}$, either xTy or there exists some $z \in X$ such that xTz and zTy. There are other ways to define the top cycle set and the uncovered set in terms of the transitive closure of the tournament and a "covering" relation respectively (see Laslier (1997) for details and other tournament solutions).

The Copeland set is a refinement of the uncovered set (see Brandt et al. (2016)). That is, the set of Copeland "winners" always belong to the uncovered set and hence the top-cycle set. We can see this for the tournament T on the set $X = \{a, x, y, z, w\}$ given in the figure below.



$$\mathrm{Cop}_T(X) = \{a\},\ UC_T(X) = \{a, x, z, w\} \text{ and } TC_T(X) = A$$

We further examine the relationship between these three tournament solutions by seeing which characterizing axioms of the other two are satisfied by Copeland rules. A revealed preference characterization of the top-cycle set was provided by Ehlers and Sprumont (2008) using three axioms: Weakened WARP (WWARP), Weak Contraction Consistency (WCC) and Binary Dominance Consistency (BDC). While BDC is used in our characterization of Copeland rules, Copeland rules fails to satisfy WWARP. However they satisfy WCC. WWARP, discussed in Section 2 is formally stated as follows.

WWARP: For any x, y and if $x \in C(A)$ and $y \in A \setminus C(A)$ for some A, then there does not exist $B \in \mathcal{P}(X)$ such that $y \in C(B)$ and $x \in B \setminus C(B)$.

A Copeland rule may not satisfy WWARP as it is clear from the tournament in the example above. We have xTa implying $\{x\} = C(\{a, x\})$. However, since $\{a\} = \mathrm{Cop}_T(A) = C(A)$, we have a violation of WWARP. The other condition required in the characterization of top-cycle rules is WCC that requires that if an alternative is chosen in a menu, then it must be chosen in *some* subset of cardinality one less. Formally,

WCC: For any $A \in \mathcal{P}(X)$, $C(A) \subseteq \bigcup_{x \in A} C(A \setminus \{x\})$.

Observation 1. *Copeland rules satisfy WCC.*

This immediately follows from BDC and IDA. If an alternative x is the Condorcet winner in a set A, then it is the Condorcet winner in every subset of A as well where it is present. Therefore, it is chosen by BDC. If x is not the Condorcet winner in A, then there exists another alternative that beats it in a pariwise comparison. Therefore x should be chosen upon the removal of that "dominating" alternative by IDA.

The characterization of uncovered set rules was given by Lombardi (2008) using four conditions: Weak Expansion (WE), Non-Discrimination (ND), Weakened Chernoff (WC) and BDC. WE is the standard expansion axiom or condition γ of Sen (1971), whereas ND requires that any alternatives x, y, z form a pairwise cycle, then all three must be chosen in the set $\{x, y, z\}$. WC is a weakening of condition α also known as Chernoff condition and requires that if an alternative is chosen in a menu of cardinality at least 4, then it must be chosen in the presence of every other alternative in that menu in some sub menu. These axioms are formally stated as follows.

Weak Expansion(WE): For any $A_1, \ldots, A_K \in \mathcal{P}(X)$, $K \in \mathbb{N}$, $\bigcap_{k=1}^{K} C(A_k) \subseteq C(\bigcup_{k=1}^{K} A_k)$.

Non-Discrimination(ND): For all distinct $x, y, z \in X$, if $C(xy) = \{x\}C(yz) = \{y\}$ and $C(xz) = \{z\}$, then $C(xyz) = \{xyz\}$.

Weakened Chernoff(WC): For all $A \in \mathcal{P}(X)$, such that $|A| > 3$, if $x \in C(A)$ and $y \in A \setminus \{x\}$, then $x \in \bigcup_{B \subsetneq A : x, y \in B} C(B)$.

A Copeland rule may not satisfy WE as can be seen from the tournament in the above example. Since x forms a cycle with y and w, it is chosen in $A = \{x, y, z\}$. It also forms a cycle with a and z and is chosen in $B = \{a, x, z\}$. However, since x is not a Copeland winner in $A \cup B = X$, it is not chosen in X. Copeland rules satisfy ND as an xyz cycle implies that $xTyTzTx$, therefore, the entire set is the Copeland set as well as the uncovered set.

Observation 2. *Copeland rules satisfy WC.*

To see why this is true, note that Copeland set is a subset of the uncovered set. Consider a set A and an alternative $x \in \mathrm{Cop}_T(A)$. Consider any arbitrary $y \in A \setminus \{x\}$. Since x is in the uncovered set, either xTy and we get $\{x\} = C(xy)$, or yTx in which case there exists $z \in A \setminus \{x, y\}$ such that $xTzTyTx$ forming a cycle implying $x \in C(xyz)$. Therefore, Copeland rules satisfies WC.

References

ABREU, D. AND A. RUBINSTEIN (1988): "The structure of Nash equilibrium in repeated games with finite automata," *Econometrica: Journal of the Econometric Society*, 1259–1281.

AGUIAR, V. H., M. J. BOCCARDI, AND M. DEAN (2016): "Satisficing and stochastic choice," *Journal of Economic Theory*, 166, 445–482.

AHUMADA, A. AND L. ÜLKÜ (2018): "Luce rule with limited consideration," *Mathematical Social Sciences*, 93, 52–56.

APESTEGUIA, J. AND M. A. BALLESTER (2013): "Choice by sequential procedures," *Games and Economic Behavior*, 77, 90–99.

ARROW, K. J. (1959): "Rational choice functions and orderings," *Economica*, 26, 121–127.

BRANDT, F., V. CONITZER, U. ENDRISS, J. LANG, AND A. D. PROCACCIA (2016): *Handbook of computational social choice*, Cambridge University Press.

CAMARA, M. K. (2021): "Computationally Tractable Choice," Tech. rep., Working Paper.

CAPLIN, A. AND M. DEAN (2011): "Search, choice, and revealed preference," *Theoretical Economics*, 6, 19–48.

CHEREPANOV, V., T. FEDDERSEN, AND A. SANDRONI (2013): "Rationalization," *Theoretical Economics*, 8, 775–800.

CHERNOFF, H. (1954): "Rational selection of decision functions," *Econometrica: journal of the Econometric Society*, 422–443.

CHEUNG, P. AND Y. MASATLIOGLU (2021): "Decision Making with Recommendation," .

COLOMER, J. M. (2013): "Ramon Llull: from 'Ars electionis' to social choice theory," *Social Choice and Welfare*, 40, 317–328.

COPELAND, A. H. (1951): "A reasonable social welfare function," Tech. rep., Mimeo, University of Michigan USA.

DIMITROV, D., S. MUKHERJEE, AND N. MUTO (2016): "'Divide-and-choose'in list-based decision problems," *Theory and Decision*, 81, 17–31.

DUTTA, R. AND S. HORAN (2015): "Inferring rationales from choice: Identification for rational shortlist methods," *American Economic Journal: Microeconomics*, 7, 179–201.

ECHENIQUE, F. AND K. SAITO (2019): "General luce model," *Economic Theory*, 68, 811–826.

EHLERS, L. AND Y. SPRUMONT (2008): "Weakened WARP and top-cycle choice rules," *Journal of Mathematical Economics*, 44, 87–94.

FUTIA, C. (1977): "The complexity of economic decision rules," *Journal of Mathemat-*

ical Economics, 4, 289–299.

Gerasimou, G. (2020): "The Decision-Conflict Logit," *arXiv preprint arXiv:2008.04229.*

Guney, B. (2014): "A theory of iterative choice in lists," *Journal of Mathematical Economics*, 53, 26–32.

Henriet, D. (1985): "The Copeland choice function an axiomatic characterization," *Social Choice and Welfare*, 2, 49–63.

Horan, S. (2010): "Sequential search and choice from lists," *Unpublished paper.[716].*

——— (2016): "A simple model of two-stage choice," *Journal of Economic Theory*, 162, 372–406.

Houthakker, H. S. (1950): "Revealed preference and the utility function," *Economica*, 17, 159–174.

Huber, J., J. W. Payne, and C. Puto (1982): "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of consumer research*, 9, 90–98.

Huber, V. L., M. A. Neale, and G. B. Northcraft (1987): "Decision bias and personnel selection strategies," *Organizational Behavior and Human Decision Processes*, 40, 136–147.

Koopmans, T. C. (1960): "Stationary ordinal utility and impatience," *Econometrica: Journal of the Econometric Society*, 287–309.

Kovach, M. and L. Ülkü (2020): "Satisficing with a variable threshold," *Journal of Mathematical Economics*, 87, 67–76.

Kramer, G. H. (1967): "An impossibility result concerning the theory of decision-making," .

Laslier, J.-F. (1997): *Tournament solutions and majority voting*, vol. 7, Springer.

Lleras, J. S., Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay (2017): "When more is less: Limited consideration," *Journal of Economic Theory*, 170, 70–85.

Lombardi, M. (2008): "Uncovered set choice rules," *Social Choice and Welfare*, 31, 271–279.

——— (2009a): "Minimal covering set solutions," *Social Choice and Welfare*, 32, 687–695.

——— (2009b): "Reason-based choice correspondences," *Mathematical Social Sciences*, 57, 58–66.

Luce, R. D. (1956): "Semiorders and a theory of utility discrimination," *Econometrica, Journal of the Econometric Society*, 178–191.

Manzini, P. and M. Mariotti (2007): "Sequentially rationalizable choice," *American Economic Review*, 97, 1824–1839.

——— (2010): "Revealed preferences and boundedly rational choice procedures: an experiment," *University of St. Andrews and IZA Working Paper*, 141–167.

——— (2012): "Categorize then choose: Boundedly rational choice and welfare," *Journal of the European Economic Association*, 10, 1141–1165.

——— (2014): "Stochastic choice and consideration sets," *Econometrica*, 82, 1153–1176.

Manzini, P., M. Mariotti, and L. Ulku (2019): "Sequential Approval: A Model of "Likes", Paper Downloads and Other Forms of Click Behaviour." .

Masatlioglu, Y. and D. Nakajima (2007): "A theory of choice by elimination," *Unpublished paper.[778, 785, 788].*

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012): "Revealed attention," *American Economic Review*, 102, 2183–2205.

Moon, J. W. (2015): *Topics on tournaments in graph theory*, Courier Dover Publications.

Munkres, J. (1974): *Topology.*

Neyman, A. (1985): "Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma," *Economics letters*, 19, 227–229.

Papi, M. (2012): "Satisficing choice procedures," *Journal of Economic Behavior & Organization*, 84, 451–462.

Richter, M. K. (1966): "Revealed preference theory," *Econometrica: Journal of the Econometric Society*, 635–645.

Richter, M. K. and K.-C. Wong (1999): "Computable preference and utility," *Journal of Mathematical Economics*, 32, 339–354.

Rubinstein, A. (1980): "Ranking the participants in a tournament," *SIAM Journal on Applied Mathematics*, 38, 108–111.

——— (1986): "Finite automata play the repeated prisoner's dilemma," *Journal of economic theory*, 39, 83–96.

——— (2012): *Lecture notes in microeconomic theory*, Princeton University Press.

Rubinstein, A. and Y. Salant (2006): "A model of choice from lists," *Theoretical Economics*, 1, 3–17.

Salant, Y. (2003): "Limited computational resources favor rationality," *NAJ Economics*, 6.

——— (2011): "Procedural analysis of choice rules with applications to bounded rationality," *American Economic Review*, 101, 724–48.

Salant, Y. and A. Rubinstein (2008): "(A, f): choice with frames," *The Review of Economic Studies*, 75, 1287–1296.

Sanver, M. R. and Ö. Selçuk (2010): "A characterization of the Copeland solution," *Economics Letters*, 107, 354–355.

Sen, A. K. (1971): "Choice functions and revealed preference," *The Review of Economic Studies*, 38, 307–317.

Simon, H. A. (1955): "A behavioral model of rational choice," *The quarterly journal of economics*, 69, 99–118.

Simonson, I. (1989): "Choice based on reasons: The case of attraction and compromise effects," *Journal of consumer research*, 16, 158–174.

Sokolova, T. and A. Krishna (2016): "Take it or leave it: How choosing versus rejecting alternatives affects information processing," *Journal of Consumer Research*, 43, 614–635.

Teppan, E. C. and A. Felfernig (2009): "Minimization of product utility estimation errors in recommender result set evaluations," in *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, IEEE, vol. 1, 20–27.

Tserenjigmid, G. (2019): "Choosing with the worst in mind: A reference-dependent model," *Journal of Economic Behavior & Organization*, 157, 631–652.

——— (2021): "The order-dependent Luce model," *Management Science*, 67, 6915–6933.

Tversky, A. (1972): "Choice by elimination," *Journal of mathematical psychology*, 9, 341–367.

Tyson, C. J. (2013): "Behavioral implications of shortlisting procedures," *Social Choice and Welfare*, 41, 941–963.

——— (2015): "Satisficing behavior with a secondary criterion," *Social Choice and Welfare*, 44, 639–661.

Wang, K. (2022): "Approval with frames," *Journal of Mathematical Economics*, 103, 102788.

Wright, P. and F. Barbour (1977): *Phased decision strategies: Sequels to an initial screening*, Graduate School of Business, Stanford University.

Zermelo, E. (1929): "Die berechnung der turnier-ergebnisse als ein maximumproblem der wahrscheinlichkeitsrechnung," *Mathematische Zeitschrift*, 29, 436–460.