On observable restrictions of choice models

Yuta Inoue

Acknowledgements

First, I would like to express my gratitude to my advisors, Dr. Ryo Nagata and Prof. Hisatoshi Tanaka, for the continuous support of my Ph.D studies and research. I would also like to thank the members of my thesis committee, Prof. Kazumi Shimizu and Prof. Yuki Hosoya, for their insightful comments and encouragement.

My sincere thanks also goes to my co-author Dr. Koji Shirai, with whom I conducted the study that constitutes Chapter 2 of this thesis. I appreciate that he kindly agreed to include our study in this thesis. He also gave comments on my research in Chapters 3 and 4, which substantially improved the quality of my research.

Last but not least, I would like to thank my family, my parents, my brothers, and my wife and son, for supporting me spiritually throughout the composition of this thesis and my life in general.

Contents

Chapter 1

Introduction

This chapter provides a brief introduction of the economic ideas dealt in this thesis, namely *revealed preference theory* and *limited consideration*. In particular, revealed preference theory is a method used in all of the subsequent chapters, while limited consideration is a subfield of bounded rationality that we deal with in Chapters 2 and 3.

1.1 Revealed preference theory

In standard economic theory, it is a common method to construct a model based on some hypotheses, and analyze the model under some suitable assumptions. For such economic analyses to be plausible, we must be sure that the hypotheses and assumptions that the analysis is based on are plausible. Put otherwise, in order to conduct an economic analysis on some observed economic phenomenon, we must clarify that the underlying hypotheses/assumptions of the economic model do not contradict with the observed phenomenon. Revealed preference theory is a field that aims to provide tests to check whether specific economic hypotheses/assumptions do not contradict observed data. The structures of an economic model that are used in such tests are called *testable implications* or *observable restrictions* of the model.

Revealed preference theory has been extensively studied subsequent to Samuelson (1938) and Houthakker (1950), where they showed a set of necessary and sufficient conditions that a demand function must satisfy if a decision maker (henceforth DM) is a utility maximizer. This result was further studied by Afriat (1967) and Varian (1982), where they gave necessary and sufficient conditions that a finite set of price-consumption data must obey if the consumer is maximizing her utility. While Afriat (1967) and Varian (1982) focused on consumption theory on the Eucledean consumption space, Richter (1966) considers observable restrictions of optimizing behavior on abstract consumption spaces. Since this thesis assumes a finite set of alternatives to be the grand set of alternatives, the revealed preference analyses in this thesis will follow that of Richter (1966).

Here we show the observable restrictions of preference maximizing behavior on a finite consumption space. Let X be a finite set of alternatives, and $\succ$ be a connected, asymmetric, and transitive preference of a decision maker (henceforth, DM), which we refer to as a *strict preference*. Typically, a choice function is observed: letting $\mathcal{D} \subseteq 2^X \backslash \varnothing$ be an arbitrary collection of nonempty feasible sets, for every feasible set $A \in \mathcal{D}$, $f(A) \in A$ is the chosen alternative from A.[1][2] If a DM obeys the rational choice model, then for every subset $A \in \mathcal{D}$, she maximizes her strict preference on A. Therefore, we can define a *(strict) direct revealed preference relation* $\succ^R$, such that $x'' \succ^R x'$ if there exists a feasible set $A \in \mathcal{D}$ with $f(A) = x''$, $x' \in A$, and $x' \neq x''$. In

[1]In Chapter 2, we assume that a finite number of pairs of choice and feasible set are observed. This is just a different way to express a choice function defined as above.

[2]Throughout this thesis, we abuse notation and abbreviate the braces, and write $2^X \backslash \varnothing$ instead of $2^X \backslash \{\varnothing\}$. Similar abbreviation of braces will be used whenever there is no fear of confusion.

words, alternative x'' is revealed to be preferred to x' if there is a feasible set where x'' is chosen when x' is available. If the DM is rational, then her direct revealed preference $\succ^R$ must be a subset of her true preference $\succ$, and thus $\succ^R$ must be acyclic. In fact, acyclicity of the direct revealed preference, namely the *Strong Axiom of Revealed Preference (SARP)* is not only a necessary condition, but also a sufficient condition for an observed choice function to be consistent with the rational choice model.

1.2 Limited consideration

While the rational choice model can easily be tested using SARP, it is widely known that observed choice data commonly violates SARP. In order to deal with such seemingly irrational behavior, various theories of bounded rationality have been proposed. Amongst others, in Chapters 2 and 3 of this thesis, we adopt the behavioral assumption of limited consideration. Under limited consideration, consciously or unconsciously, a DM excludes some feasible alternatives from consideration. In particular, it is assumed that a DM has a strict preference $\succ$, but when facing a feasible set A, she takes into consideration only a subset of what is available: $\Gamma(A) \subseteq A$. Then, the DM chooses the $\succ$-best alternative within $\Gamma(A)$ rather than the feasible set A itself. This subset $\Gamma(A)$ is called the *consideration set* of feasible set A, and the mapping $\Gamma : 2^X \backslash \varnothing \rightarrow 2^X \backslash \varnothing$ is referred to as a *consideration mapping.* Various models of limited consideration have been proposed, and the models differ depending on restrictions casted on the consideration mapping. For example, the *limited attention* model in Masatlioglu, Nakajima, and Ozbay (2012) assumes that Γ is an *attention filter*, which requires that removal of an ignored alternative does not change the consideration set; and the *overwhelming choice* model in Lleras, Masatlioglu, Nakajima, and Ozbay (2017) assumes that Γ is a *competition filter*, which requires that an alternative considered in a larger set must be considered in a smaller set.

In fact, other important theories of non-standard decision making are closely related to the limited consideration models introduced above. Under the *rational shortlisting method* by Manzini and Mariotti (2007), the *transitive rational shortlisting method* by Au and Kawai (2011), the *categorize-then-choose* model by Manzini and Mariotti (2012), and the *rationalization* model by Cherepanov, Feddersen, and Sandroni (2013), given a feasible set, a DM makes a shortlist prior to making a final decision. The profile of shortlists created under the rational shortlisting method, the categorize-then-choose model, and the rationalization model is in fact a competition filter, and that under the transitive rational shortlisting method is an attention filter and competition filter. In this sense, these models can be regarded as special cases of limited consideration.

1.3 Organization of the thesis

In this thesis, we focus on expositions of choice models based on non-exhaustive choice data, which means that the observed choice data is a choice function whose domain is merely a subset of all conceivable feasible sets, or a finite set of pairs of choice a feasible sets.

Chapter 2 develops a revealed preference analysis for limited consideration models. A revealed preference test is given for the decision model obeying two well-established hypotheses on a decision maker's consideration: the attention filter property and competition filter property. We also provide a test for a two-step decision model, namely the (transitive) rational shortlist method. We conduct simulations to compare the relative strength of observable restrictions across leading models, where we find drastic differences in Bronars' indices of models. This chapter is based on Inoue and Shirai (Forthcoming).[3]

Chapter 3 is based on Inoue (2020a), which puts forward a behavioral framework where a decision maker makes choices at multiple time periods, while she may not be aware of all available alternatives at all times. This framework inherits the main idea of the theory of limited consideration, and adds to it an assumption that a decision maker's consideration grows over time. In particular, it is required that she takes into consideration any alternative that she chose in the past. We refer to this property as growing consideration. Revealed preference tests, as well as conditions under which we can robustly infer the decision maker's preference, consideration, and non-consideration are given. Following a revealed preference analysis of a baseline framework of decision making under growing consideration, we deal with special cases where the decision maker's consideration is a competition filter/attention filter.

In Chapter 4, we develop revealed preference analysis of an individual choice model where a DM is a weak preference maximizer, under the assumption that a choice function, rather than a choice correspondence, is observed. In particular, we provide a revealed preference test for such model, and then provide conditions under which we can surely say whether some alternative is indifferent/weakly preferred/strictly preferred to another, solely from the information on the choice function. Furthermore, interpreting a choice correspondence as sets of potential candidates of alternatives that could be chosen from each feasible set, we analyze which alternatives must be, or cannot be a member of the choice correspondence: sharp lower and upper bounds of this

[3]The author's contribution to this paper is as follows. Firstly, the main idea and the setup of the model was proposed by the author. Also, the simulation in Section 2.3 was conducted by the author. Finally, the proofs of the theorems were done in tandem with the co-author, Dr. Koji Shirai.

underlying choice correspondence are given. As an assumption on observability of data, we assume that the choice function is defined on a non-exhaustive domain, so our results are applicable to data analysis even when only a limited data set is available. Chapter 4 is based on Inoue (2020b).

Following concluding remarks in Chapter 5, basic mathematical concepts regarding binary relations and extension theorems are given in the Mathematical Appendix in Chapter 6.

Chapter 2

On the observable restrictions of limited consideration models: theory and application

In the literature of decision theory, limited consideration models have been motivated by evidence that a decision maker (DM) does not consider all (physically) available alternatives. Amongst others, Masatlioglu, Nakajima, and Ozbay (2012) and Lleras, Masatlioglu, Nakajima, and Ozbay (2017) formulate choice models in which a DM's consideration set has specific restrictions called, respectively, *attention filter (AF)* and *competition filter (CF)* properties. In the former, the removal of unrecognized alternatives does not change the consideration set, while the latter requires that any alternative considered in a larger set must be considered in a smaller set as long as it is available. It is known that a CF is derived from various two-step decision models, including those by Manzini and Mariotti (2007), Au and Kawai (2011), Manzini and Mariotti (2012), and Cherepanov, Feddersen, and Sandroni (2013), while Au and Kawai (2011) also implies an AF.

These models have clean and interpretable axiomatizations in terms of an *exhaustive* choice function, which specifies choices over *all* logically possible subsets of alternatives. In this chapter, we are interested in the problem of testing these models when we observe choices from *some* but not all subsets of alternatives. In the case of AF, this has been done by De Clippel and Rozen (2021), whereas the case of CF is essentially done by Dean, Kibris, and Masatlioglu (2017).[1] The primary objective of this chapter is to provide tests for important cases that are not covered by these two papers.[2]

We formulate a test for checking models where the consideration set has both AF and CF structures, which we call a *competitive attention filter (CAF)*. In fact, there are many plausible scenarios that satisfy both AF and CF properties.[3] It should be noted that even if a data set is consistent with AF and CF separately, there is no guarantee that it is consistent with CAF. We also provide a revealed preference test for the *rational shortlisting (RS)* method proposed by Manzini and Mariotti (2007), which is a special case of the CF model. In this model, a DM makes a shortlist of alternatives that are undominated in terms of some acyclic binary relation, and then maximizes her preference within this shortlist. We also test the important special case of the RS model formulated by Au and Kawai (2011). In Au and Kawai's variation, the binary relation deciding shortlists is supposed to be

[1]Dean, Kibris, and Masatlioglu (2017) provides a revealed preference test for a status quo bias model where a DM's consideration set is a CF that depends on the status quo. It can be easily adjusted to a general test for a CF from limited data. Note also that the first version of De Clippel and Rozen (2021) dates back to 2012, and, to the best of the author's knowledge, it is the first paper to provide a revealed preference analysis of limited consideration models when the data are incomplete.

[2]In this thesis, we concentrate on deterministic choice models. Stochastic choice models with limited consideration have been also studied in the literature; see for example, Manzini and Mariotti (2014) and Brady and Rehbeck (2016). A recent paper by Allen and Rehbeck (Forthcoming) deals with a revealed preference analysis on a stochastic choice model that can comprehend a certain type of stochastic consideration models as a special case.

[3]See Section 2.1 for examples.

asymmetric and transitive, so we call this decision procedure the *transitive rational shortlisting (TRS)* method. Note that the TRS model also satisfies AF, and hence it is a special case of the CAF model.

The basic idea of our tests can be summarized as follows. To explain a data set by a certain limited consideration model, we need to elicit the information concerning a DM's consideration sets obeying relevant structures as well as that of her preference. Unlike the standard rational choice model, a data set collected from a limited consideration model can contain revealed preference cycles (i.e., the Strong Axiom of Revealed Preference is violated). As long as the DM's preference is a strict preference, each cycle must contain at least one revealed preference relation that has the opposite direction to the true preference.[4] Such a reversal contains some information concerning the DM's consideration sets, since not choosing a preferred alternative implies that it is not considered. Then, each restriction on the DM's consideration sets, such as CAF or (T)RS, would in turn impose some restriction on how cycles are broken. Then the precise restriction for each model works as a revealed preference test.

Our revealed preference tests have a similar mathematical structure to the test for AF in De Clippel and Rozen (2021), which the authors call *acyclic satisfiability*. As shown in their paper, this class of problems is in general computationally challenging, and our tests also share that computational issue. Nevertheless, the features of our tests allow us to employ a computing

[4]While we concentrate on models where a DM has a strict preference, one may also consider a more general case. Indeed, in Cherepanov, Feddersen, and Sandroni (2013), Manzini and Mariotti (2012), and even Manzini and Mariotti (2007), a DM's preference is just assumed to be asymmetric (Cherepanov, Feddersen, and Sandroni (2013) also contain the case of a strict preference). Testing models with a non-standard preference from limited data is largely open, but a recent working paper by De Clippel and Rozen (2018) deals with such a case (the paper is available at `https://www.brown.edu/Departments/Economics/Faculty/Kareen_Rozen/catrat_v67.pdf`).

method called *backtracking*, which is an efficient search method in dealing with combinatorial problems.[5]

Having developed tools for testing different limited consideration models, we apply them to investigate the empirical restrictiveness of these models. Specifically, we test AF, CF, CAF, and (T)RS models. It is obvious that limited consideration models are relatively permissive compared to the rational choice model, and there are several subclass/superclass relations within limited consideration models (e.g., CAF is obviously stronger than both of AF and CF). Following Bronars (1987), we generate random choices and apply our tests to see the fraction of data that are consistent with each model. In our simulation, we stick to an environment with 20 sets of alternatives, each of which contains 2 – 8 alternatives out of 10 alternatives. Choices are generated by using the uniform distribution over each set of alternatives following Bronars (1987). We find that the strengths of observable restrictions are strikingly different across models. In fact, AF is very hard to reject with the average pass rate of random data exceeding 99%, and CF is also permissive with the average pass rate exceeding 60%. However, CAF is far more restrictive with the average pass rate being less than 4%. The rational shortlist models also have strong testing power: the average pass rate of RS is less than 3% and that of TRS is less than 0.1%.

Organization of this chapter: In Section 2.1, we briefly review limited consideration models dealt with in this chapter. The theoretical heart of this chapter lies in Section 2.2: we firstly provide a basic idea of our approach in testing limited consideration models in Section 2.2.1, and then provide revealed preference tests for CAF and (T)RS respectively in Sections 2.2.2

[5] Classical textbook examples where backtracking is used are the eight queens puzzle, crossword puzzles, and sudoku.

and 2.2.3. In Section 2.2.4, we briefly refer to the test for AF and CF that are shown in the existing papers. In Section 2.3, we apply our tests to simulation data. The substantial parts of the proofs of main theorems are given in Appendix I, and the computation procedure using backtracking is explained in Appendix II.

2.1 Limited consideration models

Consider an individual decision problem where X is a finite set of alternatives, and $\succ$ is a connected, asymmetric, and transitive preference of a DM, which we refer to as a *strict preference*. If a DM obeys the *rational choice* model, then for every subset $A \subseteq X$, she maximizes her strict preference on A. On the other hand, in limited consideration models, either consciously or unconsciously, a DM makes a shortlist of alternatives, and then she maximizes her preference on that shortlist. That is, there exists a *consideration mapping* $\Gamma : 2^X \backslash \varnothing \to 2^X \backslash \varnothing$ such that $\Gamma(A) \subseteq A$ for every $A \in 2^X \backslash \varnothing$, and a DM maximizes her strict preference on $\Gamma(A)$, rather than A itself. Given a consideration mapping Γ, $\Gamma(A)$ is referred to as a *consideration set* on A. We call a pair of strict preference and consideration mapping $(\succ, \Gamma)$ as a *limited consideration* model.

While various types of Γ are considered in the literature, we focus on consideration mappings obeying the following two properties that are investigated by Masatlioglu, Nakajima, and Ozbay (2012) and Lleras, Masatlioglu, Nakajima, and Ozbay (2017). We say that Γ is an *attention filter (AF)*, if for every $A \subseteq X$ and $x \in A$,

$$x \notin \Gamma(A) \Longrightarrow \Gamma(A \backslash x) = \Gamma(A). \tag{2.1}$$

In words, Γ is an AF, if the consideration set is not affected when unrecognized alternatives are removed. On the other hand, we say that Γ is a *competition filter (CF)*, if for every $A' \subseteq A''$ and $x \in A'$,

$$x \notin \Gamma(A') \Longrightarrow x \notin \Gamma(A''). \tag{2.2}$$

That is, when Γ is a CF, if an alternative is not recognized in a smaller set, then it cannot be recognized in a larger set. When Γ satisfies both (2.1) and (2.2), we say that Γ is a *competitive attention filter (CAF)*.[6] When Γ is an AF, a limited consideration model $(\succ, \Gamma)$ is referred to as the *AF-model*. Similarly, we say that $(\succ, \Gamma)$ is the *CF-model* (*CAF-model*), when Γ is a CF (CAF). Many real-world examples can be described by a CAF. For example, a DM pays attention to: (a) n-most advertised commodities; (b) all commodities of a specific brand, and if there are none available, then all commodities of another specific brand; or (c) n-top candidates in each field in job markets, then all of them derive CAF.

As an important special case of the CF-model, we also deal with the *rational shortlist method* proposed by Manzini and Mariotti (2007). There, for each set $A \subseteq X$, a consideration set is defined such that

$$\Gamma(A) = \{x \in A : \nexists x' \in A \text{ such that } x' \succ' x\}, \tag{2.3}$$

for some acyclic relation $\succ'$, to which we refer as a *consideration relation*. That is, a DM only picks up undominated alternatives with respect to her consideration relation. We refer to a consideration mapping Γ defined as

[6] We realize that these notions are also considered in many other papers, with being referred to as different names. We basically follow the terminologies in Masatlioglu, Nakajima, and Ozbay (2012) and Lleras, Masatlioglu, Nakajima, and Ozbay (2017). A survey by Moulin (1985) contains many theoretical results concerning these restrictions.

(2.3) as a *rational shortlist (RS)*. Au and Kawai (2011) deals with the special case where $\succ'$ is asymmetric and transitive, and we call Γ as a *transitive rational shortlist (TRS)* in that case. It is straightforward to check that an RS is a special case of a CF and that a TRS is a special case of a CAF. For example, amongst three examples of CAFs in the preceding paragraph, only (b) can be described as a TRS. Moreover, one can also confirm that an RS is a TRS if and only if it obeys AF-property.[7]

Note also that, in the original setting of Manzini and Mariotti (2007), a DM's preference $\succ$ is just asymmetric, while in this thesis we consider the case where $\succ$ is a strict preference as in Au and Kawai (2011).

2.2 Revealed preference tests

In this section, we provide a tool for testing limited consideration models based on a *data set* in the form of $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$; where $\mathcal{T} = \{1, 2, ..., T\}$ is the set of indices of observations, $A^t \subseteq X$ is the set of (physically) available alternatives at observation t, and $a^t \in A^t$ is the chosen alternative at $t \in \mathcal{T}$. Thus, for each observation point $t \in \mathcal{T}$, we observe a DM's choice and a set A^t, while a consideration set $\Gamma(A^t)$ is not observable. Throughout this chapter, $A^s \neq A^t$ is assumed for every $s \neq t$. It should be stressed that following De Clippel and Rozen (2021), we allow the case where an economist can observe a DM's choice behavior only on *some* subsets of X, rather than observing an entire choice function.

Given a data set as above, we would like to find a pair of strict preference and consideration mapping $(\succ, \Gamma)$ such that for each $t \in \mathcal{T}$, the observed choice a^t is the $\succ$-best alternative within $\Gamma(A^t)$ and that Γ obeys a specific

[7]This connection can be easily shown by using the results in the earlier papers by Sen (1971) and Schwartz (1976).

restriction introduced in the preceding section (AF, CF, CAF and (T)RS).

Definition 1. A data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ is *rationalizable* by a limited consideration model $(\succ, \Gamma)$, if for every $t \in \mathcal{T}$, $a^t \in \Gamma(A^t)$ and $a^t \succ x$ for every $x \in \Gamma(A^t) \backslash a^t$. In particular, if $\mathcal{O}$ is rationalizable by the $\mathcal{M}$-model ($\mathcal{M}$ = AF, CF, CAF, (T)RS), then we say that $\mathcal{O}$ is $\mathcal{M}$-rationalizable.[8]

Note that when we say that $\mathcal{O}$ is $\mathcal{M}$-rationalizable, Γ is required to obey $\mathcal{M}$ on the entire domain $2^X \backslash \varnothing$, rather than just on the observed subsets. Amongst five types of restrictions raised in the preceding section, tests for the AF-model and the CF-model are known in the literature: De Clippel and Rozen (2021) formulates a test for $\mathcal{M}$ = AF, and a test for $\mathcal{M}$ = CF can be easily derived from Theorem 5 in Dean, Kibris, and Masatlioglu (2017) (see Section 2.2.4 for more details). Our contribution is to formulate revealed preference tests for the CAF-model and the (T)RS-model. Note that as we later show in an example, even if a data set is AF-rationalizable and CF-rationalizable, it may not be CAF-rationalizable, and hence, a test for the CAF-model must be independently derived. In Section 2.2.1, we provide the general idea for testing, and then proceed to a test for each specific model in Sections 2.2.2 and 2.2.3.

2.2.1 A common starting point of tests

Given a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$, define the *(strict) direct revealed preference* relation $\succ^R$ such that $x'' \succ^R x'$, if $x'' = a^t$ for some $t \in \mathcal{T}$, $x'' \neq x'$, and $x' \in A^t$. It is well known that a data set is consistent with the rational choice model, if and only if $\succ^R$ is acyclic, or the *Strong Axiom of Revealed*

[8]It is clear that any data set is rationalizable by some $(\succ, \Gamma)$ without any restriction on the shape of Γ. Indeed, we could just let $\Gamma(A^t) = \{a^t\}$ for every $t \in \mathcal{T}$, which does not work in general once some $\mathcal{M}$ (= AF, CF, CAF, (T)RS) is imposed.

Preference (SARP) is satisfied. Put otherwise, if a data set $\mathcal{O}$ obeys SARP, then we can find a strict preference $\succ$ such that $(\succ, \Gamma)$ rationalizes $\mathcal{O}$ with Γ being the identity mapping. Since the identity mapping obeys all conditions concerning Γ referred to in this thesis, testing limited consideration models becomes substantial when $\mathcal{O}$ contains revealed preference cycles.

A revealed preference cycle is formally defined as a set of pairs $\mathcal{C} = \{(x^k, x^{k+1})\}_{k=1}^{K}$ with $x^k \succ^R x^{k+1}$ for every $k = 1, 2, \ldots, K$, and $x^1 = x^{K+1}$. We refer to each (x^k, x^{k+1}) as an *arc* of a cycle. In what follows, we only consider cycles that do not contain other cycles inside. For example, while one may construct a cycle like $x \succ^R y \succ^R z \succ^R y \succ^R x$, we don't count it as a cycle. In the subsequent argument, such a cycle is automatically treated if we deal with two "independent" cylces $x \succ^R y \succ^R x$ and $y \succ^R z \succ^R y$. In addition, if a cycle is constructed by rotating elements of another cycle (e.g., $y \succ^R z \succ^R x \succ^R y$ is constructed by rotating $x \succ^R y \succ^R z \succ^R x$), we regard it as the same cycle with the original cycle.

Now, consider a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ that has revealed preference cycles, and suppose that it is collected from a DM obeying some limited consideration model $(\succ, \Gamma)$. Since a DM has a strict preference $\succ$, for each cycle, there exists at least one arc (x^k, x^{k+1}) for which $x^{k+1} \succ x^k$. When there are Q revealed preference cycles, from each q-th cycle, pick up one of those arcs c_q to make a profile of arcs $\mathbf{D} = (c_1, c_2, ..., c_Q) \in \times_{q=1}^{Q} \mathcal{C}_q$.[9] Note that, since each c_i is an ordered pair of components in X, $\mathbf{D}$ can be also regarded as a set of ordered pairs, or a binary relation on X. We interpret $\mathbf{D}$ as such whenever it is convenient. If a profile of arcs $\mathbf{D}$ is determined as above, it is effectively a part of a DM's strict preference $\succ$ in that $(x', x'') \in \mathbf{D}$ implies $x'' \succ x'$. Hence, $\mathbf{D}$ must be acyclic if it is regarded as a binary

[9] As stated in the preceding paragraph, we only consider "independent" cycles. Since X is finite, there are at most finitely many such cycles.

relation. In addition, the set of arcs $\mathbf{D}$ also has some connection with a DM's consideration mapping Γ as follows. For each $t \in \mathcal{T}$, define

$$B^t_{\mathbf{D}} = \{x \in A^t : (a^t, x) \in \mathbf{D}\}. \tag{2.4}$$

Then, every $x \in B^t_{\mathbf{D}}$ is available at A^t and preferred to a^t, which implies that $x \notin \Gamma(A^t)$. Put otherwise, $\Gamma(A^t) \subseteq A^t \backslash B^t_{\mathbf{D}}$ holds for every $t \in \mathcal{T}$. We summarize the above observation as a fact for future references.

Fact 1. *Suppose that a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ is rationalizable by some limited consideration model and has Q revealed preference cycles. Then, there exists an acyclic selection of arcs from cycles $\mathbf{D} = (c_1, c_2, ..., c_Q) \in \times_{q=1}^{Q} \mathcal{C}_q$ such that $\Gamma(A^t) \subseteq A^t \backslash B^t_{\mathbf{D}}$ for every $t \in \mathcal{T}$, where $B^t_{\mathbf{D}} = \{x \in A^t : (a^t, x) \in \mathbf{D}\}$.*

Note that the above fact is derived without using any specific property $\mathcal{M}$, and hence it is shared by any limited consideration model. On the other hand, in testing a model, we do not have a priori information concerning $(\succ, \Gamma)$. Hence, given a data set with revealed preference cycles, we have to make a "guess" of a profile of arcs $\mathbf{D}$ that is acyclic. By Fact 1, once $\mathbf{D}$ is specified, that in turn specifies for each $t \in \mathcal{T}$, a set of robustly unconsidered alternatives $B^t_{\mathbf{D}}$. Thus, at least $\mathbf{D}$ has to be chosen such that $a^t \notin B^t_{\mathbf{D}}$. In addition, depending on the structure of a family of observed subsets, the fact that $\Gamma(A^t) \subseteq A^t \backslash B^t_{\mathbf{D}}$ for every $t \in \mathcal{T}$ may conflict with some structure $\mathcal{M}$ on Γ. Put otherwise, whenever $\mathcal{M}$ is specified, we have to find $\mathbf{D}$ that does not cause any contradiction with it. This is a common structure of revealed preference tests in this chapter: we need to check the existence of $\mathbf{D}$ that can be reconciled with observed choices and a restriction on consideration in issue. Then, a crucial step, which we elaborate in the next subsection, is finding a condition that $\mathbf{D}$ must satisfy in relation to each specific $\mathcal{M}$. Before proceeding to it, the example below may help to see the argument up to this point.

Example 1. *Let $X = \{x_1, x_2, x_3\}$ and consider a data set of three observations as follows, where for each $t \in \mathcal{T}$, the chosen alternative is underlined:*

$$A^1 = \{\underline{x_1}, x_2\}, \quad A^2 = \{x_1, \underline{x_2}, x_3\}, \quad A^3 = \{x_1, \underline{x_3}\}.$$

This data set contains cycles $\mathcal{C}_1 : x_1 \succ^R x_2 \succ^R x_1$ and $\mathcal{C}_2 : x_1 \succ^R x_2 \succ^R x_3 \succ^R x_1$. As a possible selection of arcs, pick up $x_1 \succ^R x_2$ from both cycles; that is, $\mathbf{D} = \{(x_1, x_2)\}$, which is obviously acyclic as a binary relation. Then, $B^1_{\mathbf{D}} = \{x_2\}$ and $B^t_{\mathbf{D}} = \varnothing$ for $t = 2, 3$. By Fact 1, this implies that $\Gamma(A^1) \subseteq \{x_1\}$, and hence $\Gamma(A^1) = \{x_1\}$. However, this selection $\mathbf{D}$ cannot be a "correct" guess of a DM's preference when one would like to rationalize her choices by the CF-model: $x_2 \notin \Gamma(A^1)$ and $A^1 \subseteq A^2$ implies $x_2 \notin \Gamma(A^2)$ when x_2 is selected from A^2. This shows that some restriction is imposed by the model in selecting arcs from the cycles.

2.2.2 Testing CAF-model

Here, we establish a revealed preference test for the CAF-model. Suppose that a data set $\mathcal{O}$ with Q revealed preference cycles is CAF-rationalizable. To derive a characterization, by using the fact that Γ is a CAF, we strengthen Fact 1 step by step. In what follows, another expression of the definition of an AF in (2.1) is useful: for every $A, A' \subseteq X$, $\Gamma(A) \subseteq A' \subseteq A \Longrightarrow \Gamma(A') = \Gamma(A)$.

Fact 2. *Suppose that a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ is CAF-rationalizable. Then, there exists an acyclic selection of arcs from cycles $\mathbf{D} = (c_1, c_2, ..., c_Q) \in$*

$\times_{q=1}^{Q}\mathcal{C}_q$ *such that for every* $s, t \in \mathcal{T}$*,*

$$(A^s \backslash B_{\mathbf{D}}^s) \subseteq A^t \Longrightarrow \Gamma(A^t) \subseteq A^t \backslash B_{\mathbf{D}}^s. \tag{2.5}$$

Proof. Take any $\mathbf{D}$ for which $\Gamma(A^t) \subseteq A^t \backslash B_{\mathbf{D}}^t$ for every $t \in \mathcal{T}$, which exists by Fact 1. Note that when $(A^s \backslash B_{\mathbf{D}}^s) \subseteq A^t$ holds, we have $\Gamma(A^s) \subseteq A^s \backslash B_{\mathbf{D}}^s \subseteq (A^s \cap A^t) \subseteq A^s$. Then, since Γ is an AF, $\Gamma(A^s) = \Gamma(A^s \cap A^t)$ must hold, and hence, $x \in B_{\mathbf{D}}^s \Longrightarrow x \notin \Gamma(A^s \cap A^t)$. Since Γ is also a CF and $(A^s \cap A^t) \subseteq A^t$, we have $x \notin \Gamma(A^t)$. This shows that $\Gamma(A^t) \subseteq A^t \backslash B_{\mathbf{D}}^s$. □

The above can be further extended as follows.

Fact 3. *Suppose that a data set* $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ *is CAF-rationalizable. Then, there exists an acyclic selection of arcs from cycles* $\mathbf{D} = (c_1, c_2, ..., c_Q) \in \times_{q=1}^{Q}\mathcal{C}_q$ *such that for every* $r, s, t \in \mathcal{T}$*,*

$$[(A^r \cup A^s) \backslash (B_{\mathbf{D}}^r \cup B_{\mathbf{D}}^s)] \subseteq A^t \Longrightarrow \Gamma(A^t) \subseteq A^t \backslash (B_{\mathbf{D}}^r \cup B_{\mathbf{D}}^s). \tag{2.6}$$

Proof. Again, consider $\mathbf{D}$ obeying the property referred to in Fact 1. Then, both $\Gamma(A^r) \subseteq A^r \backslash B_{\mathbf{D}}^r$ and $\Gamma(A^s) \subseteq A^s \backslash B_{\mathbf{D}}^s$ hold. Since Γ is a CF, it holds that $x \in B_{\mathbf{D}}^r \Longrightarrow x \notin \Gamma(A^r \cup A^s)$ and $x \in B_{\mathbf{D}}^s \Longrightarrow x \notin \Gamma(A^r \cup A^s)$, which implies $\Gamma(A^r \cup A^s) \subseteq [(A^r \cup A^s) \backslash (B_{\mathbf{D}}^r \cup B_{\mathbf{D}}^s)]$. Since $[(A^r \cup A^s) \backslash (B_{\mathbf{D}}^r \cup B_{\mathbf{D}}^s)] \subseteq A^t$ is assumed, we have $[(A^r \cup A^s) \backslash (B_{\mathbf{D}}^r \cup B_{\mathbf{D}}^s)] \subseteq [A^t \cap (A^r \cup A^s)] \subseteq (A^r \cup A^s)$. By the fact that Γ is an AF, it holds that $\Gamma(A^t \cap (A^r \cup A^s)) = \Gamma(A^r \cup A^s) \subseteq [(A^r \cup A^s) \backslash (B_{\mathbf{D}}^r \cup B_{\mathbf{D}}^s)]$. Finally, combining $[A^t \cap (A^r \cup A^s)] \subseteq A^t$ and Γ being a CF, we have $x \in (B_{\mathbf{D}}^r \cup B_{\mathbf{D}}^s) \Longrightarrow x \notin \Gamma(A^t)$ as desired. □

By an inductive argument, ultimately we can extend (2.6) for any subset of indices $\tau \subseteq \mathcal{T}$ such that $\left(\bigcup_{r \in \tau} A^r \backslash \bigcup_{r \in \tau} B_{\mathbf{D}}^r\right) \subseteq A^t$. That is:

Fact 4. *Suppose that a data set* $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ *is CAF-rationalizable. Then, there exists an acyclic selection of arcs from cycles* $\mathbf{D} = (c_1, c_2, ..., c_Q) \in \times_{q=1}^{Q}\mathcal{C}_q$ *such that for every* $\tau \subseteq \mathcal{T}$*,*

$$\left(\bigcup_{r\in\tau} A^r \Big\backslash \bigcup_{r\in\tau} B^r_{\mathbf{D}}\right) \subseteq A^t \Longrightarrow \Gamma(A^t) \subseteq A^t \backslash \bigcup_{r\in\tau} B^r_{\mathbf{D}}. \tag{2.7}$$

The condition in Fact 4 depends on Γ, which is not observed by an econometrician, and hence we cannot directly check the existence of $\mathbf{D}$ obeying (2.7) from a data set. Nevertheless, we can convert it to a condition in terms of choices, which are observed in a data set. Indeed, Fact 4 implies that, when $\mathcal{O}$ is CAF-rationalizable, there must exist an acyclic section of arcs from cycles $\mathbf{D}$ such that $\left(\bigcup_{r\in\tau} A^r \Big\backslash \bigcup_{r\in\tau} B^r_{\mathbf{D}}\right) \subseteq A^t \Longrightarrow a^t \notin \bigcup_{r\in\tau} B^r_{\mathbf{D}}$. The right hand side follows, since $a^t \in \Gamma(A^t)$ must hold for every $t \in \mathcal{T}$.

CAF-condition: Suppose that a data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ contains Q revealed preference cycles. A selection of arcs from cycles $\mathbf{D} = (c_1, c_2, \ldots, c_Q) \in \times_{q=1}^{Q}\mathcal{C}_q$ obeys the CAF-condition, if for every $t \in \mathcal{T}$ and any set of indices $\tau \subseteq \mathcal{T}$,

$$\left(\bigcup_{r\in\tau} A^r \Big\backslash \bigcup_{r\in\tau} B^r_{\mathbf{D}}\right) \subseteq A^t \Longrightarrow a^t \notin \bigcup_{r\in\tau} B^r_{\mathbf{D}}. \tag{2.8}$$

The existence of a selection of arcs obeying the above can be checked once a data set is given, and it is necessary for a data set to be CAF-rationalizable. More substantially, if we can find such an acyclic selection of arcs, then a data set is CAF-rationalizable. That is, CAF-rationalizability is tested by checking the existence of an acyclic selection of arcs from cycles obeying

CAF-condition.

Theorem 1. *A data set* $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ *is CAF-rationalizable, if and only if there exists an acyclic selection of arcs from cycles that obeys the CAF-condition.*

Remark. As seen from the statement, testing CAF-rationalizability is equivalent to testing the existence of acyclic binary relations obeying a certain restriction. A test of this type is referred to as a test of *acyclic satisfiability* by De Clippel and Rozen (2021), whose test for AF-rationalizability also has that structure. In general, checking acyclic satisfiability is computationally hard and requires special computational techniques. In our application, we employ *backtracking*, which is explained in detail in Appendix II.

Two examples are in order. We firstly look at the example in the preceding subsection. Concerning that example, we have raised a selection of arcs that is inconsistent with the CF-model (and hence, inconsistent with the CAF-model), but the data set *is* CAF-rationalizable, since there is another selection of arcs that obeys CAF-condition. On the other hand, the second example is not CAF-rationalizable, even though it is both AF-rationalizable and CF-rationalizable separetely.

Example 1 (continued). *Reconsider the data set as follows:*

$$A^1 = \{\underline{x_1}, x_2\}, \quad A^2 = \{x_1, \underline{x_2}, x_3\}, \quad A^3 = \{x_1, \underline{x_3}\},$$

which contains two revealed preference cycles $\mathcal{C}_1 : x_1 \succ^R x_2 \succ^R x_1$ *and* $\mathcal{C}_2 : x_1 \succ^R x_2 \succ^R x_3 \succ^R x_1$. *Choose* $x_2 \succ^R x_1$ *from* $\mathcal{C}_1$ *and* $x_3 \succ^R x_1$ *from* $\mathcal{C}_2$ *so that* $\mathbf{D} = \{(x_2, x_1), (x_3, x_1)\}$, *which is obviously acyclic as a* *binary relation. For each* $t = 1, 2, 3$*, the corresponding* $B^t_{\mathbf{D}}$ *is calculated as* $B^1_{\mathbf{D}} = \varnothing$*,* $B^2_{\mathbf{D}} = \{x_1\}$*, and* $B^3_{\mathbf{D}} = \{x_1\}$*. For* $t = 1$*, only* $\tau = \{1\}$ *satisfies the LHS of (2.8), and* $a^1 = x_1 \notin B^1_{\mathbf{D}}$ *is satisfied. For* $t = 2$*, any* $\tau \subseteq \mathcal{T}$ *satisfies the LHS of (2.8). As shown in the proof of Theorem 1 in Appendix I, it suffices to confirm (2.8) for the maximal subset obeying the LHS of it, which is* $\tau = \{1, 2, 3\}$ *in this case. It is clear that* $a^2 = x_2 \notin \bigcup_{r=1}^{3} B^r_{\mathbf{D}}$*. For* $t = 3$*, only* $\tau = \{3\}$ *satisfies the LHS of (2.8), and* $a^3 = x_3 \notin B^3_{\mathbf{D}}$ *holds. Thus,* $\mathbf{D} = \{(x_2, x_1), (x_3, x_1)\}$ *obeys CAF-condition, and Theorem 1 ensures that this data set is CAF-rationalizable. For example, letting* $(\succ, \Gamma)$ *be such that* $x_1 \succ x_2 \succ x_3$*,* $\Gamma(\{x_1, x_2\}) = \{x_1, x_2\}$*,* $\Gamma(\{x_2, x_3\}) = \{x_2, x_3\}$*,* $\Gamma(\{x_1, x_3\}) = \{x_3\}$*, and* $\Gamma(\{x_1, x_2, x_3\}) = \{x_2, x_3\}$*, then it is easy to check that it is a CAF-model that rationalizes the data.*

Example 2. *Let* $X = \{x_1, x_2, x_3\}$ *and consider the following observations:*

$$A^1 = \{\underline{x_1}, x_2\}, \quad A^2 = \{x_1, \underline{x_2}, x_3\}, \quad A^3 = \{x_2, \underline{x_3}\},$$

where chosen alternatives are underlined. This data set is not CAF-rationalizable. This data set contains two revealed preference cycles, namely $\mathcal{C}_1 : x_1 \succ^R x_2 \succ^R x_1$*;* $\mathcal{C}_2 : x_2 \succ^R x_3 \succ^R x_2$*. We firstly claim that any selection of arcs containing* (x_1, x_2) *cannot satisfy CAF-condition. Let* $\mathbf{D}$ *be a selection of arcs from cycles containing* (x_1, x_2)*. Then,* $x_2 \in B^1_{\mathbf{D}}$*, and hence* $A^1 \backslash B^1_{\mathbf{D}} \subseteq A^2$ *and* $a^2 = x_2 \in B^1_{\mathbf{D}}$*. This is a violation of (2.8), and hence such a selection* $\mathbf{D}$ *cannot satisfy CAF-condition. Therefore, from* $\mathcal{C}_1$*, the arc* (x_2, x_1) *must be selected. Then, consider a selection of arcs* $\mathbf{D} = ((x_2, x_1), (x_2, x_3))$*. This derives* $B^2_{\mathbf{D}} = \{x_1, x_3\}$*, and we have* $A^2 \backslash B^2_{\mathbf{D}} = \{x_2\} \subseteq A^1$ *and* $a^1 = x_1 \in B^2_{\mathbf{D}}$*, which is a violation of (2.8). Thus, from* $\mathcal{C}_2$*, the arc* (x_3, x_2) *must be selected, and* $\mathbf{D} = ((x_2, x_1), (x_3, x_2))$ *is the only one remaining possibility. This* *selection derives* $B^3_{\mathbf{D}} = \{x_2\}$ *and* $A^3 \backslash B^3_{\mathbf{D}} = \{x_3\} \subseteq A^2$. *However, since* $a^2 = x_2 \in B^3_{\mathbf{D}}$, *this* $\mathbf{D}$ *also violates CAF-condition.*

On the other hand, this data set can be rationalized respectively by the AF-model and the CF-model (not both simultaneously). Let $(\succ, \Gamma)$ *be such that* $x_3 \succ x_1 \succ x_2$ *and* $\Gamma(\{x_1, x_2\}) = \{x_1, x_2\}$, $\Gamma(\{x_2, x_3\}) = \{x_2, x_3\}$, $\Gamma(\{x_1, x_3\}) = \{x_1, x_3\}$, *and* $\Gamma(\{x_1, x_2, x_3\}) = \{x_2\}$. *Then,* Γ *is a CF, while it is not an AF (the removal of* x_1 *from* $\{x_1, x_2, x_3\}$ *changes consideration). It is straightforward that the data set here is generated from this* $(\succ, \Gamma)$. *Similarly, a data set is rationalizable by the AF-model* $(\succ, \Gamma)$ *as follows:* $x_2 \succ x_1 \succ x_3$ *and* $\Gamma(\{x_1, x_2\}) = \{x_1\}$, $\Gamma(\{x_2, x_3\}) = \{x_3\}$, $\Gamma(\{x_1, x_3\}) = \{x_3\}$ *and* $\Gamma(\{x_1, x_2, x_3\}) = \{x_1, x_2, x_3\}$. *Indeed, at* $\{x_1, x_2, x_3\}$, *every alternative is considered, and hence the requirement of an AF is trivially satisfied. It is also easy to see that this* Γ *is not a CF.*

2.2.3 Testing (T)RS-model

We turn to the case of the (T)RS-model. Based on the nature of the model, we can extend Fact 1 as follows. Suppose that $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ contains Q revealed preference cycles and is (T)RS-rationalizable. Then, a DM has two binary relations $\succ'$ and $\succ$, where the former is acyclic (asymmetric and transitive) consideration relation and the latter is a strict preference, and the consideration mapping Γ is defined as (2.3). Then, Fact 1 implies that there exists an acyclic selection of arcs from cycles $\mathbf{D} = (c_1, c_2, \ldots, c_Q)$ such that the corresponding $\{B^t_{\mathbf{D}}\}_{t \in \mathcal{T}}$ satisfies $\Gamma(A^t) \subseteq A^t \backslash B^t_{\mathbf{D}}$ for every $t \in \mathcal{T}$. Since a DM obeys the (T)RS-model, for every $x' \in B^t_{\mathbf{D}}$, there exists some $x'' \in A^t \backslash x'$ such that $x'' \succ' x'$. This in turn implies that x' is not considered as long as x'' is available, and hence, $x' \succ^R x''$ is impossible.

Given the discussion above, we can define a binary relation $\rhd$ on X such that $x'' \rhd x'$ if $x' \in B^t_{\mathbf{D}}$ for some $t \in \mathcal{T}$, $x'' \in A^t \backslash x'$, and $x' \nsucc^R x''$. Since we start from a data set consistent with the (T)RS-model, for every $x' \in B^t_{\mathbf{D}}$, there exists at least one $x'' \in A^t \backslash x'$ with $x'' \rhd x'$ for which $x'' \succ' x'$ actually holds. Loosely speaking, $\rhd$ can be seen as a broad guess of the consideration relation $\succ'$. In addition, the acyclicity of $\succ'$ requires that we can always find a selection $\rhd' \subseteq \rhd$ that is acyclic, and for every $t \in \mathcal{T}$ and $x' \in B^t_{\mathbf{D}}$, there exists some $x'' \in A^t \backslash x'$ with $x'' \rhd' x'$. Furthermore, if the consideration relation $\succ'$ is assumed to be transitive, a selection $\rhd'$ has to be chosen so that

$$\text{for every } x' \in B^t_{\mathbf{D}} \text{ and } z^1, ..., z^k,\ x'' \rhd' z^1 \rhd' \cdots \rhd' z^k \rhd' x' \Longrightarrow x' \nsucc^R x''. \tag{2.9}$$

Now, $\rhd'$ is a "correct" guess of the consideration relation, and if transitivity is imposed, the above implies that $x'' \succ' x'$. Hence, if $x' \succ^R x''$ were to hold, then it leads to a contradiction that x' is deleted from a consideration set from which it is actually chosen. In fact, this observation is summarized in the conditions below, and plays a key role to characterize a data set that is rationalizable by the (T)RS-model.

(T)RS-condition: Suppose that a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ contains Q revealed preference cycles. An acyclic selection of arcs from cycles $\mathbf{D} = (c_1, c_2, \ldots, c_Q) \in \times_{q=1}^{Q} \mathcal{C}_q$ obeys the RS-condition, if for the corresponding $\{B^t_{\mathbf{D}}\}_{t \in \mathcal{T}}$, there exists an acyclic selection $\rhd'$ of $\rhd$, where for every $t \in \mathcal{T}$,

$$\text{for every } x' \in B^t_{\mathbf{D}},\ \text{there exists } x'' \in A^t \text{ with } x'' \rhd' x'. \tag{2.10}$$

When $\rhd'$ can be chosen so that (2.9) is also satisfied, we say that $\mathbf{D}$ obeys the TRS-condition.

Theorem 2. *A data set* $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ *is (T)RS-rationalizable, if and only if there exists an acyclic selection of arcs from cycles obeying the (T)RS-condition.*

Remark 1. As seen from the statement, testing the (T)RS-model also involves tests of acyclic satisfiability. De Clippel and Rozen (2021) showed that testing the RS-model is NP-hard, even though they did not provide a revealed preference test for that model. Backtracking is also useful for testing the (T)RS-model, as we discuss in Appendix II.

Remark 2. As referred to in Section 2.1, the original verion of the RS-model in Manzini and Mariotti (2007) requires that a DM's preference is just asymmetric, rather than a strict preference. Our theorem above (and simulation in Section 2.3) does not cover this general case.

The following examples show how we can test the (T)RS-model using Theorem 2. The first example is consistent with the TRS-model, while the second is not even rationalized by the RS-model. In addition, the latter shows that the RS-model has a strictly stronger observable restriction than the CF-model. Similarly, one can construct an example that is CAF-rationalizable, but not TRS-rationalizable.

Example 1 (continued). *Reconsider the data set as follows:*

$$A^1 = \{\underline{x_1}, x_2\}, \quad A^2 = \{x_1, \underline{x_2}, x_3\}, \quad A^3 = \{x_1, \underline{x_3}\},$$

which contains two revealed preference cycles $\mathcal{C}_1 : x_1 \succ^R x_2 \succ^R x_1$ *and* $\mathcal{C}_2 : x_1 \succ^R x_2 \succ^R x_3 \succ^R x_1$. *We claim that this data set is TRS-rationalizable, by using* $\mathbf{D} = \{(x_2, x_1), (x_3, x_1)\}$. *As shown in the preceding subsection, this derives* $B^1_{\mathbf{D}} = \varnothing$, $B^2_{\mathbf{D}} = \{x_1\}$, *and* $B^3_{\mathbf{D}} = \{x_1\}$. *In this case, only one relation* *is determined by* $\rhd$*, which is* $x_3 \rhd x_1$*. This flows from the fact that* $x_1 \in B^2_{\mathbf{D}}$*,* $x_3 \in A^2$ *and* $x_1 \not\succ^R x_3$ *(*$x_2 \not\succ x_3$*, since* $x_1 \succ^R x_2$*). The same relation can be also derived from* $t = 3$*. Obviously, we can adopt* $\rhd$ *itself as an asymmetric and transitive selection* $\rhd'$ *obeying (2.10), and hence, this data set is TRS-rationalizable.*

Example 2 (continued). *Reconsider the data set consisting of the following observations:*

$$A^1 = \{\underline{x_1}, x_2\}, \quad A^2 = \{x_1, \underline{x_2}, x_3\}, \quad A^3 = \{x_2, \underline{x_3}\},$$

where chosen alternatives are underlined. While this data set is CF-rationalizable as shown in the previous subsection, it is not RS-rationalizable. We firstly claim that any selection of arcs containing (x_1, x_2) *or* (x_3, x_2) *cannot satisfy RS-condition. Whenever* (x_1, x_2) *is contained in* $\mathbf{D}$*, we have* $x_2 \in B^1_{\mathbf{D}}$*. Meanwhile, since there is no* $x \in A^1$ *such that* $x_2 \not\succ^R x$*, it is impossible to define* $\rhd$ *so that* x_2 *is dominated by an alternative in* A^1*, which violates (2.10). A parallel logic shows that* (x_3, x_2) *cannot be an arc selected from* $\mathcal{C}_2$*. Hence* $\mathbf{D} = ((x_2, x_1), (x_2, x_3))$ *is the only remaining possibility. However, in this case,* $B^2_{\mathbf{D}} = \{x_1, x_3\}$*, and thus* $x_1 \rhd x_3$ *and* $x_3 \rhd x_1$ *hold, and there does not exist an acyclic selection of* $\rhd$ *that obeys (2.10).*

2.2.4 Tests for an AF-model and a CF-model

Here, we briefly refer to tests for the AF-model and the CF-model, which are known in the literature: De Clippel and Rozen (2021) establish the former, while the latter can be derived from Theorem 5 in Dean, Kibris, and Masatlioglu (2017) (we prove the latter in Appendix I).

Theorem A. *A data set* $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ *is AF-rationalizable, if and only* *if there exists an acyclic binary relation* $\succ^*$ *such that for every* $s, t \in \mathcal{T}$ *with* $a^s, a^t \in A^s \cap A^t$ *and* $a^s \neq a^t$*,*

$$\exists x \in A^s \backslash A^t \text{ such that } a^s \succ^* x \text{ or } \exists x \in A^t \backslash A^s \text{ such that } a^t \succ^* x. \tag{2.11}$$

Theorem B. *A data set* $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ *is CF-rationalizable, if and only if the following binary relation* $\succ^{CF}$ *is acyclic: for* $x'' \neq x'$*,*

$$x'' \succ^{CF} x', \text{ if for some } s, t \in \mathcal{T}, \ x'' = a^s, x' = a^t \text{ and } \{x'', x'\} \subseteq A^s \subseteq A^t. \tag{2.12}$$

As seen from the above, tests for the AF-model and the CF-model have quite different structures. The test for the CF-model stated in Theorem B has a “conventional” form of revealed preference tests in that it only requires testing acyclicity of a binary relation determined from a data set. On the other hand, Theorem A has a similar structure with tests in Theorems 1 and 2, in that one has to check the existence of an acyclic binary relation with specific properties. Note also that these models can be also tested by using our approach of finding out a profile of arcs from cycles with certain properties.[10]

2.3 Simulation

Given the theorems in the preceding section, we can now test the AF, CF, CAF and (T)RS-models on a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$. In this section, we apply the tests for these models to randomly generated data sets to com-

[10]See the earlier working paper version of this chapter for those results (available at `https://ideas.repec.org/p/kgu/wpaper/176-2.html`).

pare the relative strength of observable restrictions across models.[11] Since the power of each test depends on the structure of a family of the sets of alternatives, we randomly generate sets of alternatives as well as the choices over them. This can be regarded as a version of Bronars' test in the context of limited consideration models, and one can measure the strength of the observable restrictions of each model by using its pass rate.[12] We apply our tests to random choices obtained by the uniform distribution to approximate the proportion of choices that are model-consistent among all logically possible choices. If this value is very close to 1, then the model in question is very hard to refute, or its observable restrictions are weak. The detail of our data generating procedure is as follows.

Random generation of sets of alternatives. In this simulation, we use data sets with $|X| = 10$, $|\mathcal{T}| = 20$, $\min |A^t| = 2$, and $\max |A^t| = 8$, and randomly generate 100 profiles of sets of alternatives $\mathbb{A}_n := \{A_n^t\}_{t \in \mathcal{T}}$ for $n = 1, \ldots, 100$. For each n, we firstly specify the size of each A_n^t following a uniform distribution over the set of natural numbers $\{2, \ldots, 8\}$, and then choose $|A_n^t|$ elements from X following a uniform distribution over X. We also require that $A_n^s \neq A_n^t$ for $s \neq t$.

Choices by the uniform distribution. For each profile of sets $\mathbb{A}_n = \{A_n^t\}_{t \in \mathcal{T}}$, a profile of choices $\{a_{i,n}^t\}_{t \in \mathcal{T}}$ is generated for $i = 1, \ldots, 100$ using the uniform distribution. Consequently, we obtain a random choice data set $\mathcal{O}_{i,n} = \{(a_{i,n}^t, A_n^t)\}_{t \in \mathcal{T}}$ for $i = 1, \ldots, 100$ and $n = 1, \ldots, 100$, to which we apply revealed preference tests for AF, CF, CAF, (T)RS, and SARP.

[11] In the earlier working paper version of this chapter mentioned in footnote 10, we also carry out an experiment concerning these models.

[12] Bronars (1987) deals with the revealed preference test of classical consumer theory. There, the fail rate of the General Axiom of Revealed Preference (GARP) on randomly generated consumption bundles on randomly generated budgets is calculated.

Tests	SARP	AF	CF	CAF	RS	TRS
Pass rates	0.0000	0.9927	0.6298	0.0396	0.0259	0.0006

Table 2.1: Average pass rates.

Table 2.1 shows the average pass rates of tests, where the average is taken over 100 different profiles of the sets of alternatives.[13] It shows that the AF-model is extremely permissive, letting more than 99% of the random data sets pass the test, and the CF-model is also quite permissive. On the other hand, we can say that observable restrictions of the CAF and the (T)RS-models are reasonably strong. What is striking is that, while more than 60% of all data sets pass both tests for the AF-model and the CF-model, the pass rate of the CAF-model is significantly lower (lower than 4%). Similarly, the difference between the RS-model and the TRS-model is also huge, which is again due to AF-property (recall that the TRS-model is the RS-model with AF-property). Thus, although the hypothesis of Γ being an AF is very hard to reject by itself, combining it with some other restrictions could strengthen observable restrictions drastically.

2.4 Appendix I: Proofs

Proof of Theorem 1

We construct a pair of consideration mapping and strict preference that rationalizes $\mathcal{O}$ based on an acyclic selection of arcs from cycles $\mathbf{D}$ (and the corresponding $\{B^t_{\mathbf{D}}\}_{t\in\mathcal{T}}$) obeying CAF-condition. To define Γ, we need the

[13]While the pass rate of SARP is zero, there exist choice patterns consistent with it in theory — none of them turns up in our samples.

following set of indices for every $A \subseteq X$:

$$\tau(A) = \max\left\{\tau \subseteq \mathcal{T} : \bigcup_{r\in\tau} A^r \Big\backslash \bigcup_{r\in\tau} B^r_{\mathbf{D}} \subseteq A\right\}. \tag{2.13}$$

Then, by using $\tau(A)$, define Γ such that

$$\Gamma(A) = A \Big\backslash \bigcup_{r\in\tau(A)} B^r_{\mathbf{D}}. \tag{2.14}$$

Obviously, in order for the above definition to be well-defined, $\tau(A)$ must be uniquely determined for every $A \subseteq X$, which is actually the case. To see this, suppose to the contrary: there exist $\tau_1(A) \neq \tau_2(A)$ that obey (2.13). Then, we have $\left(\bigcup_{r\in\tau_1(A)} A^r \backslash \bigcup_{r\in\tau_1(A)} B^r_{\mathbf{D}}\right) \subseteq A$ and $\left(\bigcup_{r\in\tau_2(A)} A^r \backslash \bigcup_{r\in\tau_2(A)} B^r_{\mathbf{D}}\right) \subseteq A$, which implies that

$$\left[\bigcup_{r\in\tau_1(A)\cup\tau_2(A)} A^r \Big\backslash \left(\bigcup_{r\in\tau_1(A)} B^r_{\mathbf{D}} \cup \bigcup_{r\in\tau_2(A)} B^r_{\mathbf{D}}\right)\right] \subseteq A.$$

Obviously, this can be rewritten as

$$\left(\bigcup_{r\in\tau_1(A)\cup\tau_2(A)} A^r \Big\backslash \bigcup_{r\in\tau_1(A)\cup\tau_2(A)} B^r_{\mathbf{D}}\right) \subseteq A.$$

By defining $\tau(A) = \tau_1(A) \cup \tau_2(A)$, we have $\tau(A) \supsetneq \tau_i(A)$ for $i = 1, 2$, which contradicts the maximality of $\tau_1(A)$ and $\tau_2(A)$.

Given that Γ defined as (2.14) is well-defined, we move on to show that Γ is both AF and CF. Consider any $A', A'' \subseteq X$ with $A' \subseteq A''$, and $x \in A'$ such that $x \notin \Gamma(A')$. This means that $x \in \bigcup_{r\in\tau(A')} B^r_{\mathbf{D}}$. Since $\tau(\cdot)$ is clearly monotonic, it follows that $\tau(A') \subseteq \tau(A'')$, and hence, $x \in \bigcup_{r\in\tau(A'')} B^r_{\mathbf{D}}$. This assures that $x \notin \Gamma(A'')$, which shows that Γ is CF. To see AF, take any

$A \subseteq X$ and any $x \in A$ with $x \notin \Gamma(A)$. This means that $x \in \bigcup_{r\in\tau(A)} B^r_{\mathbf{D}}$, which in turn implies

$$\left(\bigcup_{r\in\tau(A)} A^r \Big\backslash \bigcup_{r\in\tau(A)} B^r_{\mathbf{D}}\right) \subseteq A\backslash x. \tag{2.15}$$

The maximality and uniqueness of $\tau(\cdot)$, combined with (2.15), imply $\tau(A) \subseteq \tau(A\backslash x)$. On the other hand, the monotonicity of $\tau(\cdot)$ implies $\tau(A\backslash x) \subseteq \tau(A)$. Hence we have $\tau(A) = \tau(A\backslash x)$. Then, $\Gamma(A\backslash x) = (A\backslash x)\backslash\bigcup_{r\in\tau(A\backslash x)} B^r_{\mathbf{D}} = A\backslash\bigcup_{r\in\tau(A)} B^r_{\mathbf{D}} = \Gamma(A)$, which shows that Γ is AF.

Let $\succ^*$ be a binary relation such that $x'' \succ^* x'$, if $x'' = a^t$, $x' \in \Gamma(A^t)$, and $x'' \neq x'$. We show that $\succ^*$ is acyclic, and thus extendable to a strict preference. By way of contradiction, suppose that there exists a cycle $x^1 \succ^* x^2 \succ^* \cdots \succ^* x^L \succ^* x^1$, which clearly implies $x^1 \succ^R x^2 \succ^R \cdots \succ^R x^L \succ^R x^1$. Then, there exists an arc $(x^\ell, x^{\ell+1})$ contained in $\mathbf{D}$. Since $x^\ell = a^t$ and $x^{\ell+1} \in A^t$ hold for some $t \in \mathcal{T}$, this means that $x^{\ell+1} \in B^t_{\mathbf{D}}$ for such an observation t. It is easy to check from the definition of Γ that $t \in \tau(A^t)$, and hence, $x^{\ell+1} \notin \Gamma(A^t) \subseteq A^t\backslash B^t_{\mathbf{D}}$. However, then, it holds that $x^\ell \not\succ^* x^{\ell+1}$, which is a contradiction.

Finally, let us show that $a^t \in \Gamma(A^t)$ for every $t \in \mathcal{T}$, which follows immediately from CAF-condition. Indeed, for every $t \in \mathcal{T}$, we have

$$\left(\bigcup_{r\in\tau(A^t)} A^r \Big\backslash \bigcup_{r\in\tau(A^t)} B^r_{\mathbf{D}}\right) \subseteq A^t,$$

and then, CAF-condition requires $a^t \notin \bigcup_{r\in\tau(A^t)} B^r_{\mathbf{D}}$, which in turn ensures $a^t \in \Gamma(A)$ for every $t \in \mathcal{T}$. Since $\succ^*$ is acyclic, it is extendable to a strict preference $\succ$ on X using Szpilrajn's theorem. Then this $\succ$ and Γ defined as (2.14) combined together is a CAF-model that rationalizes $\mathcal{O}$. $\square$

Proof of Theorem 2

The proofs for RS-model and TRS-model are almost identical, so we provide the proofs of them jointly. Since the necessity parts of them have been already discussed, we prove the sufficiency parts of them based on an acyclic selection of arcs from cycles obeying (T)RS-condition. Using an acyclic selection $\rhd'$ of $\rhd$, define Γ as

$$\Gamma(A) = \{x \in A : \nexists x' \in A \text{ such that } x' \rhd' x\}. \tag{2.16}$$

Note that the selection $\rhd'$ is acyclic, so we use it as a consideration relation for RS-model. If we can find $\rhd'$ so that it obeys (2.9) in addition to (2.10), then we use the transitive closure of it, say, $\rhd''$ as a consideration relation and define Γ by using it instead of $\rhd'$. This $\rhd''$ works as a consideration relation for TRS-model. Note further that $\Gamma(A^t) \subseteq A^t \backslash B^t_{\mathbf{D}}$ holds, by the definition of $\rhd'$ (or $\rhd''$) and the construction of Γ. The remaining substantial parts of the proof are to show that $a^t \in \Gamma(A^t)$ for every $t \in \mathcal{T}$, and the binary relation $\succ^*$ defined as $x'' \succ^* x'$ if $x'' = a^t, x' \in \Gamma(A^t)$, and $x'' \neq x'$ is acyclic.

To prove that $\succ^*$ is acyclic, suppose to the contrary, i.e., there is a cycle: $x^1 \succ^* x^2 \succ^* \cdots \succ^* x^L \succ^* x^1$. Since we have $\succ^* \subseteq \succ^R$, this cycle implies $x^1 \succ^R x^2 \succ^R \cdots \succ^R x^L \succ^R x^1$. Then, it must be the case that there exists an arc $(x^\ell, x^{\ell+1})$ contained in $\mathbf{D}$, and we have $x^{\ell+1} \in B^t_{\mathbf{D}}$ for every $t \in \mathcal{T}$ with $x^\ell = a^t$ and $x^{\ell+1} \in A^t$. By (T)RS-condition, there exists some $x \in A^t$ such that $x \rhd' (\rhd'') x^{\ell+1}$, which in turn implies $x^{\ell+1} \notin \Gamma(A^t)$. Then it is impossible to have $x^\ell = a^t \succ^* x^{\ell+1}$, and we conclude that $\succ^*$ is acyclic.

To see that $a^t \in \Gamma(A^t)$ for every $t \in \mathcal{T}$, by way of contradiction, suppose that for some $t \in \mathcal{T}$, $a^t \notin \Gamma(A^t)$. This means that there exists $x \in A^t \backslash a^t$ such that $x \rhd' a^t$, which in turn implies $x \rhd a^t$. However, this is not possible, since

$x \rhd a^t$ requires $a^t \not\succ^R x$, while we have $a^t \succ^R x$. When $\mathbf{D}$ obeys TRS-condition and Γ is defined as the set of maximal elements with respect to $\rhd''$, $a^t \notin \Gamma(A^t)$ implies the existence of some $x \in A^t \backslash a^t$ such that $x \rhd'' a^t$. However, this is also impossible, since $x \rhd'' a^t$ implies the existence of a sequence $z^1, z^2, ..., z^k$ such that $x \rhd' z^1 \rhd' \cdots \rhd' z^k \rhd' a^t$, and by TRS-condition, $a^t \not\succ^R x$, which contradicts the assumption that $x \in A^t$. The rest of the proof is to extend the transitive closure of $\succ^*$ to a strict preference by using Szpilrajn's theorem. Then it can easily be seen that the data set is rationalized by a (T)RS-model $(\succ, \Gamma)$. □

Proof of Theorem B

Let $\succ^*$ be a linear extension of $\succ^{\text{CF}}$, and define Γ such that

$$\Gamma(A) = \left[\left(\bigcup_{t:A^t \supset A} \{a^t\}\right) \cap A\right] \cup \{x \in A : y \succ^* x \text{ for all } y \in A \backslash x\}. \qquad (2.17)$$

Then $\Gamma(A) \neq \varnothing$ for every $A \subseteq X$, and $a^t \in \Gamma(A^t)$ for all $t \in \mathcal{T}$. Moreover, by definition of $\succ^{\text{CF}}$, a^t is the best alternative in $\Gamma(A^t)$ in terms of $\succ^*$. The remaining issue is whether the above defined Γ is a CF, which can be confirmed as follows. Let $\bar{x} \in A' \subseteq A''$, and let $\bar{x} \in \Gamma(A'')$. If $\bar{x} \in \{x \in A'' : y \succ^* x \text{ for all } y \in A'' \backslash x\}$, then $\bar{x}$ is the worst alternative (w.r.t. $\succ^*$) in a larger set A'', and hence it must be also the worst alternative in A'. If $\bar{x} \in [(\bigcup_{t:A^t \supset A''} \{a^t\}) \cap A'']$, then it is obvious that $\bar{x} \in [(\bigcup_{t:A^t \supset A'} \{a^t\}) \cap A']$ also holds. In both cases, it holds that $\bar{x} \in \Gamma(A')$, which implies that Γ is a CF. □

2.5 Appendix II: Backtracking

The revealed preference tests for CAF-model and (T)RS-model involve combinatorial calculations, and applying them to actual data may be computationally challenging. However, the tests become manageable with the help of a simple but powerful method called *backtracking*.[14] Here we illustrate how this method is adopted to our revealed preference tests, after a brief introduction of this method. Note that the method here is also applicable to Theorem A by De Clippel and Rozen (2021), and we actually employ the algorithm here in our data analysis.

To get the basic idea of backtracking, consider a problem where we have to select c_q from some set $\mathcal{C}_q$ for every $q = 1, 2, ..., Q$, so that the resulting selection $(c_1, c_2, \ldots, c_Q)$ obeys some constraint $\mathbf{M}_Q$. While there are $\prod_{q=1}^{Q} |\mathcal{C}_q|$ logically possible trials that we must check, the backtracking procedure may lead us to a solution with much fewer trials, especially when $\mathbf{M}_Q$ has the *cut-off* property defined below. For every $\bar{Q} < Q$, let us refer to $(c_1, c_2, \ldots, c_{\bar{Q}})$ as a *partial* selection in the sense that c_q is not yet determined for $q \in \{\bar{Q} + 1, \ldots, Q\}$. Then, we say that $\mathbf{M}_Q$ has the cut-off property if: (I) for every $\bar{Q} < Q$, there exists a constraint $\mathbf{M}_{\bar{Q}}$, which is a length-$\bar{Q}$-modified version of $\mathbf{M}_Q$; and (II) partial selection $(c'_1, c'_2, \ldots, c'_{\bar{Q}})$ violating $\mathbf{M}_{\bar{Q}}$ implies violation of $\mathbf{M}_{\bar{Q}+1}$ for any partial selection $(c'_1, c'_2, \ldots, c'_{\bar{Q}}, c_{\bar{Q}+1})$. Given the cut-off property, if some partial selection $(c'_1, c'_2, \ldots, c'_{\bar{Q}})$ violates $\mathbf{M}_{\bar{Q}}$, then there is no need to waste time on searching for subsequent components $c_{\bar{Q}+1}, \ldots, c_Q$, since there is no chance of any selection $(c'_1, c'_2, \ldots, c'_{\bar{Q}}, c_{\bar{Q}+1}, \ldots, c_Q)$ satisfying $\mathbf{M}_Q$. In fact, this feature is at the heart of backtracking, and allows us to adopt a component-by-component search for a desired selection.

[14] Some foundational references of the backtracking method are Walker (1960), Davis, Logemann and Loveland (1962), and Golomb and Baumert (1965).

Given below is a basic algorithm of the backtracking method. We consider a case where $\mathcal{C}_q$ is finite for every q, so with no loss of generality, we assume that sets $\mathcal{C}_q$ are a sets of integers.

Basic backtracking algorithm. Given sets $(\mathcal{C}_q)_{q=1}^Q$ and constraints $(\mathbf{M}_q)_{q=1}^Q$, this algorithm yields a selection $(c_1, c_2, \ldots, c_Q)$ that satisfies $\mathbf{M}_Q$, or $\varnothing$ (meaning that $\mathbf{M}_Q$ cannot be satisfied).

1. [Initialize.] Set $\bar{Q} \leftarrow 0$.
2. [Enter level $\bar{Q}+1$.] (Now $(c_1, \ldots, c_{\bar{Q}})$ obeys $\mathbf{M}_{\bar{Q}}$.) Set $\bar{Q} \leftarrow \bar{Q}+1$. Then set $c_{\bar{Q}} \leftarrow \min \mathcal{C}_{\bar{Q}}$.
3. [Test $(c_1, \ldots, c_{\bar{Q}})$.] If $(c_1, \ldots, c_{\bar{Q}})$ obeys $\mathbf{M}_{\bar{Q}}$, go to 6.
4. [Try again.] If $c_{\bar{Q}} \neq \max \mathcal{C}_{\bar{Q}}$, set $c_{\bar{Q}}$ to the next larger element of $\mathcal{C}_{\bar{Q}}$, and go to 3.
5. [Backtrack.] Set $c_{\bar{Q}} \leftarrow \min \mathcal{C}_{\bar{Q}}$ and $\bar{Q} \leftarrow \bar{Q}-1$. If $\bar{Q}=0$, return $\varnothing$ and stop. Otherwise, go to 4.
6. [Terminate.] If $\bar{Q}=Q$, return $(c_1, \ldots, c_{\bar{Q}})$ and stop. Otherwise, go to 2.

The big picture of this algorithm is as follows. The process initially starts from considering a singleton selection (c_1) and sees whether $\mathbf{M}_1$ is satisfied. If there is no such element in $\mathcal{C}_1$, then we can immediately conclude that there is no chance of finding a selection $(c_1, c_2, \ldots, c_Q)$ obeying $\mathbf{M}_Q$. If we find a successful partial selection $(c_1, c_2, \ldots, c_{\bar{Q}-1})$ and reach the $\bar{Q}$-th level, we set $c_{\bar{Q}}$ to be the minimum element in $\mathcal{C}_{\bar{Q}}$, and test whether $(c_1, c_2, \ldots, c_{\bar{Q}})$ obeys $\mathbf{M}_{\bar{Q}}$. If $\mathbf{M}_{\bar{Q}}$ is satisfied, then we proceed to the $(\bar{Q}+1)$-th level. If not, we redefine $c_{\bar{Q}}$ to be the next larger element of $\mathcal{C}_{\bar{Q}}$ and check $\mathbf{M}_{\bar{Q}}$. If we cannot find any $c_{\bar{Q}} \in \mathcal{C}_{\bar{Q}}$ such that $(c_1, c_2, \ldots, c_{\bar{Q}})$ obeys $\mathbf{M}_{\bar{Q}}$, then we go back to the $(\bar{Q}-1)$-th level and update $c_{\bar{Q}-1}$. This search algorithm terminates when we succeed in finding some $(c_1, c_2, \ldots, c_Q)$ obeying $\mathbf{M}_Q$, or it is determined that any (partial) selection with $c_1 = \max C_1$ cannot be successful.

We now show that the backtracking method is applicable to our revealed preference tests as follows. Suppose that a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ has Q revealed preference cycles. For each $q = 1, 2, ..., Q$, let $\mathcal{C}_q$ be the q-th revealed preference cycle. Then, for every $\mathcal{M} \in \{$CAF, RS, TRS$\}$, if we set $\mathbf{M}_Q$ as the joint of acyclicity and $\mathcal{M}$-condition, the revealed preference test for $\mathcal{M}$-model is equivalent to the existence problem of a selection of arcs from cycles $\mathbf{D} = (c_1, c_2, ..., c_Q)$ obeying constraint $\mathbf{M}_Q$. We claim that the above defined $\mathbf{M}_Q$ obeys the cut-off property for every $\mathcal{M} \in \{$CAF, RS, TRS$\}$.

Condition (I): We define $\mathbf{M}_{\bar{Q}}$ for every $\bar{Q} \leqslant Q$ as follows. Given a partial selection of arcs $\mathbf{D}_{\bar{Q}} = (c_1, c_2, \ldots, c_{\bar{Q}})$, note that $\mathbf{D}_{\bar{Q}}$ can be regarded as a binary relation. Therefore, acyclicity is a well-defined constraint. Now we define a partial sequence version of $\mathcal{M}$-condition, to which we refer as $\mathcal{M}_{\bar{Q}}$-condition as follows. Similar to (2.4), we can define for every $t \in \mathcal{T}$,

$$B^t_{\bar{Q}} = \{x \in A^t : (a^t, x) \in \mathbf{D}_{\bar{Q}}\}. \tag{2.18}$$

We say that a partial selection of arcs from cycles $\mathbf{D}_{\bar{Q}} = (c_1, c_2, ..., c_{\bar{Q}})$ obeys $\mathcal{M}_{\bar{Q}}$-condition, if the corresponding $\{B^t_{\bar{Q}}\}_{t \in \mathcal{T}}$ satisfies the restriction in $\mathcal{M}$-condition; specifically, $\mathbf{D}_{\bar{Q}}$ obeys CAF$_{\bar{Q}}$-condition, if it holds that for every $t \in \mathcal{T}$ and any set of indices $\tau \subseteq \mathcal{T}$,

$$\left(\bigcup_{r \in \tau} A^r \setminus \bigcup_{r \in \tau} B^r_{\bar{Q}}\right) \subseteq A^t \Longrightarrow a^t \notin \bigcup_{r \in \tau} B^r_{\bar{Q}}. \tag{2.19}$$

Similar terminology is used for other models as well. With this $\mathcal{M}_{\bar{Q}}$-condition, we let $\mathbf{M}_{\bar{Q}}$ be the joint of acyclicity and $\mathcal{M}_{\bar{Q}}$-condition, which is clearly a well-defined constraint.

Condition (II): We show that if a partial selection of arcs from cycles $\mathbf{D}_{\bar{Q}} = (c_1, c_2, ..., c_{\bar{Q}})$ does not satisfy $\mathbf{M}_{\bar{Q}}$ for some $\bar{Q} < Q$, then $\mathbf{D}_{\bar{Q}+1} = (c_1, c_2, ..., c_{\bar{Q}}, c_{\bar{Q}+1})$ cannot satisfy $\mathbf{M}_{\bar{Q}+1}$ for any $c_{\bar{Q}+1} \in \mathcal{C}_{\bar{Q}+1}$. It is obvious, if $\mathbf{D}_{\bar{Q}}$ is cyclic, then $\mathbf{D}_{\bar{Q}+1}$ cannot be acyclic. Therefore, the substantial part is $\mathcal{M}_{\bar{Q}}$-condition. However, this follows straightforwardly by taking a look at our revealed preference conditions and the construction of B^t-sets, which is shown below.

Fact 5. *If a partial selection of arcs from cycles* $\mathbf{D}_{\bar{Q}} = (c_1, c_2, ..., c_{\bar{Q}})$ *fails* $\mathcal{M}_{\bar{Q}}$*-condition, then partial selection of arcs from cycles* $\mathbf{D}_{\bar{Q}+1} = (c_1, c_2, ..., c_{\bar{Q}}, c_{\bar{Q}+1})$ *fails* $\mathcal{M}_{\bar{Q}+1}$*-condition.*

Proof. Note that the selection of arcs from cycles $1, 2, \ldots, \bar{Q}$ are the same in $\mathbf{D}_{\bar{Q}}$ and $\mathbf{D}_{\bar{Q}+1}$. Hence, it follows from (2.18) that $B^t_{\bar{Q}} \subseteq B^t_{\bar{Q}+1}$ for every $t \in \mathcal{T}$. By the structure of $\mathcal{M}_{\bar{Q}}$-condition and $\mathcal{M}_{\bar{Q}+1}$-condition, we can see the following: whenever we have "larger" B^t-sets, (i) the LHS of CAF-condition is more permissive and (ii) $\rhd$ is stronger and thus more difficult to find an acyclic (asymmetric and transitive) selection of it in (T)RS-condition. Both (i) and (ii) imply that $\mathbf{D}_{\bar{Q}+1}$ fails $\mathcal{M}_{\bar{Q}+1}$-condition whenever $\mathbf{D}_{\bar{Q}}$ fails $\mathcal{M}_{\bar{Q}}$-condition. □

Example 3. *Let* $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ *and consider a data set of six observations as follows, where for each* $t \in \mathcal{T}$*, the chosen alternative is* *underlined:*

$$A^1 = \{\underline{x_1}, x_2, x_4\}, \quad A^2 = \{x_1, \underline{x_2}\}, \quad A^3 = \{\underline{x_3}, x_4, x_6\},$$
$$A^4 = \{x_3, \underline{x_4}\}, \quad A^5 = \{x_2, \underline{x_5}, x_6\}, \quad A^6 = \{x_5, \underline{x_6}\}.$$

Let us walk through the backtracking algorithm, and see how we determine that the data set is not rationalizable by an RS-model. Note that the data set has four cycles (we order the cycles and the arcs in them as below):

1. $\mathcal{C}_1 = \{(x_1, x_2), (x_2, x_1)\}$,

2. $\mathcal{C}_2 = \{(x_3, x_4), (x_4, x_3)\}$,

3. $\mathcal{C}_3 = \{(x_5, x_6), (x_6, x_5)\}$,

4. $\mathcal{C}_4 = \{(x_1, x_4), (x_4, x_3), (x_3, x_6), (x_6, x_5), (x_5, x_2), (x_2, x_1)\}$.

For every $\bar{Q} \in \{1, 2, 3, 4\}$*, let us denote by* $\mathbf{M}_{\bar{Q}}$ *the joint of acyclicity and* $RS_{\bar{Q}}$*-condition. Following our backtracking procedure, we first set* $\bar{Q} = 1$ *and set* $c_1 = (x_1, x_2)$*, which is the first arc of the first cycle. Since single element selection* $((x_1, x_2))$ *obeys* $\mathbf{M}_1$*, we proceed to the second cycle by setting* $\bar{Q} = 2$*. Here we set* $c_2 = (x_3, x_4)$ *and check whether* $((x_1, x_2), (x_3, x_4))$ *obeys* $\mathbf{M}_2$*, which is affirmative. Then we go to the third cycle by setting* $\bar{Q} = 3$ *and set* $c_3 = (x_5, x_6)$*. In fact, this partial selection of arcs from cycles* $((x_1, x_2), (x_3, x_4), (x_5, x_6))$ *fails to satisfy* $\mathbf{M}_3$*, specifically* RS_3*-condition. In this case, we keep* $\bar{Q} = 3$*, and update* c_3 *to the next arc in* $\mathcal{C}_3$*, and set* $c_3 = (x_6, x_5)$*. Then, we test whether this updated selection of arcs* $((x_1, x_2), (x_3, x_4), (x_6, x_5))$ *obeys* $\mathbf{M}_3$*, which is negative. At this point, we can determine that it is impossible to find a selection* (c_1, c_2, c_3, c_4) *obeying RS-condition and acyclicity as long as* $(x_1, x_2), (x_3, x_4)$ *are selected from* $\mathcal{C}_1, \mathcal{C}_2$

$\bar{Q}$	(partial) selection	$\mathbf{M}_{\bar{Q}}$
1	$((x_1, x_2))$	PASS
2	$((x_1, x_2), (x_3, x_4))$	PASS
3	$((x_1, x_2), (x_3, x_4), (x_5, x_6))$	FAIL
3	$((x_1, x_2), (x_3, x_4), (x_6, x_5))$	FAIL
2	$((x_1, x_2), (x_4, x_3))$	FAIL
1	$((x_2, x_1))$	FAIL
0	$\varnothing$	**STOP**

Table 2.2: Backtracking procedure applied to Example 2 for testing RS-model.

respectively. Thus we backtrack $\bar{Q}$ to 2, and update c_2 to (x_4, x_3). Looking at $((x_1, x_2), (x_4, x_3))$, it fails $\mathbf{M}_2$. Since there is no chance of success unless (x_1, x_2) is discarded from the selection, we rewind $\bar{Q}$ to 1, and update c_1 to (x_2, x_1). Then we check whether $((x_2, x_1))$ obeys $\mathbf{M}_1$, which is negative. Then $\bar{Q}$ is set to 0 and the algorithm terminates, which means that the data set is not rationalizable by RS-model.

Remark 1: One advantage of the backtracking approach is that we may be able to determine, at an early stage of the process of search, that a data set fails the test. Due to this feature, calculation time does depend on how we order the cycles. We suggest that the cycles are sorted so that shorter cycles come first: whenever $q' < q''$, q'-th cycle is weakly shorter than q''-th cycle. The cycles in Example 2 are sorted in this way. Whenever this takes too much calculation time, it seems natural to list "problematic" cycles first. Problematic cycles are those such that a (partial) selection of arcs from cycles fails when adding an arc at that cycle. This may allow us to determine that a data set fails the test at an early stage of the backtracking process (and we actually adopt this type of strategy).

Remark 2: Backtracking can be applied to De Clippel and Rozen (2021)'s

AF-test as well. Recall that their test requires the existence of an acyclic binary relation $\succ^*$ such that, for every $s, t \in \mathcal{T}$, with $a^s, a^t \in A^s \cap A^t$ and $a^s \neq a^t$,

$$\exists x \in A^s \backslash A^t : a^s \succ^* x \text{ or } \exists x \in A^t \backslash A^s : a^t \succ^* x.$$

Suppose there are $Q > 0$ pairs of observations (s, t) such that $a^s, a^t \in A^s \cap A^t$ and $a^s \neq a^t$. It can be seen that backtracking is applicable to De Clippel and Rozen's test, by letting $\mathbf{M}_Q$ be acyclicity, and for q-th pair (s, t), defining

$$\mathcal{C}_q = \left\{(x'', x') : [x'' = a^s \text{ and } x' \in A^s \backslash A^t] \text{ or } [x'' = a^t \text{ and } x' \in A^t \backslash A^s]\right\}.$$

Chapter 3

Growing consideration

This chapter puts forward a behavioral framework where a decision maker (DM) makes choices at multiple time periods, while she may not be aware of all available alternatives at all times. This framework inherits the main idea of the theory of limited consideration, and adds to it an assumption that the DM's consideration grows over time. In particular, it is required that she takes into consideration any alternative that she chose in the past. We refer to this property as growing consideration. Revealed preference tests, as well as conditions under which we can robustly infer the decision maker's preference, consideration, and non-consideration are given. Following a revealed preference analysis of a baseline framework of decision making under growing consideration, we deal with special cases where the DM's consideration is a competition filter/attention filter.

Let X be the grand set of alternatives, which we assume to be finite. The rational choice model assumes that a DM maximizes her strict preference $\succ$, given any feasible set $A \subseteq X$. In this framework, it is typically assumed that an economist observes a choice function f of the DM: letting $\mathcal{D} \subseteq 2^X \backslash \varnothing$ be an arbitrary collection of nonempty feasible sets, for every feasible set $A \in \mathcal{D}$,

$f(A) \in A$ is the chosen alternative from A. It is well known that a choice function is consistent with the rational choice model, if and only if it obeys the Strong Axiom of Revealed Preference (SARP). However, it is reported in a number of experimental studies that violation of SARP is commonly observed. In order to deal with such seemingly irrational behavior, various theories of bounded rationality have been proposed. Amongst others, in this chapter as well as Chapter 2, we adopt the behavioral assumption of limited consideration.

Here, we supplement this limited consideration framework by adding "time" into it: we put forward a framework where a DM makes decisions at multiple time periods, while she may not take into consideration all available alternatives at all times. In particular, we assume that a DM has a time-invariant strict preference $\succ$, and letting $\mathcal{T} = \{1, 2, \ldots, T\}$ be a set of time periods, for every time period $t \in \mathcal{T}$, there exists a consideration mapping $\Gamma_t : 2^X \backslash \varnothing \to 2^X \backslash \varnothing$. Then, given any feasible set A at time period t, the DM chooses the $\succ$-best alternative within $\Gamma_t(A)$. In addition, we assume that the DM's consideration depends on past choices. Specifically, we assume that any alternative chosen in a past period must be considered. We refer to this property as *growing consideration.* For example, any commodity that the DM consumed in the past should be familiar to her and thus be easier to spot, or particular websites show what choices she made previously, so it is an assumption that reflects everyday decision making.

We formalize the framework of decision making under growing consideration, and provide a necessary and sufficient condition for a DM's choices to be consistent with it. Furthermore, we derive conditions under which we can robustly infer the DM's preference, consideration, and non-consideration, provided that observed choices are in line with the growing consideration framework. By conditions for robust inference, we mean conditions under which we can surely say that (i) some alternative is preferred to another; (ii) some alternative is considered at some feasible set and time period; and (iii) some alternative is not considered at some set and period. Such inferences are useful from the viewpoint of welfare analysis, since we cannot pin down a DM's preference or consideration even when choices are consistent with the growing consideration framework. For example, under this framework, an alternative x'' being chosen over x' does not directly imply that x'' is preferred to x': we must also take into account the possibility that x' is preferred to x'', but x' is overlooked. Thus, inferring the DM's preference/(non-)consideration is a non-trivial exercise. In this chapter, following a revealed preference analysis of a baseline framework of decision making under growing consideration, we deal with special cases where we require that consideration mapping Γ_t is a competition filter/attention filter for every time period t.

A paper closely related to this chapter is Ferreira and Gravel (2017), in which they explicitly analyze a situation where choices are observed across multiple time periods. They provide revealed preference tests for choice models with (i) changing preference; (ii) preference formation by trial and error; and (iii) endogenous status-quo bias. One difference between Ferreira and Gravel (2017) and our analyses, apart from the models analyzed, is the structure of the data set assumed to be observed. While Ferreira and Gravel (2017) assume that a choice from only one feasible set is observed for every time period, we adopt a more general assumption and deal with the case where a choice function is observed for each time period. Other related papers are Bernheim and Rangel (2007, 2009) and Salant and Rubinstein (2008): the framework we put forward have some similarity with theirs. These papers analyze a situation where each feasible set is supplemented with an additional condition, which is referred to as an *ancillary condition* by the former and as a *frame* by the latter. Such conditions represent either "some characteristic of the choice environment that is consequently irrelevant to outside observer" or "how alternatives are framed". In principle, "time of choice" can be regarded as such a condition, and in fact Bernheim and Rangel give it as an example of an ancillary condition.[1] In this regard, our framework can be seen as a special case of the framework dealt with in these papers. However, there are substantial differences that distinguish our framework from those of Bernheim and Rangel (2007, 2009) and Salant and Rubinstein (2008). Firstly, the limited consideration/growing consideration assumption is a feature not covered by Bernheim and Rangel (2007, 2009) or Salant and Rubinstein (2008).[2] Secondly, as a technical issue, Bernheim and Rangel/Salant and Rubinstein assume that for every ancillary condition/frame, a choice function (or correspondence) defined on an exhaustive domain is observed, while we assume that the choice function observed for each time period is defined on an arbitrary collection of feasible sets.[3] Therefore, results derived by Bernheim and Rangel/Salant and Rubinstein are not directly applicable to our context. Furthermore, Bernheim and Rangel (2007, 2009) propose a conservative criterion under which welfare judgements can be made, but we show that our results regarding robust inference of preference may lead to completely opposite welfare implications.

[1]Salant and Rubinstein, in their paper, do not refer to time as a frame.

[2]In fact, Salant and Rubinstein (2008) deal with a specific type of limited consideration, where they apply the "number of alternatives that the DM can pay attention to" as the frame. Such a behavioral assumption is different from the limited consideration that we consider. Moreover, if we attempt to illustrate our framework in terms of Salant and Rubinstein (2008), we must apply both "time of choice" *and* "structure of limited consideration" as the frame. Such a case is not considered by Salant and Rubinstein (2008).

[3]By an exhaustive domain of a choice function, we mean that the choice function is defined on all nonempty subsets of X, i.e., $f : 2^X \backslash \varnothing \to X$.

Organization of this Chapter: In Section 3.1, we formalize the framework of decision making under growing consideration, and define the concepts of rationalizability and robust inference of preference/(non-)consideration. Section 3.2 is devoted to an analysis of observable restrictions of decision making under growing consideration. In particular, we provide a revealed preference test in Section 3.2.1, and then derive conditions for robust inference of preference/(non-)consideration in Section 3.2.2. In Section 3.3, we analyze a special case of growing consideration, where we require that consideration mapping of each time period is a competition filter. A revealed preference test for decision making under growing consideration with attention filters is given in Appendix.

3.1 Preliminaries

Let X be a finite set, and consider a choice framework where a DM makes choices over multiple time periods. We assume that the DM has a time-invariant, connected, transitive, and asymmetric preference $\succ$, which we refer to as a *strict preference*. In addition, we assume that the DM exhibits limited consideration. Letting $\mathcal{T} = \{1, 2, \ldots, T\}$ be the set of time periods, for every period $t \in \mathcal{T}$, there is a consideration mapping $\Gamma_t : 2^X \backslash \varnothing \to 2^X \backslash \varnothing$ such that $\Gamma_t(A) \subseteq A$ for every $A \subseteq X$. Then, given any feasible set $A \subseteq X$ and period $t \in \mathcal{T}$, the DM chooses the $\succ$-best alternative within $\Gamma_t(A)$. Regarding the structure of DM's consideration, assume that alternatives chosen in past periods must be considered. We refer to this property as *growing consideration*.

A DM's choices at time period $t \in \mathcal{T}$ are summarized in a choice function f_t. In particular, for every $t \in \mathcal{T}$, let $\mathcal{D}_t \subseteq 2^X \backslash \varnothing$ be the collection of feasible sets observed at period t. The profile of choice functions $(f_t)_{t \in \mathcal{T}}$ is such that $f_t : \mathcal{D}_t \to X$ and $f_t(A) \in A$ for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$, where $f_t(A)$ is the alternative chosen by the DM from feasible set A at time period t. Let us abuse terminology and refer to $(f_t)_{t\in\mathcal{T}}$ simply as a "choice function" when there is no fear of confusion. Then, growing consideration is formally defined as below:

Definition 2. *Given a profile of choice functions $(f_t)_{t\in\mathcal{T}}$, a profile of consideration mappings $(\Gamma_t)_{t\in\mathcal{T}}$ obeys growing consideration, if alternatives chosen in the past are included in consideration sets. That is, for every $t' \in \mathcal{T}$, $A' \subseteq X$, and $x \in A'$,*

$$x \in \Gamma_{t'}(A') \text{ if there exist } t < t' \text{ and } A \in \mathcal{D}_t \text{ such that } x = f_t(A). \tag{3.1}$$

We refer to a pair of strict preference and profile of consideration mappings $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ as a growing consideration model, or in short a GC-model, whenever $(\Gamma_t)_{t\in\mathcal{T}}$ obeys growing consideration.

In this chapter, we derive a necessary and sufficient condition that a profile of choice functions $(f_t)_{t\in\mathcal{T}}$ must obey, in order for it to be rationalizable by a GC-model. A formal definition of rationalizability is as follows:

Definition 3. *A profile of choice functions $(f_t)_{t\in\mathcal{T}}$ is rationalizable by a growing consideration model (GC-rationalizable), if there exists a growing consideration model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ such that, for every $t' \in \mathcal{T}$ and $A' \in \mathcal{D}_{t'}$, $f_{t'}(A')$ is the $\succ$-best alternative within $\Gamma_{t'}(A')$. In this case, we say that such a GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ rationalizes $(f_t)_{t\in\mathcal{T}}$.*

When choices are GC-rationalizable, an alternative x'' chosen over x' does not necessarily imply that the DM prefers x'' to x': it may be the case that x' is preferred to x'', but x' was overlooked. Moreover, the GC-model that rationalizes choices is not uniquely determined in general. Nevertheless, it is possible to pin down the relative ranking between particular alternatives, or robustly infer that some alternative is considered/ignored at some feasible set and time period. These conditions are derived in Section 3.2.2 (and in Section 3.3.2 for growing consideration with competition filters). Henceforth, for notational simplicity, let us use the expression (t, A) when dealing with feasible set A at time period t.

Definition 4. *Let $(f_t)_{t\in\mathcal{T}}$ be rationalizable by a growing consideration model. Then:*

- *x'' is robustly preferred to x' if $x'' \succ x'$ holds under every GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$;*
- *x' is robustly considered at (t', A') if $x' \in \Gamma_{t'}(A')$ holds under every GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$;*
- *x' is robustly* **not** *considered at (t', A') if $x' \notin \Gamma_{t'}(A')$ holds under every GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$.*

In what follows, we derive observable restrictions of decision making under growing consideration, namely conditions for rationalizability and robust inference. A baseline framework, where no intra-temporal restriction is casted on $(\Gamma_t)_{t\in\mathcal{T}}$, is dealt with first, followed by a case where we require Γ_t to be a competition filter for every $t \in \mathcal{T}$. In Appendix, we provide a revealed preference test for decision making under growing consideration, where each Γ_t is an attention filter.

3.2 Observable restrictions of growing consideration

In this section, we derive observable restrictions of a baseline framework of decision making under growing consideration, where we cast no intra-temporal restriction on consideration mappings $(\Gamma_t)_{t\in\mathcal{T}}$. It is worth noting that when $T = 1$, which corresponds to the standard limited consideration framework with no "time," rationalizability becomes vacuous without any restriction on the structure of Γ. Given a choice function f, we can simply set $\Gamma(A) = \{f(A)\}$ for every $A \in \mathcal{D}$, and set $\Gamma(A) \subseteq A$ arbitrarily for $A \notin \mathcal{D}$. Then, any strict preference $\succ$ accompanied with this consideration mapping Γ would rationalize the choice function. When $T \geqslant 2$, testing for rationalizability would have a bite, which is shown below.

3.2.1 Test for GC-rationalizability

Suppose that choice function $(f_t)_{t\in\mathcal{T}}$ is generated by a DM obeying a growing consideration model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$. Fixing any $t \geqslant 2$ and $A \in \mathcal{D}_t$, whenever there exist $t' < t$ and $A' \in \mathcal{D}_{t'}$ with $f_{t'}(A') \in A$, it follows by growing consideration that $f_{t'}(A') \in \Gamma_t(A)$. This motivates us to define a binary relation P_t for every period $t \geqslant 2$ as follows: $x'' P_t x'$ if there exists $A \in \mathcal{D}_t$ such that $x'' = f_t(A)$, and

$$\text{there exist } t' < t \text{ and } A' \in \mathcal{D}_{t'} \text{ such that } x' = f_{t'}(A') \in A \backslash x''. \tag{3.2}$$

Note that whenever $x'' P_t x'$ holds, then we have $x'' \succ x'$. Now let us define a binary relation P as a union of all P_t's, i.e., $P = \bigcup_{t=2}^{T} P_t$. Then under growing consideration, $x'' P x'$ implies $x'' \succ x'$, and therefore, acyclicity of P is a necessary condition for $(f_t)_{t\in\mathcal{T}}$ to be GC-rationalizable.[4] In fact, the opposite direction is true as well.

Proposition 1. *A profile of choice functions $(f_t)_{t\in\mathcal{T}}$ is rationalizable by a*

[4] A binary relation P is *acyclic*, if for any $x^1, x^2, \ldots, x^K \in X$, $x^1 P x^2 P \cdots P x^{K-1} P x^K$ implies not $x^K P x^1$.

growing consideration model if and only if binary relation P is acyclic.

Proof. We show sufficiency here by constructing a GC-model that rationalizes $(f_t)_{t\in\mathcal{T}}$. Since binary relation P is acyclic, by Szpilrajn's Theorem, there exists a connected, asymmetric, and transitive extension of P, which we denote by $\succ$. We define consideration mapping Γ_t for every $t \in \mathcal{T}$ as follows:

$$\text{if } A \in \mathcal{D}_t, \quad \Gamma_t(A) = \{f_t(A)\} \cup \{x \in A : f_t(A) \succ x\}; \tag{3.3}$$

$$\begin{aligned} \text{if } A \notin \mathcal{D}_t, \quad \Gamma_t(A) =& \{x \in A : x = f_{t'}(A'), \exists A' \in \mathcal{D}_{t'}, \exists t' < t\} \\ & \cup \{x \in A : y \succ x, \forall y \in A \backslash x\}. \end{aligned} \tag{3.4}$$

Note that $\Gamma_t(A) \neq \varnothing$ for every (t, A), and $f_t(A) \in \Gamma_t(A)$ for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$. Thus $(\Gamma_t)_{t\in\mathcal{T}}$ is a well-defined profile of consideration mappings, and it is clear by construction that $f_t(A)$ is the $\succ$-best alternative within $\Gamma_t(A)$ for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$. It remains to show that $(\Gamma_t)_{t\in\mathcal{T}}$ obeys growing consideration. Fix any $t \geqslant 2$ and $A \subseteq X$, and take any $x' \in A$ such that $x' = f_{t'}(A')$ for some $t' < t$ and $A' \in \mathcal{D}_{t'}$. If $A \in \mathcal{D}_t$, then we have either $f_t(A) = x'$ or $f_t(A) P_t x'$, and the latter case in turn implies $f_t(A) \succ x'$. In both cases, by (3.3), we have $x' \in \Gamma_t(A)$. If $A \notin \mathcal{D}_t$, then it follows immediately from (3.4) that $x' \in \Gamma_t(A)$. Summarizing, we conclude that $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ is a GC-model that rationalizes $(f_t)_{t\in\mathcal{T}}$. □

Example 4. *We give here a profile of choice functions that is not GC-rationalizable. Let $X = \{x_1, x_2, x_3, x_4\}$, $\mathcal{T} = \{1, 2\}$, and consider (f_1, f_2) summarized in Table 3.1. Since x_2 is chosen at $t = 2$ from $\{x_1, x_2, x_3\}$, and x_1 is an alternative chosen in the past ($t = 1$), it follows that $x_2 P x_1$. Similarly, we have $x_1 P x_2$. Since binary relation P has a cycle, $(f_t)_{t\in\mathcal{T}}$ is not GC-rationalizable. Indeed, if we attempt to find a preference and profile of consideration mappings that constitute a choice-rationalizing GC-model, it*

$A \in \mathcal{D}_1$	$\{x_1, x_2\}$	$\{x_1, x_2, x_3\}$
$f_1(A)$	x_1	x_2
$A \in \mathcal{D}_2$	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_3, x_4\}$
$f_2(A)$	x_2	x_1

Table 3.1: Choice function $(f_t)_{t\in\mathcal{T}}$ of Example 4.

must be the case that x_1, x_2 are considered at $t = 2$ at both $\{x_1, x_2, x_3\}$ and $\{x_1, x_2, x_3, x_4\}$. Then, since x_2 and x_1 are chosen alternatives at $t = 2$ from $\{x_1, x_2, x_3\}$ and $\{x_1, x_2, x_3, x_4\}$ respectively, it must follow that x_2 is strictly preferred to x_1 and vice versa, which is impossible.

3.2.2 Robust inference

Even when a profile of choice functions $(f_t)_{t\in\mathcal{T}}$ is GC-rationalizable, the GC-model that rationalizes it is not uniquely determined in general. Nevertheless, there are cases where we can robustly infer the DM's preference, consideration, and non-consideration. Here we derive conditions under which such inferences can be made. Let us denote by P^{TC} the transitive closure of P.[5] Henceforth, we use superscript "TC" to denote the transitive closure of any binary relation.

Proposition 2. *Suppose that a profile of choice functions $(f_t)_{t\in\mathcal{T}}$ is rationalizable by a growing consideration model. Then:*

1. *x'' is robustly preferred to x' if and only if $x'' P^{TC} x'$;*

2. *$x' \in A'$ is robustly considered at (t', A') if and only if*

[5] We have $x'' P^{TC} x'$, if there exist $z^0, z^1, \ldots, z^K \in X$ such that $x'' = z^0, x' = z^K$, and $z^{k-1} P z^k$ for every $k \in \{1, 2, \ldots, K\}$.

(a) $x' = f_{t'}(A')$, or

(b) there exist $t'' < t'$ and $A'' \in \mathcal{D}_{t''}$ such that $x' = f_{t''}(A'')$;

3. $x' \in A'$ is robustly not considered at (t', A') if and only if $x' P^{TC} f_{t'}(A')$.[6]

Proof. We first show 1. Since sufficiency of 1 is clear, we prove necessity by showing the contrapositive. Suppose that $x'' P^{TC} x'$ does not hold. Then, it is known that there exists an extension $\succ$ of P^{TC} such that (i) $\succ$ is connected, asymmetric, and transitive; and (ii) $x' \succ x''$.Defining $(\Gamma_t)_{t\in\mathcal{T}}$ as (3.3) and (3.4), we have a GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ with $x' \succ x''$ that rationalizes $(f_t)_{t\in\mathcal{T}}$.

To show sufficiency of 2, take any GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$. Whenever (a) or (b) holds, $x' \in \Gamma_{t'}(A')$ follows, by definition of consideration mappings in the case of (a), and by definition of growing consideration in the case of (b). We prove necessity of 2 by showing the contrapositive. Suppose that neither (a) nor (b) holds. Letting $\succ$ be a connected, transitive, and asymmetric extension of P, define $(\Gamma_t)_{t\in\mathcal{T}}$ as follows: for every $(t, A) \neq (t', A')$,

$$\begin{aligned} \text{if } A \in \mathcal{D}_t, \quad \Gamma_t(A) =& \{f_t(A)\} \cup \{x \in A : x = f_{t''}(A''), \exists A'' \in \mathcal{D}_{t''}, \exists t'' < t\}; \\ \text{if } A \notin \mathcal{D}_t, \quad \Gamma_t(A) =& \{x \in A : x = f_{t''}(A''), \exists A'' \in \mathcal{D}_{t''}, \exists t'' < t\} \\ & \cup \{x \in A : y \succ x, \forall y \in A \backslash x\}; \end{aligned}$$

and for (t', A'),

$$\begin{aligned} \text{if } A' \in \mathcal{D}_{t'}, \quad \Gamma_{t'}(A') =& \{f_{t'}(A')\} \cup \{x \in A' \backslash x' : x = f_{t''}(A''), \exists A'' \in \mathcal{D}_{t''}, \exists t'' < t'\}; \\ \text{if } A' \notin \mathcal{D}_{t'}, \quad \Gamma_{t'}(A') =& \{x \in A' \backslash x' : x = f_{t''}(A''), \exists A'' \in \mathcal{D}_{t''}, \exists t'' < t'\} \\ & \cup \{x \in A' \backslash x' : y \succ x, \forall y \in A' \backslash \{x, x'\}\}. \end{aligned}$$

[6] Note that 2-(a) and 3 have a bite only if $A' \in \mathcal{D}_{t'}$.

$A \in \mathcal{D}_1$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_1, x_3, x_4\}$
$f_1(A)$	x_2	x_3	x_1

$A \in \mathcal{D}_2$	$\{x_1, x_2\}$	$\{x_2, x_3\}$
$f_2(A)$	x_1	x_2

Table 3.2: Choice function $(f_t)_{t\in\mathcal{T}}$ of Example 5.

Then, for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$, $f_t(A)$ is the $\succ$-best alternative within $\Gamma_t(A)$, since $f_t(A)P_t x$ holds for every $x \in \Gamma_t(A)\backslash f_t(A)$. By construction of $(\Gamma_t)_{t\in\mathcal{T}}$, growing consideration holds, and we have $x' \notin \Gamma_{t'}(A')$. Thus, $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}}\rangle$ is a GC-model that rationalizes $(f_t)_{t\in\mathcal{T}}$, while x' is not considered at (t', A').

To show sufficiency of 3, take any GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}}\rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$, and suppose that $x' P^{TC} f_{t'}(A')$. It follows from 1 that $x' P^{TC} f_{t'}(A')$ implies $x' \succ f_{t'}(A')$. Meanwhile, since $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}}\rangle$ rationalizes $(f_t)_{t\in\mathcal{T}}$, $f_{t'}(A')$ is the $\succ$-best element in $\Gamma_{t'}(A')$. Hence $x' \notin \Gamma_{t'}(A')$ follows. We prove necessity of 3 by showing the contrapositive. Suppose that $x' P^{TC} f_{t'}(A')$ does not hold. Then there exists an extension $\succ$ of P^{TC} such that (i) $\succ$ is connected, asymmetric, and transitive; and (ii) $f_{t'}(A') \succ x'$. Defining $(\Gamma_t)_{t\in\mathcal{T}}$ as (3.3) and (3.4), we have a GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}}\rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$ with $x' \in \Gamma_{t'}(A')$. $\square$

Below we give an example of a choice function that is GC-rationalizable, and demonstrate how robust inference can be conducted.

Example 5. *Let $X = \{x_1, x_2, x_3, x_4\}$ and $\mathcal{T} = \{1, 2\}$, and consider choice function $(f_t)_{t\in\mathcal{T}}$ summarized in Table 3.2. We first show that $(f_t)_{t\in\mathcal{T}}$ is GC-rationalizable. By definition of binary relation P, it follows that $x_1 P x_2$ and $x_2 P x_3$. Since P is acyclic, $(f_t)_{t\in\mathcal{T}}$ is GC-rationalizable. Regarding* *robust inference, it follows that x_1 is robustly preferred to x_2, x_2 is robustly preferred to x_3, and x_1 is robustly preferred to x_3, i.e., we can surely say that $x_1 \succ x_2 \succ x_3$. Looking at 2-(a) of Proposition 2, we see that, for example, x_2 is robustly considered at $(t' = 1, A' = \{x_1, x_2\})$. Focusing on 2-(b), first note that x_1, x_2, x_3 are chosen at $t = 1$. Therefore, for $i \in \{1, 2, 3\}$, x_i is robustly considered at $A \subseteq X$ at $t = 2$ whenever $x_i \in A$. Finally, since $x_1 P^{TC} x_3$ and $x_3 = f_1(x_1, x_3)$, x_1 is robustly not considered at $(t' = 1, A' = \{x_1, x_3\})$.*

3.3 Growing consideration with competition filters

In this section, we derive observable restrictions of decision making under growing consideration, where we cast an intra-temporal restriction on consideration mapping of each time period. In particular, for every time period $t \in \mathcal{T}$, we require Γ_t to be a *competition filter (CF)*, which is defined as follows: for every $A' \subseteq A''$ and $x \in A'$, $x \in \Gamma(A'')$ implies $x \in \Gamma(A')$; that is, an alternative considered at a larger feasible set must be considered at a smaller feasible set.[7] We first provide a necessary and sufficient condition under which $(f_t)_{t\in\mathcal{T}}$ is rationalizable by a growing consideration with competition filters, i.e., Γ_t is a CF for every $t \in \mathcal{T}$. We refer to such a GC-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ as a *growing consideration model with CF*, or in short a *GC(CF)-model*, and refer to rationalizability by such a model as *GC(CF)-rationalizability.* Then, we derive conditions for robust inference of preference/(non-)consideration, provided that $(f_t)_{t\in\mathcal{T}}$ is GC(CF)-rationalizable.

[7]Note that the expression of CF given here is equivalent to (2.2). For details of CF, see Section 2.1 in Chapter 2.

3.3.1 Test for GC(CF)-rationalizability

Suppose that choice function $(f_t)_{t\in\mathcal{T}}$ is generated by a DM obeying a growing consideration model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}}\rangle$ with CF. We derive observable implications of decision making under growing consideration with CF, by inferring the DM's consideration and preference from the choice function $(f_t)_{t\in\mathcal{T}}$. To begin with, consider period $t = 1$. For distinct alternatives $x', x'' \in X$, $x'' \succ x'$ can be inferred by the following logic: if there exist $A, A' \in \mathcal{D}_t$ such that $x'' = f_t(A)$,

$$x', x'' \in A \subseteq A', \text{ and } x' = f_t(A'), \tag{3.5}$$

then, since Γ_t is a CF, we must have $x' \in \Gamma_t(A)$, which in turn implies $f_t(A) = x'' \succ x'$. This motivates us to define a binary relation Q_1 as follows: for distinct alternatives $x', x'' \in X$, $x''Q_1x'$ if there exist $A, A' \in \mathcal{D}_1$ such that $x'' = f_1(A)$, and (3.5) holds for $t = 1$. Now fix any period $t \geqslant 2$ and any $A \subseteq X$. We can infer $x' \in \Gamma_t(A)$ for some $x' \in A$, if (i) there exists $A' \supseteq A$ such that $x' = f_t(A')$; or (ii) there exist $t' < t$ and $A' \in \mathcal{D}_{t'}$ with $x' = f_{t'}(A')$. Then let us define binary relation Q_t as follows: for distinct $x', x'' \in X$, $x''Q_tx'$ if there exists $A \in \mathcal{D}_t$ such that $x'' = f_t(A)$ and (a) there exists $A' \in \mathcal{D}_t$ that obeys (3.5), or (b) there exist $t' < t$ and $A' \in \mathcal{D}_{t'}$ as in (3.2). Note that letting $Q = \bigcup_{t\in\mathcal{T}} Q_t$, $x''Qx'$ implies $x'' \succ x'$. Thus, binary relation Q is acyclic under growing consideration with CF. In fact, acyclicity of Q is not only necessary but also sufficient for rationalizability.

Proposition 3. *A profile of choice functions $(f_t)_{t\in\mathcal{T}}$ is rationalizable by a growing consideration model with CF, if and only if binary relation Q is acyclic.*

Proof. We show sufficiency here. Let $\succ$ be a connected, asymmetric, and transitive extension of Q, and define consideration mapping $(\Gamma_t)_{t\in\mathcal{T}}$ as follows: for every $t \in \mathcal{T}$ and every $A \subseteq X$,

$$\begin{aligned}\Gamma_t(A) = &\{x \in A : x = f_t(A''), \exists A'' \supseteq A\} \\ &\cup \{x \in A : x = f_{t''}(A''), \exists A'' \in \mathcal{D}_{t''}, \exists t'' < t\} \\ &\cup \{x \in A : y \succ x, \forall y \in A \backslash x\}.\end{aligned} \tag{3.6}$$

Note that $\Gamma_t(A) \neq \varnothing$ for every (t, A), and $f_t(A) \in \Gamma_t(A)$ for every t and $A \in \mathcal{D}_t$, so $(\Gamma_t)_{t\in\mathcal{T}}$ is a well-defined profile of consideration mappings. Moreover, it follows immediately from construction that $(\Gamma_t)_{t\in\mathcal{T}}$ obeys growing consideration. Now fix any $t \in \mathcal{T}$ and $A \in \mathcal{D}_t$, and take any $x \in \Gamma_t(A)\backslash f_t(A)$. This means that (i) there exists $A'' \supset A$ such that $x = f_t(A'')$; (ii) there exist $t'' < t$ and $A'' \in \mathcal{D}_{t''}$ such that $x = f_{t''}(A'')$; or (iii) x is the $\succ$-worst alternative in A. In cases (i) and (ii), we have $f_t(A)Q_t x$, which in turn implies $f_t(A) \succ x$; and in case (iii) we have $f_t(A) \succ x$. This shows that $f_t(A)$ is the $\succ$-best alternative within $\Gamma_t(A)$ for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$. It remains to show that for every $t \in \mathcal{T}$, Γ_t is a CF. Fix any $A, A' \subseteq X$ such that $A \subseteq A'$, and take any $x \in \Gamma_t(A') \cap A$. Note that $x \in \Gamma_t(A')$ implies: (a) there exists $A'' \supseteq A'$ with $x = f_t(A'')$; (b) there exist $t'' < t$ and $A'' \in \mathcal{D}_{t''}$ with $x = f_{t''}(A'')$; or (c) x is the $\succ$-worst alternative within A'. Taking a look at (3.6), we have $x \in \Gamma_t(A)$ in all of these cases. □

Example 5 (continued). *We show that $(f_t)_{t\in\mathcal{T}}$ in Table 3.2 is not GC(CF)-rationalizable. Note that we have $x_1Qx_2Qx_3Qx_1$: $x_1Q_2x_2$ because x_1 is chosen from $\{x_1, x_2\}$ at $t = 2$ and x_2 is a chosen alternative at $t = 1$; $x_2Q_2x_3$ holds following an analogous logic; and $x_3Q_1x_1$ holds because, at $t = 1$,*

$x_1, x_3 \in \{x_1, x_3\} \subset \{x_1, x_3, x_4\}$, x_1 is chosen from $\{x_1, x_3, x_4\}$, and x_3 is chosen from $\{x_1, x_3\}$. Since Q has a cycle, $(f_t)_{t\in\mathcal{T}}$ is not rationalizable by a GC(CF)-model.

3.3.2 Robust inference

Suppose that a profile of choice functions $(f_t)_{t\in\mathcal{T}}$ is GC(CF)-rationalizable. As stated before, the GC(CF)-model that rationalizes $(f_t)_{t\in\mathcal{T}}$ is not uniquely determined in general. Nevertheless, it is possible to make inference regarding the DM's preference and consideration, which is shown in the proposition below.

Proposition 4. *Suppose that a profile of choice functions $(f_t)_{t\in\mathcal{T}}$ is rationalizable by a growing consideration with CF. Then:*

1. *x'' is robustly preferred to x' if and only if $x'' Q^{TC} x'$;*
2. *$x' \in A'$ is robustly considered at (t', A') if and only if*
   (a) *there exists $A'' \supseteq A'$ such that $x' = f_{t'}(A'')$, or*
   (b) *there exist $t'' < t'$ and $A'' \in \mathcal{D}_{t''}$ such that $x' = f_{t''}(A'')$;*
3. *$x' \in A'$ is robustly not considered at (t', A') if and only if $x' Q^{TC} f_{t'}(A'')$ for some A'' such that $x' \in A'' \subseteq A'$.*[8]

Proof. To show sufficiency of 1, take any GC(CF)-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$, and suppose $x'' Q^{TC} x'$. In fact, it suffices to show that $x'' Q x'$ implies $x'' \succ x'$. Whenever we have $x'' Q x'$, one of the following holds: (i) there exist $t \in \mathcal{T}$ and $A, A' \in \mathcal{D}_t$ such that $x', x'' \in A \subset A'$, $x'' = f_t(A)$, and $x' = f_t(A')$; or (ii) there exist t, t' with $t' < t$ and $A \in \mathcal{D}_t, A' \in \mathcal{D}_{t'}$ such

[8]Note that Q^{TC} is the transitive closure of Q.

that $x'' = f_t(A)$, $x' = f_{t'}(A')$, and $x' \in A$. In both cases, we have $x' \in \Gamma_t(A)$, which in turn implies $x'' \succ x'$. We prove necessity of 1 by showing the contrapositive. Suppose that $x'' Q^{TC} x'$ does not hold. Then there exists and extension $\succ$ of Q^{TC} such that (i) $\succ$ is connected, asymmetric, and transitive; and (ii) $x' \succ x''$. Defining $(\Gamma_t)_{t\in\mathcal{T}}$ as (3.6), $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ is a GC(CF)-model that rationalizes $(f_t)_{t\in\mathcal{T}}$, while $x' \succ x''$ holds.

To show sufficiency of 2, take any GC(CF)-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$. If (a) holds, then since Γ_t is a CF, we must have $x' \in \Gamma_{t'}(A')$; and if (b) holds, then by growing consideration we must have $x' \in \Gamma_{t'}(A')$. We prove necessity of 2 by showing the contrapositive. Suppose that neither (a) nor (b) holds. Then let $\succ$ be a connected, asymmetric, and transitive extension of Q^{TC} and define $(\Gamma_t)_{t\in\mathcal{T}}$ as follows:

$$\begin{aligned}\Gamma_t(A) =& \{x \in A : x = f_t(A''), \exists A'' \supseteq A\} \\ & \cup \{x \in A : x = f_{t''}(A''), \exists A'' \in \mathcal{D}_{t''}, \exists t'' < t\} \\ & \cup \{x \in A \backslash x' : y \succ x, \forall y \in A \backslash \{x, x'\}\}.\end{aligned}$$

This is identical to (3.6), apart from the final part. Following the proof of Proposition 3, under the assumption that neither (a) nor (b) holds, we have a GC(CF)-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$, while $x' \notin \Gamma_{t'}(A')$.

To show sufficiency of 3, take any GC(CF)-model $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}} \rangle$ that rationalizes $(f_t)_{t\in\mathcal{T}}$. By robust inference of preference in 1, $x' Q^{TC} f_{t'}(A'')$ implies $x' \succ f_{t'}(A'')$, which means that $x' \notin \Gamma_{t'}(A'')$ holds. Then, since $A'' \subseteq A'$, and since $\Gamma_{t'}$ is a CF, it follows that $x' \notin \Gamma_{t'}(A')$. Necessity of 3 is proved by showing the contrapositive. Suppose that $x' Q^{TC} f_{t'}(A'')$ holds for no $A'' \in \mathcal{D}_{t'}$ with $x' \in A'' \subseteq A'$. Now define a binary relation $\tilde{Q}$ as follows: for distinct elements $x, y \in X$, $x \tilde{Q} y$ if (i) $x Q^{TC} y$ or (ii) $y = x'$ and $\neg[x' Q^{TC} x]$. This binary relation ranks x' as low as possible, as long as it does not contradict Q^{TC}.

Lemma 1. *Binary relation $\tilde{Q}$ is acyclic.*

Proof of Lemma 1. Suppose by way of contradiction that $\tilde{Q}$ has a cycle. Since Q^{TC} is acyclic (or asymmetric) and transitive, this cycle must involve x'. Furthermore, it must be in the form: $x'Q^{TC}x\tilde{Q}x'$, and $x\tilde{Q}x'$ must be defined via (ii) in the construction of $\tilde{Q}$. However, this is impossible when $x'Q^{TC}x$, which shows that $\tilde{Q}$ is acyclic. *Lemma 1, Q.E.D.*

Now let $\succ$ be a connected, asymmetric, and transitive extension of $\tilde{Q}$, and define $(\Gamma_t)_{t\in\mathcal{T}}$ as (3.6).[9] Note that, following the proof of Proposition 3, $\langle \succ, (\Gamma_t)_{t\in\mathcal{T}}\rangle$ is a GC(CF)-model that rationalizes $(f_t)_{t\in\mathcal{T}}$. Additionally, define $(\tilde{\Gamma}_t)_{t\in\mathcal{T}}$ so that $\tilde{\Gamma}_t = \Gamma_t$ for every $t \neq t'$, and for t':

$$\tilde{\Gamma}_{t'}(A) = \begin{cases} \Gamma_{t'}(A) \cup \{x'\} & \text{if } x' \in A \subseteq A', \\ \Gamma_{t'}(A) & \text{otherwise.} \end{cases}$$

By construction, we have $x' \in \tilde{\Gamma}_{t'}(A')$. It remains to show that $\langle \succ, (\tilde{\Gamma}_t)_{t\in\mathcal{T}}\rangle$ is a GC(CF)-model that rationalizes $(f_t)_{t\in\mathcal{T}}$: that is, $(\tilde{\Gamma}_t)_{t\in\mathcal{T}}$ obeys growing consideration; $\tilde{\Gamma}_t$ is a CF for every $t \in \mathcal{T}$; and $f_t(A)$ is the $\succ$-best alternative in $\tilde{\Gamma}_t(A)$ for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$. Note that $(\Gamma_t)_{t\in\mathcal{T}}$ obeys growing consideration, and since $\Gamma_t(A) \subseteq \tilde{\Gamma}_t(A)$ for every (t, A), $(\tilde{\Gamma}_t)_{t\in\mathcal{T}}$ obeys growing consideration as well. Since Γ_t is a CF for every t, and $\tilde{\Gamma}_t = \Gamma_t$ for every $t \neq t'$, it follows that $\tilde{\Gamma}_t$ is a CF for every $t \neq t'$. Now focusing on period t', take any $A, A'' \subseteq X$ such that $A \subseteq A''$ and any $x \in \tilde{\Gamma}_{t'}(A'') \cap A$. If $x \in \Gamma_{t'}(A'')$, then since $\Gamma_{t'}$ is a CF, we have $x \in \Gamma_{t'}(A)$, and thus $x \in \tilde{\Gamma}_{t'}(A)$.

[9] Since $Q^{TC} \subseteq \tilde{Q}$, $\succ$ is an extension of Q^{TC} as well.

$A \in \mathcal{D}_1$	$\{x_1, x_2\}$	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$
$f_1(A)$	x_1	x_3	x_2
$A \in \mathcal{D}_2$		$\{x_2, x_3\}$	
$f_2(A)$		x_2	

Table 3.3: Choice function $(f_t)_{t \in \mathcal{T}}$ of Example 6.

If $x \notin \Gamma_{t'}(A'')$, this means that $x = x'$ and $x' \in A \subseteq A'' \subseteq A'$. Then, by construction of $\tilde{\Gamma}_{t'}$, it follows that $x \in \tilde{\Gamma}_{t'}(A)$. Finally, we show that $f_t(A)$ is the $\succ$-best alternative in $\tilde{\Gamma}_t(A)$ for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$. We already know that this holds at every $t \neq t'$, so fix period t', and consider any $A \in \mathcal{D}_{t'}$ and any $x \in \tilde{\Gamma}_{t'}(A) \backslash f_{t'}(A)$. If $x \in \Gamma_{t'}(A)$, then $f_{t'}(A) \succ x$ follows. If $x \notin \Gamma_{t'}(A)$, this means that $x = x'$ and $x' \in A \subseteq A'$. By assumption, we have $\neg[x' Q^{TC} f_{t'}(A)]$, and thus $f_{t'}(A) \tilde{Q} x'$ holds, which in turn implies $f_{t'}(A) \succ x$. Summarizing, $\langle \succ, (\tilde{\Gamma}_t)_{t \in \mathcal{T}} \rangle$ is a GC(CF)-model that rationalizes $(f_t)_{t \in \mathcal{T}}$, while $x' \in \tilde{\Gamma}_{t'}(A')$. □

The following example gives a choice function $(f_t)_{t \in \mathcal{T}}$ that is GC(CF)-rationalizable, and we demonstrate how robust inference can be conducted.

Example 6. *Let $X = \{x_1, x_2, x_3, x_4\}$ and $\mathcal{T} = \{1, 2\}$, and consider a choice function $(f_t)_{t \in \mathcal{T}}$ in Table 3.3. We first show that $(f_t)_{t \in \mathcal{T}}$ is GC(CF)-rationalizable. It follows that $x_1 Q x_2$ and $x_2 Q x_3$: we have $x_1 Q_1 x_2$ by (3.5); and $x_2 Q_2 x_3$ since x_2 is chosen from $\{x_2, x_3\}$ at $t = 2$ and x_3 is a chosen alternative at $t = 1$. Since Q is acyclic, $(f_t)_{t \in \mathcal{T}}$ is rationalizable by a GC(CF)-model. Regarding robust inference, we see that x_1 is robustly preferred to x_2, x_2 is robustly preferred to x_3, and x_1 is robustly preferred to x_3. Therefore, we can surely say that $x_1 \succ x_2 \succ x_3$. Moreover, x_2 is robustly considered at $(t' = 1, A' = \{x_1, x_2\})$ following 2-(a) in Proposition 4, and x_2, x_3*

$A \in \mathcal{D}_1$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\{x_1, x_2, x_3\}$
$f_1(A)$	x_2	x_3	x_2	x_2
$A \in \mathcal{D}_2$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\{x_1, x_2, x_3\}$
$f_2(A)$	x_1	x_3	x_2	x_2

Table 3.4: Choice function $(f_t)_{t \in \mathcal{T}}$ of Example 7.

are robustly considered at $(t' = 2, A' = \{x_2, x_3\})$ *following 2-(b). Finally, since* $x_i Q^{TC} x_3$ *for* $i \in \{1, 2\}$*, it holds that* x_1, x_2 *are robustly not considered at* $(t' = 1, A' = \{x_1, x_2, x_3\})$*.*

We show below an example where our robust inference of preference and the welfare criterion by Bernheim and Rangel (2007, 2009) lead to opposite welfare implications. The welfare criterion proposed by Bernheim and Rangel, which is referred to as an *unambiguous choice* relation, ranks alternative x'' over x', if x' is never chosen when x'' is available. Note that the choice function of Example 7 is defined on an exhaustive domain, so the discussion regarding Bernheim and Rangel's unambiguous choice relation is well-defined.

Example 7. *Let* $X = \{x_1, x_2, x_3\}$ *and* $\mathcal{T} = \{1, 2\}$*, and consider a choice function* $(f_t)_{t \in \mathcal{T}}$ *in Table 3.4. Note that we have* $x_1 Q x_2 Q x_3$*. Since* Q *is acyclic,* $(f_t)_{t \in \mathcal{T}}$ *is GC(CF)-rationalizable, and applying Proposition 4, we can robustly infer that* x_1 *is preferred to* x_3*. On the other hand, the unambiguous choice relation by Bernheim and Rangel concludes that* x_3 *is welfare-improving over* x_1*, since* x_1 *is never chosen when* x_3 *is available.*

3.4 Appendix: Growing consideration with attention filters

Here we derive a revealed preference test for decision making under growing consideration, where we require consideration mapping Γ_t to be an *attention filter* (henceforth, *AF*) for every $t \in \mathcal{T}$. Let us refer to such a GC-model $\langle \succ, (\Gamma_t)_{t \in \mathcal{T}} \rangle$ as a *growing consideration model with AF*, or in short a GC(AF)-model. Recall that Γ is an AF, if for every $A \subseteq X$ and $x \in A$, $x \notin \Gamma(A)$ implies $\Gamma(A) = \Gamma(A \backslash x)$. An equivalent expression of an AF is: for every $A', A'' \subseteq X$, $\Gamma(A'') \subseteq A' \subseteq A''$ implies $\Gamma(A') = \Gamma(A'')$. In words, an AF requires that removal of an ignored alternative does not alter the DM's consideration: if she ignores alternative x at set A, then the set $A \backslash x$ should be treated in the same way as A. A revealed preference characterization, as well as conditions for robust inference, for limited consideration model with AF (i.e., the limited attention model) are given by Masatlioglu, Nakajima, and Ozbay (2012). However, their approach is not directly applicable to ours, since they assume that the choice function is defined on an exhaustive domain.

A revealed preference test under the assumption that a choice function is defined on a non-exhaustive domain is given by De Clippel and Rozen (2021).[10] They derive a restriction that any limited-attention-consistent preference relation must obey, and then express a revealed preference test in terms of existence of a binary relation obeying that restriction. Here, we derive a revealed preference test for decision making under growing consideration with AF à la De Clippel and Rozen (2021). It is worth noting that

[10] Under the non-exhaustive domain assumption, a revealed preference test for the overwhelming choice model is essentially given by Dean, Kibris, and Masatlioglu (2017), and tests for a limited consideration model with both AF and CF and the (transitive) rational shortlisting method are given by Inoue and Shirai (Forthcoming).

the revealed preference test for the limited attention model provided by De Clippel and Rozen (2021) may be computationally challenging, since the test requires combinatorial search.[11] Moreover, under the non-exhaustive domain assumption on the observed data set, conditions for robust inference are yet to be discovered. This is partly due to sparsity of data, which makes it relatively difficult to make deterministic statements regarding behavior. In fact, the reason why their revealed preference test requires combinatorial search is because it is not possible to pin down a DM's consideration in general. Since robust inference under the limited consideration model with AF is a challenge even in the static case, here we focus on showing that there exists a revealed preference test for growing consideration with AF, and postpone the issue of robust inference for future research.

Now let us derive a revealed preference test for decision making under growing consideration with AF. Suppose that choice function $(f_t)_{t \in \mathcal{T}}$ is generated by a DM obeying a GC(AF)-model $\langle \succ, (\Gamma_t)_{t \in \mathcal{T}} \rangle$. To see observable restrictions of AF, take any $t \in \mathcal{T}$, and suppose that there exist feasible sets $A', A'' \in \mathcal{D}_t$ such that

$$f_t(A'), f_t(A'') \in A' \cap A'' \text{ and } f_t(A') \neq f_t(A''). \tag{3.7}$$

In this case, it must follow that

$$[y \in \Gamma_t(A') \text{ for some } y \in A' \backslash A''] \text{ or } [z \in \Gamma_t(A'') \text{ for some } z \in A'' \backslash A']. \tag{3.8}$$

To see this, suppose not. Then, we have $\Gamma_t(A') \subseteq (A' \cap A'') \subseteq A'$ and

[11]Nevertheless, De Clippel and Rozen (2021) propose a tractable method to conduct their revealed preference test, provided that the data set is not too "large": the test can be conducted by a method called *enumeration*.

$\Gamma_t(A'') \subseteq (A' \cap A'') \subseteq A''$. Since Γ_t is an AF, it must follow that $\Gamma_t(A') = \Gamma_t(A' \cap A'') = \Gamma_t(A'')$. However, this contradicts $f_t(A') \neq f_t(A'')$. Whenever (3.8) holds, this means that $[f_t(A') \succ y$ for some $y \in A' \backslash A'']$ or $[f_t(A'') \succ z$ for some $z \in A'' \backslash A']$. Therefore, any binary relation $\mathcal{P}$ that reflects the DM's preference must obey the following: for every $t \in \mathcal{T}$ and every $A', A'' \in \mathcal{D}_t$ such that (3.7) hold,

$$[f_t(A')\mathcal{P}y \text{ for some } y \in A' \backslash A''] \text{ or } [f_t(A'')\mathcal{P}z \text{ for some } z \in A'' \backslash A']. \quad (3.9)$$

In addition to this, there are observable restrictions of growing consideration. Any binary relation $\mathcal{P}$ that reflects the DM's preference must obey the following: for distinct $x', x'' \in X$, $x''\mathcal{P}x'$ if there exist $t', t'' \in \mathcal{T}$ with $t' < t''$, and $A' \in \mathcal{D}_{t'}$, $A'' \in \mathcal{D}_{t''}$ such that

$$f_{t'}(A') = x', f_{t''}(A'') = x'', \text{ and } x' \in A''. \quad (3.10)$$

Summarizing, we have a condition for rationalizablity by a growing consideration model with AF. Lemmas used in the proof of the following proposition are proved at the end of the Appendix.

Proposition 5. *A profile of choice functions $(f_t)_{t\in\mathcal{T}}$ is rationalizable by a GC(AF)-model, if and only if there exists an acyclic binary relation $\mathcal{P}$ that obeys (3.9) and (3.10).*

Proof. Since necessity in shown in the discussion preceding this proposition, we show sufficiency here by constructing a preference and profile of consideration mappings that constitute a GC(AF)-model that rationalizes $(f_t)_{t\in\mathcal{T}}$. The proofs of the lemmas used here are given after this proof is complete. Let $\succ$ be a connected, asymmetric, and transitive extension of $\mathcal{P}$. Then, for every $t \in \mathcal{T}$, define consideration mapping Γ_t as follows:

$$\text{if } A \in \mathcal{D}_t, \quad \Gamma_t(A) = \{f_t(A)\} \cup \{x \in A : f_t(A) \succ x\}; \tag{3.11}$$

$$\text{if } A \notin \mathcal{D}_t, \quad \Gamma_t(A) = \begin{cases} \Gamma_t(A'') & \text{if } \Gamma_t(A'') \subseteq A \subseteq A'', \exists A'' \in \mathcal{D}_t; \\ A & \text{otherwise.} \end{cases} \tag{3.12}$$

Lemma 2. *Consideration mapping Γ_t is well-defined for every $t \in \mathcal{T}$.*

Given that $(\Gamma_t)_{t\in\mathcal{T}}$ is a well-defined profile of consideration mappings, it remains to show that (i) $f_t(A)$ is the $\succ$-best alternative within $\Gamma_t(A)$ for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$; (ii) $(\Gamma_t)_{t\in\mathcal{T}}$ obeys growing consideration; and (iii) Γ_t is an AF for every $t \in \mathcal{T}$. By constriction of $(\Gamma_t)_{t\in\mathcal{T}}$, it follows immediately that $f_t(A)$ is the $\succ$-best alternative within $\Gamma_t(A)$ for every $A \in \mathcal{D}_t$ at every $t \in \mathcal{T}$.

We proceed to show that $(\Gamma_t)_{t\in\mathcal{T}}$ obeys growing consideration. Take any $t \geqslant 2$, $A \subseteq X$, and any $x' \in A$ such that $x' = f_{t'}(A')$ for some $t' < t$ and $A' \in \mathcal{D}_{t'}$. There are three cases to deal with in showing that $x' \in \Gamma_t(A)$. If $A \in \mathcal{D}_t$, then it must be the case that $f_t(A)\mathcal{P}x'$ or $f_t(A) = x'$, and by (3.11), they both imply $x' \in \Gamma_t(A)$. Now consider the case where $A \notin \mathcal{D}_t$ and $\Gamma_t(A'') \subseteq A \subseteq A''$ for some $A'' \in \mathcal{D}_t$. Note that we have $x' \in \Gamma_t(A'')$, as shown right above. Then, by construction of Γ_t, $x' \in \Gamma_t(A)$ follows. Finally, if $A \notin \mathcal{D}_t$ and $\Gamma_t(A'') \subseteq A \subseteq A''$ for no $A'' \in \mathcal{D}_t$, then since $\Gamma_t(A) = A$, we have $x' \in \Gamma_t(A)$. Summarizing, $(\Gamma_t)_{t\in\mathcal{T}}$ obeys growing consideration.

As the final step of the proof, we show that Γ_t is an AF for every $t \in \mathcal{T}$. Fix any $t \in \mathcal{T}$ and $A \subseteq X$ such that $x' \in A$ and $x' \notin \Gamma_t(A)$ hold. Denoting $A' = A \backslash x'$, it suffices to show that $\Gamma_t(A') = \Gamma_t(A)$. Note that $x' \notin \Gamma_t(A)$ implies either (a) $A \in \mathcal{D}_t$ or (b) $A \notin \mathcal{D}_t$ and $\Gamma_t(A'') \subseteq A \subseteq A''$ for some $A'' \in \mathcal{D}_t$, which are the two cases that we must deal with.

Case (a): $A \in \mathcal{D}_t$. Note that $x' \notin \Gamma_t(A)$ means that $x' \succ f_t(A)$ and that $\Gamma_t(A) \subseteq A' \subseteq A$. Therefore, if $A' \notin \mathcal{D}_t$, then by (3.12), it follows that $\Gamma_t(A') = \Gamma_t(A)$. When $A' \in \mathcal{D}_t$, we shall apply the following lemma.

Lemma 3. *Suppose that $B, B' \in \mathcal{D}_t$ obeys $\Gamma_t(B) \subseteq B' \subseteq B$. Then $f_t(B') = f_t(B)$, and moreover, we have $\Gamma_t(B') = \Gamma_t(B)$.*

Then, since $A, A' \in \mathcal{D}_t$ and $\Gamma_t(A) \subseteq A' \subseteq A$, it follows that $\Gamma_t(A') = \Gamma_t(A)$.

Case (b): $A \notin \mathcal{D}_t$ and $\Gamma_t(A'') \subseteq A \subseteq A''$ for some $A'' \in \mathcal{D}_t$. In this case, $\Gamma_t(A) = \Gamma_t(A'')$ holds. Since $x' \notin \Gamma_t(A)$, we have $x' \notin \Gamma_t(A'')$ as well. This in turn implies $\Gamma_t(A'') \subseteq A' \subseteq A''$. Then, note that $\Gamma_t(A'') = \Gamma_t(A')$ holds: this follows from Lemma 3 if $A' \in \mathcal{D}_t$; and from (3.12) if $A' \notin \mathcal{D}_t$. Hemce, $\Gamma_t(A') = \Gamma_t(A)$ holds. □

Proof of Lemma 2

Fixing any $t \in \mathcal{T}$, the substantial case that we must consider is when there exists $A \notin \mathcal{D}_t$ such that for some $A', A'' \in \mathcal{D}_t$, $[\Gamma_t(A') \subseteq A \subseteq A']$ and $[\Gamma_t(A'') \subseteq A \subseteq A'']$ hold. In this case, it suffices to show that $\Gamma_t(A') = \Gamma_t(A'')$. Note that we have $\Gamma_t(A') \subseteq (A' \cap A'')$ and $\Gamma_t(A'') \subseteq (A' \cap A'')$. Then by construction of Γ_t, it follows that $[y \succ f_t(A')$ for every $y \in A' \backslash A'']$ and $[z \succ f_t(A'')$ for every $z \in A'' \backslash A']$. This in turn implies that $[f_t(A')\mathcal{P}y$ for no $y \in A' \backslash A'']$ and $[f_t(A'')\mathcal{P}z$ for no $z \in A'' \backslash A']$. Since $\mathcal{P}$ obeys (3.9), it must be the case that $f_t(A') = f_t(A'') =: x''$. Then, $\Gamma_t(A') \subseteq A$ means that $\Gamma_t(A') = \{x''\} \cup \{x \in A : x'' \succ x\}$. Analogously, we have $\Gamma_t(A'') = \{x''\} \cup \{x \in A : x'' \succ x\}$, and thus it follows that $\Gamma_t(A') = \Gamma_t(A'')$. □

Proof of Lemma 3

Firstly, to see that $f_t(B') = f_t(B)$, suppose to the contrary. Note that we have $f_t(B), f_t(B') \in (B \cap B')$ and $B' \backslash B = \varnothing$. Moreover, $\Gamma_t(B) \subseteq B'$ implies that

$$x \succ f_t(B) \text{ for every } x \in B \backslash B'. \tag{3.13}$$

This in turn implies that $f_t(B)\mathcal{P}x$ holds for no $x \in B\backslash B'$. However, this contradicts that $\mathcal{P}$ obeys (3.9), and thus we conclude $f_t(B') = f_t(B)$. Letting $x'' := f_t(B') = f_t(B)$, since (3.13) holds, it follows that $\Gamma_t(B) = \{x''\} \cup \{x \in B : x'' \succ x\} = \{x''\} \cup \{x \in B' : x'' \succ x\} = \Gamma_t(B')$. □

Chapter 4

Rationalizing choice functions with a weak preference

This chapter develops revealed preference analysis of an individual choice model where a decision maker (DM) is a weak preference maximizer, under the assumption that a choice function, rather than a choice correspondence, is observed. In particular, we provide a revealed preference test for such model, and then provide conditions under which we can surely say whether some alternative is indifferent/weakly preferred/strictly preferred to another, solely from the information of the choice function. Furthermore, interpreting a choice correspondence as sets of potential candidates of alternatives that could be chosen from each feasible set, we analyze which alternatives must be, or cannot be a member of the choice correspondence: sharp lower and upper bounds of this underlying choice correspondence are given. As an assumption on observability of data, we assume that the choice function is defined on a non-exhaustive domain, so our results are applicable to data analysis even when only a limited data set is available.

In this chapter, we consider a fully-rational DM, but relax the classical model in two intuitive aspects. Firstly, we relax the common assumption that the DM has a strict preference, and consider a DM who has a weak preference. In the case where a DM has a weak preference, it is typically assumed in the literature that an economist observes the DM's choice correspondence. Letting X be a finite set of alternatives, and letting $\mathcal{D} \subseteq 2^X \backslash \varnothing$, a choice function is a mapping $F : \mathcal{D} \to 2^X \backslash \varnothing$. Given feasible set $A \in \mathcal{D}$, $F(A)$ is interpreted as the set of alternatives that *could* have been chosen from A. However, it is practically not possible to observe multiple choices simultaneously. Hence as a second departure from standard theory, we relax this "full-observation" assumption, and assume that only a choice function f is observed. Therefore, we consider a fully-rational DM with a weak preference, who chooses *one* of her most preferred alternatives from each feasible set. In particular, in this chapter we provide a necessary and sufficient condition under which a choice function f is consistent with a DM maximizing her weak preference: i.e., we can find a weak preference $\succsim$ such that for every feasible set $A \in \mathcal{D}$, $f(A) \succsim x$ for every $x \in A$.

In fact, without any additional constraint on the weak preference, a choice function is trivially consistent with weak preference maximizing behavior: any choice function is a result of maximizing behavior of a DM who is indifferent between all alternatives in X. Therefore, we restrict our attention to *non-degenerate* weak preferences, ones where there exist $x', x'' \in X$ with $x'' \succ x'$. Then, given a choice function that is consistent with weak preference maximization, we go one step further and provide conditions for welfare analysis. In particular, we provide conditions under which we can surely say that some alternative is indifferent/weakly preferred/strictly preferred to another, solely from the information given in the choice function. Moreover, we provide sharp lower and upper bounds of the "underlying" choice correspondence, namely, alternatives that could have been chosen from each feasible set. This allows us to make extrapolation over out-of-sample feasible sets.

Taking into account the practicality of our results, we assume that the domain of choice function f is not necessarily exhaustive: we allow $\mathcal{D}$, the domain of f, to be a strict subset of $2^X \backslash \varnothing$. There is a growing literature in choice theory that adopts this limited data assumption, which allows us to carry out empirical applications. Some papers that adopt this limited data assumption are Inoue and Shirai (Forthcoming) and De Clippel and Rozen (2021).

Organization of this chapter: In Section 4.1, we introduce our model and the concept of rationalizability. A necessary and sufficient condition for rationalizability is given in Section 4.1.2. Then, Section 4.2 is devoted to discussions regarding robust inference of DM's preference and underlying choice correspondence. In particular, in Section 4.2.1, we derive necessary and sufficient conditions under which we can surely say that some alternative is indifferent/weakly preferred/strictly preferred to another; and in Section 4.2.2 sharp lower and upper bounds of the underlying choice correspondence are given. We conclude the chapter by showing in Section 4.3 how our model relates with some of the models in the literature akin to ours. Proofs are contained in Appendix.

4.1 The model and rationalization condition

4.1.1 Preliminaries

Let X be a finite set of alternatives, and let $\mathcal{D} \subseteq 2^X \backslash \varnothing$ be a collection of feasible sets. A *weak preference*, denoted by $\succsim$, is a connected, reflexive, and transitive binary relation on X, and a *strict preference* is a connected, asymmetric, and transitive binary relation on X. A *choice function* is a mapping $f : \mathcal{D} \to X$ with $f(A) \in A$ for every $A \in \mathcal{D}$: that is, $f(A)$ is the chosen alternative from feasible set A. It is common in the literature that a choice function is associated with a DM maximizing a strict preference, while in models where DMs with weak preferences are considered, choice correspondences are assumed to be observed. In this chapter, we adopt a natural assumption that a DM has a weak preference, and the observationally practical assumption that an economist can observe only *one* choice made from each feasible set. Put otherwise, given any feasible set, while the DM's most preferred alternatives is a set in general, i.e., the DM has a choice correspondence, only a part of the underlying choice correspondence is observed. Given this assumption, we shall first address the following question: under what condition on f is it possible to interpret f as a result of weak preference maximizing behavior? A formal definition of this issue is given below.

Definition 5. A choice function f is *rationalizable by a weak preference (or weak preference rationalizable)*, if there exists a connected, reflexive, and transitive binary relation $\succsim$ on X such that for every $A \in \mathcal{D}$, $f(A) \succsim x$ for every $x \in A$.

It is worth noting that rationalizability of a choice function is vacuous if we cast no further restriction on weak preferences that rationalize f. That is, any choice function is rationalizable by a weak preference such that all the alternatives in X are indifferent. Therefore, in this chapter, we assume in addition that a DM has a *non-degenerate* weak preference, meaning that there exists a pair of alternatives where one alternative is strictly preferred to the other.

Definition 6. A weak preference $\succsim$ is *non-degenerate* if there exist $x', x'' \in X$ with $x'' \succ x'$.

Throughout this chapter, when we use the term "weak preference," let us implicitly assume that the weak preference is non-degenerate unless otherwise stated.

4.1.2 Rationalization condition

Here we derive a necessary and sufficient condition for f to be rationalizable by a weak preference. It is worth noting that there are two papers that refer to this issue. Nishimura, Ok, and Quah (2017) provide a general condition which is applicable to this model, and De Clippel and Rozen (2021) state that this issue is solvable by applying the *enumeration procedure* (see Nishimura, Ok, and Quah (2017) and De Clippel and Rozen (2021) for details). However, for completeness of the thesis, we explicitly derive a revealed preference condition here as well.

To begin with, let us assume that choice function f is generated by a DM maximizing her weak preference $\gtrsim$. Then it is natural to define a *revealed preference relation* R on X such that $x''Rx'$ if there exists $A \in \mathcal{D}$ with $x'' = f(A)$ and $x' \in A$. Note that whenever $x''Rx'$ holds, $x'' \gtrsim x'$ holds as well. Now let R^{TC} be the transitive closure of R, and define binary relation I as follows: $x''Ix'$ if (i) $x''R^{TC}x'$ and $x'R^{TC}x''$; or (ii) $x' = x''$. Then note that binary relation I is an equivalence relation (reflexive, symmetric, and transitive), and it provides equivalence classes of X. Let us denote by X/I the collection of equivalent classes with respect to I, and assume that X is partitioned into $K \in \mathbb{N}$ equivalent classes: $X/I = \{E_1, \ldots, E_K\}$. Then, since we have $x'' \sim x'$ for every $x', x'' \in E_k$ and every $k \in \{1, \ldots, K\}$, and since the DM's weak preference is non-degenerate, it must be the case that $K \geqslant 2$. In fact, this simple condition is not only necessary, but also sufficient for the rationalizability of a choice function by a weak preference.

A	$\{x_1, x_2, x_3, x_4\}$	$\{x_1, x_2, x_3\}$	$\{x_1, x_2\}$
$f(A)$	x_2	x_1	x_2

Table 4.1: Choice function of Example 8.

Proposition 6. *A choice function f is rationalizable by a weak preference, if and only if the equivalence classes of X with respect to binary relation I has more than or equal to 2 elements.*

Below we give an example of a choice function that is weak preference rationalizable, and show how we can test whether f is rationalizable or not. Choice functions in Examples 10 and 11 are ones that are not rationalizable by a weak preference.

Example 8. *Let $X = \{x_1, x_2, x_3, x_4\}$ and consider choice function f as in Table 4.1. We show that f is weak preference rationalizable. Note that $f(x_1, x_2, x_3, x_4) = x_2$ implies x_2Rx_1, x_2Rx_2, x_2Rx_3, x_2Rx_4; $f(x_1, x_2, x_3) = x_1$ implies x_1Rx_1, x_1Rx_2, x_1Rx_3; and $f(x_1, x_2) = x_2$ implies x_2Rx_1, x_2Rx_2. Then, $I = \{(x_1, x_1), (x_1, x_2), (x_2, x_1), (x_2, x_2), (x_3, x_3), (x_4, x_4)\}$, so we have $X/I = \{\{x_1, x_2\}, \{x_3\}, \{x_4\}\}$. Since $|X/I| = 3 \geqslant 2$, f is weak preference rationalizable. One example of a weak preference that rationalizes f is $x_1 \sim x_2 \succ x_3 \succ x_4$.*

4.2 Making robust inference

In this section, we consider how we can make robust inference about the DM's preference and underlying choice correspondence, given a choice function that is weak preference rationalizable. In particular, we derive conditions under which we can surely determine the relative preference ranking between two specific alternatives, using only the information of the choice function. In addition, we derive sharp lower and upper bounds of the underlying choice correspondence, which is the set of alternatives that the DM could have chosen. Since the bounds of choice correspondence are derived not only for sets in $\mathcal{D}$, but also for unobserved feasible sets, this allows extrapolation on what the DM may choose from feasible sets outside of the observed data set. Throughout this section, we assume that the choice function is rationalizable by a weak preference.

4.2.1 Robust inference of preference

Here we analyze how we can infer the DM's preference from a choice function. Even when a choice function is rationalizable by a weak preference, such weak preference is not uniquely determined in general. Meanwhile, it may still be possible to pin down the relative ranking between two alternatives. Below, we introduce the concept of robust inference of preference, and provide necessary and sufficient conditions for such robust inference.

Definition 7. Let choice function f be rationalizble by a weak preference. Then for $x', x'' \in X$, say that:

- *x'' and x' are robustly indifferent*, if $x'' \sim x'$ holds under every weak preference $\succsim$ that rationalizes f, and denote this by $x'' \sim^{\mathrm{r}} x'$;
- *x'' is robustly weakly preferred to x'*, if $x'' \succsim x'$ holds under every weak preference $\succsim$ that rationalizes f, and denote this by $x'' \succsim^{\mathrm{r}} x'$;
- *x'' is robustly strictly preferred to x'*, if $x'' \succ x'$ holds under every weak preference $\succsim$ that rationalizes f, and denote this by $x'' \succ^{\mathrm{r}} x'$.[1]

[1] In the case of robust inference of strict preference, it is implicitly assumed that $x' \neq x''$.

A	$\{x_1, x_2, x_3, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_1, x_2, x_5\}$	$\{x_2, x_5\}$
$f(A)$	x_1	x_3	x_2	x_5

Table 4.2: Choice function of Example 9.

The proposition below gives necessary and sufficient conditions for robust inference of preference. Intuitively, $x'' \sim^{\mathrm{r}} x'$ if and only if x', x'' are in the same equivalence class; $x'' \gtrsim^{\mathrm{r}} x'$ holds if and only if x'' is in a "weakly superior" equivalence class than that of x'; and $x'' \succ^{\mathrm{r}} x'$ holds if and only if any $\gtrsim$ that rationalizes f with $x' \gtrsim x''$ becomes degenerate.

Proposition 7. *Let choice function f be rationalizable by a weak preference. Then:*

1. *x' and x'' are robustly indifferent, if and only if $x' R^{TC} x''$ and $x'' R^{TC} x'$;*
2. *x'' is robustly weakly preferred to x', if and only if $x'' R^{TC} x'$;*
3. *x'' is robustly strictly preferred to x', if and only if $x'' R^{TC} x R^{TC} x'$ holds for every $x \in X$.*

In the following examples we show how robust inference of preference is done.

Example 8 (continued). *In this example, since $x_1 R x_2$ and $x_2 R x_1$, it follows that $x_1 \sim^r x_2$. Similarly, for $i \in \{1, 2\}$ and $j \in \{3, 4\}$, we have $x_i R^{TC} x_j$, so $x_i \gtrsim^r x_j$ follows. Therefore, while there are multiple weak preferences that rationalize f, any one of them must obey $x_1 \sim x_2$ and $x_1, x_2 \gtrsim x_3, x_4$.*

Example 9. *Let $X = \{x_1, x_2, x_3, x_4, x_5\}$ and consider choice function f as in Table 4.2. Note that we have $X/I = \{\{x_1, x_2, x_3, x_5\}, \{x_4\}\}$, where $x'' R^{TC} x_4$*

for every $x'' \in \{x_1, x_2, x_3, x_5\}$*. Note that for every* $x'' \in \{x_1, x_2, x_3, x_5\}$*, we have* $x'' R^{TC} x R^{TC} x_4$ *for every* $x \in X$*. Therefore, it must be the case that* $x'' \succ^r x_4$ *for every* $x'' \in \{x_1, x_2, x_3, x_5\}$*. In fact, the weak preference that rationalizes* f *is uniquely determined:* $x_1 \sim x_2 \sim x_3 \sim x_5 \succ x_4$*.*

4.2.2 Inference of underlying choice correspondence

In many papers, a weak preference is commonly associated with a choice correspondence. Given any feasible set A, a choice correspondence $F(A)$ is the set of maximum alternatives with respect to the DM's weak preference $\succsim$: $F(A) = \{x' \in A : x' \succsim x \text{ for every } x \in A\}$. In this chapter, we make a more practical assumption, and assume that the economist only has access to a choice function f. Here, for every feasible set A, $f(A)$ can be interpreted as *one* alternative of the "underlying" choice correspondence $F(A)$, where $F(A)$ is the set of alternatives that can potentially be chosen by the DM. Then, a natural question would be: can we make any inference about the underlying choice correspondence F, when we observe only a choice function f? In this section, this question is addressed, and provide sharp lower and upper bounds of the underlying choice correspondence.

Consider any observed choice function f that is rationalizable by a weak preference. Given a feasible set $A' \subseteq X$, let us infer what alternatives must/must not be in $F(A')$. First we consider what alternatives must be included in $F(A')$, in other words, we consider the lower bound. Using the results of Proposition 7, it is certain that if x' is robustly weakly preferred to every other alternative in A', then x' must be a member of $F(A')$. This holds because under any $\succsim$ that rationalizes f, $x' \succsim x$ holds for every $x \in A'$. In fact, this simple condition characterizes the lower bound for the underlying choice correspondence.

Lemma 4. *Let f be rationalizable by a weak preference. Then, given any $\succsim$ that rationalizes f and any $A' \subseteq X$, $x' \in A'$ must be in $F(A')$ if any only if $x' \succsim^r x$ for every $x \in A'$. Moreover, defining for every $A' \subseteq X$ the set $L(A')$ as below, $L(A')$ is the greatest lower bound of $F(A')$:*

$$L(A') = \{x' \in A' : x' \succsim^r x \text{ for every } x \in A'\}. \tag{4.1}$$

Deriving upper bound of $F(A')$ is a bit more elaborate. In doing this, we focus on the alternatives that cannot be a member of $F(A')$. First consider an alternative $x' \in A'$ where there exists $x'' \in A'$ such that x'' is robustly strictly preferred to x'. Then, under any $\succsim$ that rationalizes f, we have $x'' \succ x'$, and thus x' can never be in $F(A')$. There is another case where x' cannot be a member of $F(A')$. Note that $x' \in F(A')$ would mean that x' is weakly preferred to x for every $x \in A'$. If setting $x' \succsim x$ for every $x \in A'$ inevitably results in a degenerate $\succsim$, it is not possible for x' to be a member of $F(A')$. The discussion above is summarized in the lemma below.

Lemma 5. *Let f be rationalizable by a weak preference. Then, given any $\succsim$ that rationalizes f and any $A' \subseteq X$, $x' \in A'$ cannot be in $F(A')$ if and only if 1 and/or 2 below holds,*

1. *there exists $x'' \in A'$ with $x'' \succ^r x'$,*
2. (a) *$x'' R^{TC} x'$ for every $x'' \in A'$, and*

   (b) *$X = \bigcup_{x'' \in A'} \{x \in X : x'' R^{TC} x R^{TC} x'\}$.*

Moreover, defining for every $A' \subseteq X$ the set $U(A')$ as below, $U(A')$ is the least upper bound of $F(A')$:

$$U(A') = A' \backslash \{x' \in A' : x' \text{ obeys 1 or 2 above }\}. \tag{4.2}$$

Summarizing the lemmas above, we have the lower and upper bounds of the underlying choice correspondence.

Proposition 8. *Let f be rationalizable by a weak preference. Then for any $\succsim$ that rationalizes f, and for every $A \subseteq X$, we have*

$$L(A) \subseteq F(A) \subseteq U(A), \tag{4.3}$$

where $L(A)$ and $U(A)$ are defined as (4.1) and (4.2) respectively.

Note that Proposition 8 gives a lower and upper bound of the potentially chosen alternatives for all conceivable feasible sets $A \subseteq X$, rather than feasible sets from which choices are observed ($A \in \mathcal{D}$). Hence this result allows us to make predictions on what the DM may choose when confronting an out-of-sample feasible set. This may be useful in practice, since it is realistically not possible (in many cases) to observe choices from all conceivable feasible sets.

Example 8 (continued). *Here we show how inference of the underlying choice correspondence can be made. Let us focus on feasible set $\{x_1, x_2, x_3, x_4\}$. Then we have $L(x_1, x_2, x_3, x_4) = \{x_1, x_2\}$ and $U(x_1, x_2, x_3, x_4) = \{x_1, x_2, x_3, x_4\}$. Therefore, we can infer that $\{x_1, x_2\} \subseteq F(x_1, x_2, x_3, x_4) \subseteq \{x_1, x_2, x_3, x_4\}$.*

4.3 Relation with existing models

In this section, we relate the observable restrictions of our model with existing choice models. First we see how our model relates with standard rational choice models, and then compare observable restrictions of some closely related non-standard choice models.

4.3.1 Relation with standard rational choice models

Here we take a look at the observable restrictions of our model and standard rational choice models. It is well known that a choice *function* is consistent with maximization of a *strict* preference, if and only if it obeys the Strong Axiom of Revealed Preference (SARP), while a choice *correspondence* is consistent with maximization of a *weak* preference, if and only if it obeys the Congruence Axiom (CA). The formal definition of CA is:

Congruence Axiom: A choice correspondence $F : \mathcal{D} \to 2^X \backslash \varnothing$ obeys the Congruence Axiom (CA), if for every $A \in \mathcal{D}$ and $x', x'' \in A$, $x' \in F(A)$ and $x'' R^{TC} x'$ imply $x'' \in F(A)$.[2]

First of all, note that a strict preference is a special case of a weak preference, so whenever a choice function obeys SARP, it is rationalizable by a weak preference, i.e., SARP implies weak preference rationalizability. On the other hand, f can be weak preference rationalizable even when there is a cycle with respect to $\succ^R$, so the other direction does not hold.[3]

Note that CA is a condition on a choice correspondence, so it is not possible to directly compare with weak preference rationalizability. In fact, the only case where a choice correspondence F that obeys CA is weak preference rationalizable is when F turns out to be "single-valued," i.e., $|F(A)| = 1$ for every $A \in \mathcal{D}$. However, in this case, CA boils down to SARP. Meanwhile, it is possible to see how the underlying choice correspondence of f relates with CA. Suppose that choice function f is rationalizable by a weak preference $\succsim$. Then we have a choice correspondence F such that $F(A)$ is the

[2]See Richter (1966) for details of both the strong axiom of revealed preference and the congruence axiom.

[3]Choice functions in Examples 8 and 9 are weak preference rationalizable, but violate SARP.

set of $\succsim$-maximum alternatives for every $A \in \mathcal{D}$. This choice correspondence obviously satisfies CA. On the other hand, suppose that some choice correspondence obeys CA. Then $F(A)$ is the set of $\succsim$-maximum alternatives for every $A \in \mathcal{D}$, for some weak preference $\succsim$. Then, defining a choice function f so that $f(A) \in F(A)$ for every $A \in \mathcal{D}$, this choice function is weak preference rationalizable. Therefore, it seems plausible to regard weak preference rationalizability as a counterpart of CA, under the assumption that there is limitation of observability of choices: we cannot observe multiple simultaneous choices from a given feasible set.

4.3.2 Relation with some non-standard choice models

Here we show how our model relates with some non-standard choice models in the literature. In particular, we show that the limited consideration models and weak preference rationalizability are observationally independent, and we give an example of a choice function that is r-rationalizable (with $r = 2$) but not weak preference rationalizable.

To begin with, let us go through a brief summary of the models we deal with in this section. As introduced in Chapter 2, limited consideration models assume that some feasible alternatives are a priori excluded from DM's consideration, due to limitation of recognition capacity. That is, given a feasible set $A \in \mathcal{D}$, a DM maximizes her strict preference on her consideration set $\Gamma(A) \subseteq A$. Various limited consideration models differ depending on the structure of consideration mapping $\Gamma : 2^X \backslash \varnothing \rightarrow 2^X \backslash \varnothing$, a mapping that specifies consideration set for each feasible set $A \subseteq X$. Two well-established properties on the consideration mapping are the attention filter (AF) and competition filter (CF): we refer to a limited consideration model where Γ is

A	$\{x_1, x_2, x_3\}$	$\{x_1, x_2\}$	$\{x_2, x_3\}$
$f(A)$	x_1	x_2	x_3

Table 4.3: Choice function of Example 10.

an AF as an AF-model, and one where Γ is a CF as a CF-model.[4]

In fact, the rational shortlisting method in Manzini and Mariotti (2007), categorize-then-choose model in Manzini and Mariotti (2012), and the rationalization model in Cherepanov, Feddersen, and Sandroni (2013) are special cases of the CF-model, and the transitive rational shortlisting method in Au and Kawai (2011) is a special case of both AF and CF-models. Therefore, showing observational independence between AF and CF-models and weak preference rationalizability shows that weak preference rationalizability is observationally independent from many important bounded rationality models. It is shown in Example 10 that consistency with AF and CF-models does not imply weak preference rationalizability, and Examples 8 and 9 respectively show that weak preference rationalizability does not imply consistency with CF or AF-models. Thus, while limited consideration is a plausible model to explain cyclical choices, rationalizing cyclical choices with a weak preference may be a good alternative explanation of seemingly irrational behavior.

Example 10. *Let $X = \{x_1, x_2, x_3\}$ and consider choice function f as in Table 4.3. We first show that f is not rationalizable by a weak preference. It holds that $x_i I x_j$ for $i, j \in \{1, 2, 3\}$, and thus $X/I = \{\{x_1, x_2, x_3\}\}$. Since $|X/I| = 1$, f is not weak preference rationalizable. Now see that the preference $x_2 \succ x_1 \succ x_3$ and consideration mapping Γ as in Table 4.4 are consistent with both AF and CF-models. Therefore, choice function f is not rationalizable by a weak preference, but is rationalizable by AF and CF-models.*

[4]See Chapter 2 for details of limited consideration models.

A	$\{x_1, x_2, x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\Gamma(A)$	$\{x_1, x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_3\}$

Table 4.4: Consideration mapping Γ for Example 10.

Example 8 (continued). *We show that f is not consistent with the CF-model (recall that it is weak preference rationalizable). Note that under the CF-model, it must be the case that x_1 is strictly preferred to x_2, and x_2 is strictly preferred to x_1.[5] Thus there is no strict preference that can rationalize f, so f is not consistent with the CF-model.*

Example 9 (continued). *We show that f is not consistent with the AF-model (recall that it is weak preference rationalizable). Note that under the AF-model, it must be the case that x_1 is strictly preferred to x_2, and x_2 is strictly preferred to x_1.[6] Thus there is no strict preference that can rationalize f, so f is not consistent with the AF-model.*

Another model that is closely related to ours, is the *r-rationality* model in Barberà and Neme (2016). In this model, it is assumed that a DM has a strict preference, and given feasible set A, a DM chooses *one* of her r-best alternatives ($r \in \mathbb{N}$) within A, which is similar to our setting where we assume that an economist observes *one* of her most preferred alternatives within A. Note that any choice function is r-rationalizble when $r = |X|$, and r-rationalizability boils down to the rational choice model when $r = 1$. While it is obvious

[5] Recall that the definition of Γ being a CF is as follows: for every $A' \subset A''$ and $x \in A'$, $x \in \Gamma(A'')$ implies $x \in \Gamma(A')$. Thus it follows that x_2 is considered at $\{x_1, x_2, x_3\}$, which in turn implies x_1 is strictly preferred to x_2. Similarly, provided f follows the CF-model, it must be the case that x_1 attracts attention at $\{x_1, x_2, x_3, x_4\}$, which in turn implies that x_2 is strictly preferred to x_1.

[6] An expression of Γ being an AF that is equivalent to (2.1) is the following: $f(A) \neq f(A \backslash x)$ implies $x \in \Gamma(A)$. Taking a look at this, it follows from that x_2 attracts attention at $\{x_1, x_2, x_3, x_4\}$, which in turn implies x_1 is strictly preferred to x_2. Similarly, provided f follows the AF-model, it must be the case that x_1 attracts attention at $\{x_1, x_2, x_5\}$, which in turn implies that x_2 is strictly preferred to x_1.

that weak preference rationalizability implies $|X|$-rationalizability, and that 1-rationalizability implies weak preference rationalizability, there is no logical inclusion between weak preference rationality and r-rationalizability when $1 < r < |X|$. Example 11, which is an example presented in Barberà and Neme (2016), gives a choice function that is 2-rationalizable but not weak preference rationalizable. This means that weak preference rationalizability can be more restrictive than the most restrictive case of r-rationalizability (when $r > 1$).

As a technical difference, it is worth noting that Barberà and Neme (2016) assume that choices are observed on an exhaustive domain, namely $\mathcal{D} = 2^X \backslash \emptyset$. Therefore, when an economist only has access to a choice function defined on $\mathcal{D} \subsetneq 2^X \backslash \emptyset$, it is not possible to test r-rationalizability. On the other hand, weak preference rationalizability can be tested regardless of the domain of the choice function, which is when an economist observes choices made on only some subsets of X. This is in line with the observational assumption adopted in this chapter, which is that the economist cannot "fully" observe the DM's choices: only *one* of the alternatives that could have been chosen is observed, and only *some* of the logically possible feasible sets are observed.

Example 11. *Let* $X = \{x_1, x_2, x_3, x_4, x_5\}$ *and consider choice function* f *defined on* $2^X \backslash \emptyset$ *as below:*

- *for* $A \in 2^X \backslash \emptyset$ *such that* $|A| = 2$*,*
  - *if* $A = \{x_1, x_5\}$*,* $f(A) = x_5$*,*
  - *otherwise,* $f(x_i, x_j) = x_i$*, where* $i < j$*,*
- *for* $A \in 2^X \backslash \emptyset$ *such that* $|A| = 3$*,*

- *if* $x_3 \in A$*, then* $f(A) = x_3$*,*
- *if* $x_4 \in A$ *and* $x_3 \notin A$*, then* $f(A) = x_4$*,*
- $f(x_1, x_2, x_5) = x_2$*,*

- *for* $A \in 2^X \backslash \varnothing$ *such that* $|A| = 4$*,*

$$f(x_1, x_2, x_3, x_4) = x_4, \quad f(x_1, x_2, x_3, x_5) = x_3, \quad f(x_1, x_2, x_4, x_5) = x_4,$$
$$f(x_1, x_3, x_4, x_5) = x_4, \quad f(x_2, x_3, x_4, x_5) = x_3,$$

- $f(X) = x_3$*.*

We first show that f *is 2-rationalizable. Consider strict preference* $x_4 \succ x_3 \succ x_2 \succ x_1 \succ x_5$*. Then, at every feasible set* $A \in 2^X \backslash \varnothing$*,* $f(A)$ *is either her favorite, or 2nd favorite alternative, and thus* f *is 2-rationalizable by preference* $\succ$*. To show that* f *is not weak preference rationalizable, it suffices to see choices on* A *where* $|A| = 2$*: we have* $x_1 R x_2 R x_3 R x_4 R x_5 R x_1$*, which implies that* $X/I = \{\{x_1, x_2, x_3, x_4, x_5\}\}$*. Since* $|X/I| = 1$*,* f *is not weak preference rationalizable.*

4.4 Appendix

Proof of Proposition 6

Since necessity is already proved above, here we show that the other direction holds as well. Proofs of the lemmas used here are given after the proof of this proposition is complete. Suppose that $X/I = \{E_1, \ldots, E_K\}$ with $K \geqslant 2$. Let us define a binary relation $\rhd$ on X/I as follows: $E_j \rhd E_k$ if there exist $x'' \in E_j$ and $x' \in E_k$ with $x'' R^{TC} x'$ and not $x' R^{TC} x''$. For future reference, let us present the following lemma.

Lemma 6. *$E_j \rhd E_k, x'' \in E_j$, and $x' \in E_k$ implies $x'' R^{TC} x'$ and* **not** *$x' R^{TC} x''$.*

Using this binary relation $\rhd$, let us define binary relations $\mathcal{P}$ and $\mathcal{I}$ as follows:

- $x''\mathcal{P}x'$ if there exist j, k such that $x'' \in E_j, x' \in E_k$, and $E_j \rhd E_k$,
- $x''\mathcal{I}x'$ if there exist k such that $x', x'' \in E_k$.

Then, define binary relation $\mathcal{R}$ to be the union of $\mathcal{I}$ and $\mathcal{P}$, i.e., $\mathcal{R} = \mathcal{I} \cup \mathcal{P}$. This means that $\mathcal{I}$ and $\mathcal{P}$ are the symmetric and asymmetric components of $\mathcal{R}$ respectively. Note that by definition of $\mathcal{R}$ and Lemma 6, $x''\mathcal{R}x'$ implies $x'' R^{TC} x'$.

Lemma 7. *Binary relation $\mathcal{R}$ is consistent, that is, for every $x^1, \ldots, x^L \in X$,*

$$x^1 \mathcal{R} x^2 \mathcal{R} \cdots \mathcal{R} x^L \implies \textbf{not } x^L \mathcal{P} x^1. \tag{4.4}$$

Since binary relation $\mathcal{R}$ is consistent, by Suzumura's Extension Theorem, there exists a connected, reflexive, and transitive extension of $\mathcal{R}$, which we denote by $\succsim$. Under this extension, $x''\mathcal{P}x'$ implies $x'' \succ x'$, so this binary relation is non-degenerate. Now take any $A \in \mathcal{D}$. Note that we have $f(A)\mathcal{R}x$ for every $x \in A$, so $f(A) \succsim x$ holds for every $x \in A$. Thus, $\succsim$ rationalizes choice function f. This completes the proof of Proposition 6. □

Proof of Lemma 6

By definition of $\rhd$, $E_j \rhd E_k$ means that there exist $y^j \in E_j$ and $y^k \in E_k$ such that $y^j R^{TC} y^k$ and not $y^k R^{TC} y^j$. Meanwhile, $x'', y^j \in E_j$ means that $x'' R^{TC} y^j$ and $y^j R^{TC} x''$, and $x', y^k \in E_k$ means that $x' R^{TC} y^k$ and $y^k R^{TC} x'$. Since binary relation R^{TC} is transitive, $x'' R^{TC} x'$ follows. If we assume by way of contradiction that $x' R^{TC} x''$, then transitivity of R^{TC} implies that $y^k R^{TC} y^j$, which is a contradiction. □

Proof of Lemma 7

Note that by Lemma 6, we have $x^1 R^{TC} x^L$. By definition of $\mathcal{P}$, we have $x''\mathcal{P}x'$ if and only if there exist $E', E'' \in X/I$ with $x' \in E'$, $x'' \in E''$, and $E'' \rhd E'$, which in turn implies $x'' R^{TC} x'$ and not $x' R^{TC} x''$. Hence, since we have $x^1 R^{TC} x^L$, it is impossible to have $x^L \mathcal{P} x^1$. □

Proof of Proposition 7

Note that the proofs of the lemmas used here are given after the proof of this proposition is complete. First we prove sufficiency of 1. Take any weak preference $\succsim$ that rationalizes f. Then, for every $A \in \mathcal{D}$, it must be the case that $f(A) \succsim x$ for every $x \in A$, so whenever $x''Rx'$, then $x'' \succsim x'$ holds. Therefore, $x' R^{TC} x''$ and $x'' R^{TC} x'$ imply $x' \succsim x''$ and $x'' \succsim x'$, which in turn imply $x' \sim x''$. Necessity of 1 is proved by showing the contrapositive. Suppose that "$x' R^{TC} x''$ and $x'' R^{TC} x'$" does not hold. There are two essential cases to consider: (i) $x'' R^{TC} x'$ but not $x' R^{TC} x''$; and (ii) x', x'' are not related through R. Under case (i), we can follow the proof of Proposition 6, and we have a weak preference $\succsim$ that rationalizes f with $x'' \succ x'$. Under case (ii), we have $x' \in E'$, $x'' \in E''$ such that $E', E'' \in X/I$ are not related through $\rhd$.[7] Let us define binary relation $\rhd'$ on X/I such that $\rhd' = \rhd \cup \{(E'', E')\}$, and parallel to the proof of Proposition 6, define binary relations $\mathcal{P}', \mathcal{I}'$ on X using $\rhd'$.

Lemma 8. *Binary relation $\mathcal{R}' = \mathcal{I}' \cup \mathcal{P}'$ is consistent.*

Then, there is a completion of $\mathcal{R}'$, namely $\succsim$, that rationalizes f with $x'' \succ x'$. In both cases (i) and (ii), we have the desired result.

[7] Recall that binary relation $\rhd$ is defined in the proof of Proposition 6.

To prove sufficiency of 2, first suppose that $x'' R^{TC} x'$ holds: i.e., there exist $y^1, y^2, \ldots, y^L \in X$ such that $x'' = y^1 R y^2 R \cdots R y^L = x'$. This in turn means that for every $\ell \in \{1, \ldots, L-1\}$, there exists $A \in \mathcal{D}$ such that $y^\ell = f(A)$ and $y^{\ell+1} \in A$. Then, for any $\gtrsim$ that rationalizes f, it must be the case that $y^\ell \gtrsim y^{\ell+1}$ for every ℓ. By transitivity of $\gtrsim$, we have $x'' \gtrsim x'$. Necessity of 2 is proved by showing the contrapositive. Suppose that $x'' R^{TC} x'$ does not hold. This means that there exist $E', E'' \in X/I$ with $x' \in E'$, $x'' \in E''$, and $E' \neq E''$. There are two cases that we must consider: (i) E', E'' are not related through $\rhd$; and (ii) $E' \rhd^{TC} E''$.[8] For case (i), apply the proof of necessity of 1, and for case (ii), apply the proof in Proposition 6. In either case, we have $\gtrsim$ that rationalizes f with $x' \succ x''$.

To prove sufficiency of 3, suppose that $x'' R^{TC} x R^{TC} x'$ holds for every $x \in X$. By results in 2, under any non-degenerate weak preference $\gtrsim$ that rationalizes f, it must be the case that $x'' \gtrsim x \gtrsim x'$ holds for every $x \in X$. Now suppose by way of contradiction that $x'' \sim x'$ holds. Then, by assumption, it follows that $y' \sim y''$ for every $y', y'' \in X$, which contradicts that $\gtrsim$ is non-degenerate.

Proof of necessity of 3 will be done by showing the contrapositive. As in the previous cases, we construct a non-degenerate weak preference $\gtrsim$ that rationalizes f with $x' \gtrsim x''$, when $X \neq \{x : x'' R^{TC} x R^{TC} x'\}$. There are two major cases that we consider.

Case I: $x'' R^{TC} x'$ does not hold. Within this case, if $x' R^{TC} x''$, then by results in 2, we have $x' \gtrsim^r x''$, and $x' \gtrsim x''$ holds under any $\gtrsim$ that rationalizes f. If x', x'' are unrelated via R, then again applying the logic in the proof on necessity of 2, we have $\gtrsim$ that rationalizes f with $x' \succ x''$. This completes the proof for case I.

[8]Note that $\rhd^{TC}$ is the transitive closure of $\rhd$.

Case II: $x'' R^{TC} x'$, but there exists $x \in X$ that does not exhibit $x'' R^{TC} x R^{TC} x'$.

Prior to presenting a proof for this case, let us partition X into the following three sets:

$$Y = \{x \in X : x'' R^{TC} x R^{TC} x'\},$$
$$Y_1 = \{x \in X : x R^{TC} x' \text{ does not hold}\},$$
$$Y_2 = \{x \in X : x R^{TC} x' \text{ holds, but } x'' R^{TC} x \text{ does not hold}\}.$$

Note that in this case, $x', x'' \in Y$ and $Y_1 \cup Y_2 \neq \varnothing$. There are two subcases that we consider: (II-i) $Y_1 \neq \varnothing$; and (II-ii) $Y_2 \neq \varnothing$.

In case (II-i), take any $\bar{y} \in Y_1$ such that $\bar{y} \in \bar{E}$ for some $\bar{E} \in X/I$, where $\bar{E}$ is minimal with respect to $\rhd$, i.e., $\bar{E} \rhd^{TC} E$ for no $E \in X/I$. Such an $\bar{E}$ exists because X, and thus X/I, is finite. Now define binary relations $\mathcal{P}$, $\mathcal{I}_1$ and $\mathcal{I}_2$ as follows: $y'' \mathcal{P} y'$ if $y'' \in X \backslash \bar{E}$ and $y' \in \bar{E}$; $y'' \mathcal{I}_1 y'$ if $y', y'' \in X \backslash \bar{E}$; and $y'' \mathcal{I}_2 y'$ if $y', y'' \in \bar{E}$. Then, let $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ and $\mathcal{R} = \mathcal{I} \cup \mathcal{P}$.

Lemma 9. *Binary relations $\mathcal{I}$ and $\mathcal{P}$ are symmetric and asymmetric components of $\mathcal{R}$ respectively, and $\mathcal{R}$ is consistent.*

Then there is a connected, reflexive, and transitive extension $\succsim$ of $\mathcal{R}$, and the following lemma shows that this weak preference rationalizes f.

Lemma 10. *The weak preference $\succsim$ is non-degenerate, and $f(A) \succsim x$ holds for every $x \in A$ and every $A \in \mathcal{D}$.*

Finally, since $x', x'' \in X \backslash \bar{E}$, $x'' \mathcal{I} x'$ holds by construction of $\mathcal{I}$. This implies that we have $x'' \sim x'$, which completes the proof for case (II-i).

In case (II-ii), take any $\bar{y} \in Y_2$ such that $\bar{y} \in \bar{E}$ for some $\bar{E} \in X/I$, where $\bar{E}$ is maximal with respect to $\rhd$, i.e., $E \rhd^{TC} \bar{E}$ for no $E \in X/I$. Now define binary relations $\mathcal{P}$, $\mathcal{I}_1$ and $\mathcal{I}_2$ as follows: $y'' \mathcal{P} y'$ if $y'' \in \bar{E}$ and $y' \in X \backslash \bar{E}$;

$y''\mathcal{I}_1 y'$ if $y', y'' \in X \backslash \bar{E}$; and $y''\mathcal{I}_2 y'$ if $y', y'' \in \bar{E}$. Then, let $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ and $\mathcal{R} = \mathcal{I} \cup \mathcal{P}$.

Lemma 11. *Binary relations $\mathcal{I}$ and $\mathcal{P}$ are symmetric and asymmetric components of $\mathcal{R}$ respectively, and $\mathcal{R}$ is consistent.*

Then there is a connected, reflexive, and transitive extension $\gtrsim$ of $\mathcal{R}$, and the following lemma shows that this weak preference rationalizes f.

Lemma 12. *The weak preference $\gtrsim$ is non-degenerate, and $f(A) \gtrsim x$ holds for every $x \in A$ and every $A \in \mathcal{D}$.*

Finally, since $x', x'' \in X \backslash \bar{E}$, $x''\mathcal{I}x'$ holds by construction of $\mathcal{I}$. This implies that we have $x'' \sim x'$, which completes the proof for case (II-ii). □

Proof of Lemma 8

Take any $x^1, x^2, \ldots, x^L \in X$ such that $x^1\mathcal{R}'x^2\mathcal{R}' \cdots \mathcal{R}'x^L$. If there is no ℓ such that $x^\ell \in E'', x^{\ell+1} \in E'$, then applying the exact same logic in the proof of Lemma 7, $x^L\mathcal{P}'x^1$ cannot hold. The substantial case to consider is when there exist ℓ such that $x^\ell \in E''$, $x^{\ell+1} \in E'$, which implies $x^\ell\mathcal{P}'x^{\ell+1}$. In fact, such ℓ is unique, i.e., there cannot be $m \neq \ell$ such that $x^m, x^\ell \in E''$ and $x^{m+1}, x^{\ell+1} \in E'$.[9] Thus we have $x^1\mathcal{R}'x^2\mathcal{R}' \cdots \mathcal{R}'x^\ell\mathcal{P}'x^{\ell+1}\mathcal{R}' \cdots \mathcal{R}'x^L$, which implies $x^1R^{TC}x^\ell\mathcal{P}'x^{\ell+1}R^{TC}x^L$, where $x^1 \notin E'$ and $x^L \notin E''$. Suppose by way of contradiction that $x^L\mathcal{P}'x^1$. This means that $x^LR^{TC}x^1$ must follow, which in turn implies that $x^{\ell+1}R^{TC}x^\ell$, a contradiction. □

[9] Suppose to the contrary that such m exists: $x^1\mathcal{R}'x^2\mathcal{R}' \cdots \mathcal{R}'x^\ell\mathcal{P}'x^{\ell+1}\mathcal{R}' \cdots \mathcal{R}'x^m\mathcal{P}'x^{m+1}\mathcal{R}' \cdots \mathcal{R}'x^L$. Then $x^{\ell+1}R^{TC}x^m$ must follow, which contradicts that E', E'' are not related via $\rhd$.

Proof of Lemma 9

By construction, it is obvious that $\mathcal{I}$ is symmetric and that $\mathcal{P}$ is asymmetric. Now we show that it is not possible to have $x^1\mathcal{R}x^2\mathcal{R}\cdots\mathcal{R}x^L$ and $x^L\mathcal{P}x^1$ simultaneously. Note that $x^L\mathcal{P}x^1$ means that $x^L \in X\backslash\bar{E}$, and $x^1 \in \bar{E}$. Then it follows by definition of $\mathcal{I}$ that $x^1, \ldots, x^L \in \bar{E}$, which is a contradiction. $\square$

Proof of Lemma 10

Since $y'' \succ y'$ holds for $y'' \in X\backslash\bar{E}$ and $y' \in \bar{E}$, $\gtrsim$ is non-degenerate. Now take any $A \in \mathcal{D}$. To show that $f(A) \gtrsim x$ for every $x \in A$, it suffices to show that $f(A)\mathcal{R}x$ for every $x \in A$ Note that this holds whenever there does not exist $x \in A$ with $x\mathcal{P}f(A)$. Suppose by way of contradiction that $x\mathcal{P}f(A)$ holds for some $x \in A$. This means that $x \in X\backslash\bar{E}$ and $f(A) \in \bar{E}$. Meanwhile, we have $f(A)Rx$, so it follows that $\bar{E} \rhd E(x)$, where $E(x)$ is the equivalence class of x. This contradicts that $\bar{E}$ is chosen to be minimal with respect to $\rhd$. $\square$

Proof of Lemma 11

By construction, it is obvious that $\mathcal{I}$ is symmetric and that $\mathcal{P}$ is asymmetric. Now we show that it is not possible to have $x^1\mathcal{R}x^2\mathcal{R}\cdots\mathcal{R}x^L$ and $x^L\mathcal{P}x^1$ simultaneously. Note that $x^L\mathcal{P}x^1$ means that $x^L \in \bar{E}$ and $x^1 \in X\backslash\bar{E}$, and then it follows by definition of $\mathcal{I}$ that $x^1, \ldots, x^L \in X\backslash\bar{E}$, which is a contradiction.

$\square$

Proof of Lemma 12

Since $y'' \succ y'$ holds for $y'' \in \bar{E}$ and $y' \in X\backslash\bar{E}$, $\gtrsim$ is non-degenerate. Now take any $A \in \mathcal{D}$. To show that $f(A) \gtrsim x$ for every $x \in A$, it suffices to show that

$f(A)\mathcal{R}x$ for every $x \in A$. Note that this holds whenever there does not exist $x \in A$ with $x\mathcal{P}f(A)$. Suppose by way of contradiction that $x\mathcal{P}f(A)$ holds for some $x \in A$. This means that $x \in \bar{E}$ and $f(A) \in X \backslash \bar{E}$. Meanwhile, we have $f(A)Rx$, so it follows that $E(f(A)) \rhd \bar{E}$, where $E(f(A))$ is the equivalence class of $f(A)$. This contradicts that $\bar{E}$ is chosen to be maximal with respect to $\rhd$. □

Proof of Lemma 4

Take any $A' \subseteq X$, and suppose that for $x' \in A'$, we have $x' \succsim^{\text{r}} x$ for every $x \in A'$. This means that $x' \succsim x$ for every $x \in A'$, under any $\succsim$ that rationalizes f. Therefore, $x' \in F(A')$ must hold for every $\succsim$ that rationalizes f. Now we prove necessity by showing the contrapositive. Suppose that there exists $x'' \in A'$ such that $x' \succsim^{\text{r}} x''$ does not hold. Applying Proposition 7, this means that there exists $\succsim$ that rationalizes f with $x'' \succ x'$. Under this $\succsim$, we do not have $x' \in F(A')$. □

Proof of Lemma 5

Take any $A' \subseteq X$, and suppose first that 1 holds: there exists $x'' \in A'$ with $x'' \succ^{\text{r}} x'$. Then for every $\succsim$ that rationalizes f, we have $x'' \succ x'$. This in turn means $x' \notin F(A')$ for every $\succsim$ that rationalizes f. Suppose that 2 holds: (a) $x''R^{TC}x'$ for every $x'' \in A'$; and (b) $X = \bigcup_{x'' \in A'}\{x \in X : x''R^{TC}xR^{TC}x'\}$. Take any $\succsim$ that rationalizes f, and suppose by way of contradiction that $x' \in F(A')$. This means that $x' \succsim x''$ for every $x'' \in A'$. Meanwhile, $x''R^{TC}x'$ for every $x'' \in A'$ means that $x'' \succsim x'$ for every $x'' \in A'$. Moreover, 2-(b) means that for every $x \in X$, there exists $x'' \in A'$ such that $x'' \succsim x \succsim x'$. Summarizing, we have $y'' \sim y'$ for every $y', y'' \in X$, which contradicts that $\succsim$ is non-degenerate.

Necessity is proved through showing the contrapositive. Suppose that both 1 and 2 fail to hold. Then there are two cases to consider: case I is when 1 and 2-(a) fail; and case II is when 1 and 2-(b) fail. Here, for any $x \in X$, let us denote by $E(x)$ the equivalence class that x belongs to.

Case I: 1 and 2-(a) fail. In this case, there exists $x'' \in A'$ such that $x'' R^{TC} x'$ does not hold. Then, we have $E(x') \rhd^{TC} E(x'')$, or $E(x'), E(x'')$ are not related via $\rhd$. Let $\bar{E} = \{x \in X : E(x'') = E(x) \text{ or } E(x'') \rhd^{TC} E(x)\}$. Note that $f(A') \notin \bar{E}$.[10] Now define binary relations $\mathcal{P}, \mathcal{I}_1, \mathcal{I}_2$ as follows: $y'' \mathcal{P} y'$ if $y'' \in X \backslash \bar{E}$ and $y' \in \bar{E}$; $y'' \mathcal{I}_1 y'$ if $y', y'' \in X \backslash \bar{E}$; and $y'' \mathcal{I}_2 y'$ if $y', y'' \in \bar{E}$. Then let $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ and $\mathcal{R} = \mathcal{I} \cup \mathcal{P}$. By construction of $\mathcal{I}$ and $\mathcal{P}$, it is not possible to have $x^1 \mathcal{R} x^2 \mathcal{R} \cdots \mathcal{R} x^L$ and $x^L \mathcal{P} x^1$ simultaneously: $x^L \mathcal{P} x^1$ means that $x^L \in X \backslash \bar{E}$ and $x^1 \in \bar{E}$, and then $x^1, \ldots, x^L \in \bar{E}$ must hold, which in turn implies that $x^L \mathcal{P} x^1$ is not possible. Thus $\mathcal{R}$ is consistent, and there is a connected, reflexive, and transitive extension $\succsim$ of $\mathcal{R}$. Since $y'' \succ y'$ holds for $y'' \in X \backslash \bar{E}$ and $y' \in \bar{E}$, $\succsim$ is non-degenerate. Now take any $A \in \mathcal{D}$. Since $f(A) R x$ for every $x \in A$, we have $E(f(A)) = E(x)$ or $E(f(A)) \rhd^{TC} E(x)$. This implies that it is not possible to have $x \in A \cap (X \backslash \bar{E})$ and $f(A) \in \bar{E}$, and thus $x \mathcal{P} f(A)$ never holds for any $x \in A$. Therefore, we have $f(A) \succsim x$ for every $x \in A$. By construction of $\mathcal{R}$, we have $x' \mathcal{R} x$ for every $x \in A'$, so $x' \succsim x$ holds for every $x \in A'$.[11] This results in $x' \in F(A')$.

Case II: 1 and 2-(b) fail; 2-(a) holds. In this case, there exists $\bar{x} \notin \bigcup_{x'' \in A'} \{x \in X : x'' R^{TC} x R^{TC} x'\}$. Note that we have $\bar{x} \notin \bigcup_{x'' \in A'} E(x'')$, which in turn implies that $E(\bar{x}) \cap [\bigcup_{x'' \in A'} E(x'')] = \varnothing$.[12]

[10] Otherwise, since $f(A') R x'$, we have $x'' R^{TC} x'$, which is a contradiction.

[11] Note that $x' \notin \bar{E}$, so for every $x \in A'$, we have $x' \mathcal{P} x$ or $x' \mathcal{I}_1 x$.

[12] Otherwise, we have some $x'' \in A'$ with $x'' R^{TC} \bar{x}$ and $\bar{x} R^{TC} x''$, which in turn implies that $x'' R^{TC} \bar{x} R^{TC} x'' R^{TC} x'$. Thus $x'' R^{TC} \bar{x} R^{TC} x'$, which contradicts $\bar{x} \notin \bigcup_{x'' \in A'} \{x \in X : x'' R^{TC} x R^{TC} x'\}$.

Fact 6. *One of the following holds:*

(i). there exists $\bar{x} \notin \bigcup_{x'' \in A'} \{x \in X : x'' R^{TC} x R^{TC} x'\}$ *such that* $E \rhd^{TC} E(\bar{x})$ *for no* $E \in X/I$*;*

(ii). there exists $\bar{x} \notin \bigcup_{x'' \in A'} \{x \in X : x'' R^{TC} x R^{TC} x'\}$ *such that* $E(\bar{x}) \rhd^{TC} E$ *for no* $E \in X/I$*.*

Proof of Fact 6. Take any $\bar{E} \in X/I$ such that $E \rhd^{TC} \bar{E}$ holds for no $E \in X/I$. Finiteness of X and rationalizability of f assures the existence of such $\bar{E}$. If such $\bar{E}$ exhibits $\bar{E} \cap [\bigcup_{x'' \in A'} E(x'')] = \emptyset$, then (i) holds. Otherwise, $E \rhd^{TC} \bar{E}$ for no $E \in X/I$ would imply $\bar{E} \subseteq \bigcup_{x'' \in A'} E(x'')$, meaning that for any $\bar{E} \in X/I$ that is $\rhd$-maximal, there exists $x'' \in A'$ such that $\bar{E} = E(x'')$. To show that (ii) must hold in this case, suppose to the contrary that (ii) fails: for every $\bar{E}$ such that $\bar{E} \rhd^{TC} E$ for no $E \in X/I$, we have $\bar{E} \subseteq \bigcup_{x'' \in A'} E(x'')$.[13] This means that for every $\bar{E} \in X/I$ that is $\rhd$-minimal, there exists $x'' \in A'$ such that $\bar{E} = E(x'')$. Then, for every $E \in X/I$, it follows that "there exists $\hat{x} \in A'$ such that $E(\hat{x}) = E$ or $E(\hat{x}) \rhd^{TC} E$" and "there exists $\tilde{x} \in A'$ such that $E(\tilde{x}) = E$ or $E \rhd^{TC} E(\tilde{x})$." This in turn implies that $\bar{x} \in \bigcup_{x'' \in A'} \{x \in X : x'' R^{TC} x R^{TC} x'\}$ for every $\bar{x} \in X$, contradicting the assumption that 2-(b) fails. *Fact 6, Q.E.D.*

Now suppose that (i) in Fact 6 holds, and take any $\bar{x}$ as stated there. Then define binary relations $\mathcal{P}, \mathcal{I}_1$, and $\mathcal{I}_2$ as follows: $y'' \mathcal{P} y'$ if $y'' \in E(\bar{x})$ and $y' \in X \backslash E(\bar{x})$; $y'' \mathcal{I}_1 y'$ if $y', y'' \in X \backslash E(\bar{x})$; and $y'' \mathcal{I}_2 y'$ if $y', y'' \in E(\bar{x})$. Now let $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ and $\mathcal{R} = \mathcal{I} \cup \mathcal{P}$. By construction of $\mathcal{I}$ and $\mathcal{P}$, it is not possible to have $x^1 \mathcal{R} x^2 \mathcal{R} \cdots \mathcal{R} x^L$ and $x^L \mathcal{P} x^1$ simultaneously: $x^L \mathcal{P} x^1$ means that $x^L \in E(\bar{x})$ and $x^1 \in X \backslash E(\bar{x})$, and then $x^1, \ldots, x^L \in X \backslash E(\bar{x})$ must hold, which in turn implies that $x^L \mathcal{P} x^1$ is not possible. Thus $\mathcal{R}$ is consistent, and

[13] Note that this is an equivalent statement to the failure of (ii).

there is a connected, reflexive, and transitive extension $\succsim$ of $\mathcal{R}$. Since $y'' \succ y'$ holds for $y'' \in E(\bar{x})$ and $y' \in X\backslash E(\bar{x})$, $\succsim$ is non-degenerate. Now take any $A \in \mathcal{D}$. Since it is not possible to have $f(A) \notin E(\bar{x})$ and $x \in A \cap E(\bar{x})$, $x\mathcal{P}f(A)$ never holds for any $x \in A$. Therefore, we have $f(A) \succsim x$ for every $x \in A$. By construction of $\mathcal{R}$, we have $x'\mathcal{R}x$ for every $x \in A'$, so $x' \succsim x$ holds for every $x \in A'$. This results in $x' \in F(A')$.

Suppose that (ii) in Fact 6 holds, and take any $\bar{x}$ as stated there. Then define binary relations $\mathcal{P}, \mathcal{I}_1$, and $\mathcal{I}_2$ as follows: $y''\mathcal{P}y'$ if $y'' \in X\backslash E(\bar{x})$ and $y' \in E(\bar{x})$; $y''\mathcal{I}_1 y'$ if $y', y'' \in X\backslash E(\bar{x})$; and $y''\mathcal{I}_2 y'$ if $y', y'' \in E(\bar{x})$. Now let $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ and $\mathcal{R} = \mathcal{I} \cup \mathcal{P}$. By construction of $\mathcal{I}$ and $\mathcal{P}$, it is not possible to have $x^1\mathcal{R}x^2\mathcal{R}\cdots\mathcal{R}x^L$ and $x^L\mathcal{P}x^1$ simultaneously: $x^L\mathcal{P}x^1$ means that $x^L \in X\backslash E(\bar{x})$ and $x^1 \in E(\bar{x})$, and then $x^1, \ldots, x^L \in E(\bar{x})$ must hold, which in turn implies that $x^L\mathcal{P}x^1$ is not possible. Thus $\mathcal{R}$ is consistent, and there is a connected, reflexive, and transitive extension $\succsim$ of $\mathcal{R}$. Since $y'' \succ y'$ holds for $y'' \in X\backslash E(\bar{x})$ and $y' \in E(\bar{x})$, $\succsim$ is non-degenerate. Now take any $A \in \mathcal{D}$. Since it is not possible to have $f(A) \in E(\bar{x})$ and $x \in A \cap [X\backslash E(\bar{x})]$, $x\mathcal{P}f(A)$ never holds for any $x \in A$. Therefore, we have $f(A) \succsim x$ for every $x \in A$. By construction of $\mathcal{R}$, we have $x'\mathcal{R}x$ for every $x \in A'$, so $x' \succsim x$ holds for every $x \in A'$. This results in $x' \in F(A')$. □

Chapter 5

Concluding remarks

In this thesis, we have explored observable implications of various limited consideration models, observable implications of decision-making models with growing consideration, and observable implications of a weak-preference maximizing behavior when a choice function is observed. This chapter, which offers concluding remarks, summarizes the results derived in this thesis, as well as features of our analyses and some issues that are open for future research.

Dealing with cyclical choices

Providing a model that rationalizes cyclical choices is important because it is common to observe such behavior. All of the models we analyzed in thesis are motivated to provide an alternative explanation when the rational choice model cannot be used to explain the observed choices. In particular, limited consideration models, which are dealt with in Chapters 2 and 3, allow choices to be cyclic by assuming that a decision maker (DM) may overlook some alternatives that are available. Furthermore, in Chapter 3, we put forward a framework that adopts the behavioral assumption of limited consideration, and adds to it an assumption that the DM pays attention to alternatives chosen in the past. In fact, the model dealt in Chapter 3 is different from the models in the other chapters, in that we assume that an economist observes the DM's choices over multiple time periods.

Chapter 4 deals with cyclical choices from a different viewpoint. Unlike the other chapters, it is assumed that a DM is fully rational, and that the DM maximizes a weak preference rather than a strict preference. In addition, it is assumed that an economist can observe, from each feasible set, only one of her most preferred alternatives. Under this framework, there may be revealed preference cycles that consists of indifferent alternatives.

Applicability to actual choice data

In all three chapters (Chapters 2, 3, and 4), we adopt a realistic observational assumption that choices are observed from only *some* of conceivable feasible sets. This assumption is important for the following reasons. Considering Chapters 2 and 3, limited consideration models were created because DMs commonly exhibit irrational behavior in reality. Thus, revealed preference tests to check whether a DM's behavior is in line with these models should be applicable under a realistic assumption on the DM's choices. As for Chapter 4, we adopted a realistic assumption that only a choice function is observed, that is, it is not realistically possible to observe multiple choices made simultaneously. Therefore, it is natural to accept a realistic assumption that the choice function of the DM is defined on a non-exhaustive domain.

Since all our revealed preference tests, and the robust inference results that follow, are constructed under this observational assumption, all of the results are practically applicable to actual choice data.[1] This is partially why

[1] If we require an exhaustive choice function to be observed, choices must be observed from 1,013 feasible sets, which is not very realistic.

it was possible to conduct a simulation in Chapter 2. In fact, it is not difficult to run a similar simulation for the results in Chapters 3 and 4, which we postpone as an interesting future research. Furthermore, it is also possible to conduct an experiment to collect actual choice data. In fact, Dr. Koji Shirai and the author conducted an experiment at Waseda University and gathered choice data of 113 subjects. The experimental data was used to compute Selten's indices of each limited consideration model. The author decided to omit the experimental analysis from this thesis, since it was omitted from the published version of Inoue and Shirai (Forthcoming). This experimental choice data can be directly used to compute Selten's index of weak-preference rationalizability as well. An experiment for growing consideration model may be a bit more troublesome, since we must gather choice data of subjects over multiple time periods.

Robust inference and extrapolation

In Chapters 3 and 4, we provided robust inference conditions for preference/(non-)consideration and indifference/weak preference/strict preference respectively. Being able to robustly infer a DM's behavior is meaningful, since it is not generally possible to pin down the DM's preference (or consideration) even when her behavior is rationalizable by a specific model. A case where this information may be useful, is when an economist wants to make welfare assessments between multiple alternatives. Being able to pin down that a specific alternative is preferred to another, will be convenient in saying that one is more welfare-improving than another.

The robust inference of consideration/non-consideration in Chapter 3 allows us to pin-down what alternative was considered/ignored at some specific time and feasible set. This feasible set does not necessarily have to be a set from which a choice is observed. The upper and lower bounds of the underlying choice correspondence in Chapter4 tells us what alternatives are (not) candidates to be chosen from some feasible set. This feasible set need not be observed as well. That is, these robust inference results allow us to make extrapolation upon feasible sets that are yet to be observed, which may be profitable in analyzing a DM's behavior.

It is worth noting that the assumption of growing consideration in Chapter 3 is a very general one, and other plausible cases of "growing consideration." For example, we may consider a more general case of growing consideration where a DM pays attention to alternatives that were chosen in the last several time periods as opposed to all of the past periods. Instead, we may think of a more restrictive case where a DM pays attention to all alternatives that she considered in the past.[2] These cases are interesting variations that are open for future research. In the latter case, since the restriction of the model is stronger, sharper results of robust inference and extrapolation can be expected.

The requirement on the weak preference in Chapter 4, namely non-degeneracy, is a very general requirement as well, and still conditions of robust inference existed. Thus, by casting more mathematical structure in the consumption space, and casting stronger assumptions on a DM's preference (e.g., non-satiation, continuity, etc.), would lead to stronger results of robust inference, and sharper lower and upper bounds of the underlying choice correspondence. This case is another interesting extension of the results derived in this thesis.

[2]In the general case of growing consideration where we cast no intra-temporal assumption on consideration mapping, this case boils down to the same result as growing consideration assumption that we adopted in this thesis.

Chapter 6

Mathematical Appendix

Here we provide some mathematical concepts that are used throughout this thesis. In particular, we list some properties of binary relations, then introduce two extension theorems that are essential in understanding the proofs of the theorems/propositions presented in this thesis.[1]

Let X be a non-empty arbitrary set, and let $R \subseteq X \times X$ be a binary relation on X. In this thesis, we often abbreviate $(x'', x') \in R$ as $x''Rx'$. A binary relation R is *connected*, if for every distinct $x', x'' \in X$, we have $x'Rx''$ or $x''Rx'$. We say that R is *reflexive*, if for every $x \in X$, we have xRx; R is *symmetric*, if for every $x', x'' \in X$, $x''Rx'$ implies $x'Rx''$; and R is *asymmetric*, if for every $x', x'' \in X$, $x''Rx'$ implies that $x'Rx''$ does not hold. The *asymmetric part* P of binary relation R, is defined as the set $P = \{(x'', x') \in X \times X : x''Rx' \text{ and not } x'Rx''\}$. Similarly, the *symmetric part* I of binary relation R, is defined as $I = \{(x'', x') \in X \times X : x''Rx' \text{ and } x'Rx''\}$. Say that binary relation R is *transitive*, if for every $x, x', x'' \in X$, $x''Rx'$ and $x'Rx$ imply $x''Rx$. The *transitive closure* of binary relation R, which we

[1] A comprehensive summary of extension theorems are given in Andrikopoulos (2009). For details of Suzumura's Extension Theorem, see Suzumura (1976).

denote by R^{TC}, is defined as the smallest transitive relation that contains R: i.e., for every $x', x'' \in X$ with $x'' R^{TC} x'$, there exist $z^0, z^1, \ldots, z^K \in X$ such that $x'' = z^0$, $x' = z^K$, and $z^{k-1} R z^k$ for $k = 1, \ldots, K$. Binary relation R^{TC} is transitive by definition, and it follows that $R \subseteq R^{TC}$.

Now, we introduce two properties of binary relations that are directly required in the extension theorems. A binary relation R is *acyclic*, if for every $x^1, x^2, \ldots, x^K \in X$, $x^1 R x^2 R \cdots R x^K$ implies not $x^K R x^1$. We say that R is *consistent*, if for every $x^1, x^2, \ldots, x^K \in X$, $x^1 R x^2 R \cdots R x^K$ implies not $x^K P x^1$. Acyclicity and consistency are two very similar properties, but the former is a stronger property than the latter. For example, a weak preference $\succsim$ defined on X, which we assume to be connected, reflexive, and transitive, is consistent, but not acyclic: there may be a cycle $x^1 \succsim x^2 \succsim \cdots \succsim x^K \succsim x^1$, which implies that alternatives $x^1 x^2, \ldots, x^K$ are all indifferent. Meanwhile, it cannot be the case that x^k is strictly preferred to x^{k+1} is such a cycle, which means that consistency is satisfied. A strict preference $\succ$ on X, which is assumed to be connected, asymmetric, and transitive, is acyclic: $x^1 \succ x^2 \succ \cdots \succ x^K$ implies that $x^K \succ x^1$ does not hold.

Given below are two important extension theorems that are used in the proofs of our theorems/propositions. A binary relation $\bar{R}$ is an *extension* of a binary relation R, if $x'' R x'$ implies $x'' \bar{R} x'$ and $x'' P x'$ implies $x'' \bar{P} x'$, where $\bar{P}$ is the asymmetric part of $\bar{R}$. The first extension theorem, namely Szpilrajn's Extension Theorem, is used in Chapters 2 and 3, where we assume that a DM has a strict preference.

Szpilrajn's Extension Theorem. *Every acyclic binary relation R has an extension $\bar{R}$ that is connected, asymmetric, and transitive. Moreover, if x', x'' are non-comparable with respect to R, there is a connected, asymmetric, and transitive extension $\bar{R}_1$ with $x' \bar{R}_1 x''$ and a connected, asymmetric, and* *transitive extension $\bar{R}_2$ with $x''\bar{R}_2x'$.*

A common method used in revealed preference analyses of a choice model where a DM has a strict preference, is to (i) construct a binary relation based on observed choices and the structure of the model, (ii) show that it is acyclic, and (iii) regard the connected, asymmetric, and transitive extension of it, which is assured by Szpilrajn's Extension Theorem, as the DM's strict preference.

The extension theorem below, is used in Chapter 4, where we assume that a DM has a weak preference.

Suzumura's Extension Theorem. *A binary relation R has an extension $\bar{R}$ that is connected, reflexive, and transitive, if and only if R is consistent. Moreover, if x', x'' are non-comparable with respect to R, there is a connected, reflexive, and transitive extension $\bar{R}_1$ with $x'\bar{R}_1x''$ and a connected, reflexive, and transitive extension $\bar{R}_2$ with $x''\bar{R}_2x'$.*

Parallel to the case of strict preference, it is common to use Suzumura's Extension Theorem when conducting revealed preference analyses of choice model where a DM has a weak preference. In particular, in the proofs in Chapter 4, we (i) construct binary relations based on the DM's choices, (ii) show that the binary relation is acyclic, and (iii) regard the extension of it as DM's weak preference.

Bibliography

[1] Afriat, S.N. (1967): The construction of utility function from expenditure data. *International Economic Review*, 8 (1), 67-77.

[2] Allen, R. and J. Rehbeck: Revealed stochastic choice with attributes. *Economic Theory*, Forthcoming (`https://doi.org/10.1007/s00199-021-01405-1`) .

[3] Andrikopoulos, A. (2009): Szpilrajn-type theorems in economics. MPRA paper No. 14345.

[4] Au, P.H. and K. Kawai (2011): Sequentially rationalizable choice with transitive rationales. *Games and Economic Behavior*, **73**(2), 608-614.

[5] Barberà, S., and A. Neme (2016): Ordinal relative satisficing behavior: theory and experiments. Mimeo.

[6] Bernheim, B.D. and A. Rangel (2007): Toward choice-theoretic foundations for behavioral welfare economics. *American Economic Review: Papers and Proceedings*, 97(2), 464-470.

[7] Bernheim, B.D. and A. Rangel (2009): Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, 124(1), 51-104.

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

[9] Bronars, S.G. (1987): The power of nonparametric tests of preference maximization. *Econometrica*, **55**(3), 693-698.

[10] Cherepanov, V., T. Feddersen, and A. Sandroni (2013): Rationalization. *Theoretical Economics*, **8**(3), 775-800.

[11] Davis, M., G. Logemann, and D. Loveland (1962): A machine program for theorem proving. *Communications of the ACM*, **5**(7), 394-397.

[12] Dean, M., Ö. Kibris, and Y. Masatlioglu (2017): Limited attention and status quo bias. *Journal of Economic Theory*, **169**, 93-127.

[13] De Clippel, G., and K. Rozen (2021): Bounded rationality and limited datasets. *Theoretical Economics*, **16**(2), 359 - 380 (2021)

[14] Ferreira, J.V. and N. Gravel (2017): Choice with time. Aix-Marseille School of Economics, Working Papers, WP 2017-Nr 29.

[15] Golomb, S.W. and L.D. Baumert (1965): Backtrack Programming. *Journal of the ACM*, **12**(4), 516-524.

[16] Houthakker, H.S. (1950): Revealed preference and the utility function, *Economica*, 17 (66), 159-174.

[17] Inoue, Y. (2020a): Growing consideration. Waseda University WINPEC Working Paper Series No. E2003.

[18] Inoue, Y. (2020b): Rationalizing choice functions with a weak preference. Waseda University WINPEC Working Paper Series No. E2004.

[19] Inoue, Y. and K. Shirai: On the observable restrictions of limited consideration models: theory and application, *Economic Theory*, Forthcoming (`https://doi.org/10.1007/s00199-022-01418-4`).

[20] 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.

[21] Manzini, P. and M. Mariotti (2007): Sequentially rationalizable choice. *American Economic Review*, **97**(5), 1824-1839.

[22] Manzini, P. and M. Mariotti (2012): Categorize then choose: boundedly rational choice and welfare. *Journal of the European Economic Association*, **10**(5), 1141-1165.

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

[24] Masatlioglu, Y., D. Nakajima, and E.Y. Ozbay (2012): Revealed attention. *American Economic Review*, **102**(5), 2183-2205.

[25] Moulin, H. (1985): Choice functions over a finite set: a summary. *Social Choice and Welfare*, **2**, 147-160.

[26] Nishimura, H., E.A. Ok, and J.K.-H. Quah (2017): A comprehensive approach to revealed preference theory. *American Economic Review*, 107(4), 1239-1263.

[27] Richter, M.K. (1966): Revealed preference theory. *Econometrica*, 34(3), 635-645.

[28] Salant, Y. and A. Rubinstein (2008): (A, f): Choice with frames. *Review of Economic Studies*, 75, 1287-1297.

[29] Samuelson, P.A. (1938): The Empirical Implications of Utility Analysis, *Econometrica*, 6(4), 344-356.

[30] Schwartz, T. (1976): Choice functions, "rationality" conditions, and variations on the weak axiom of revealed preference. *Journal of Economic Theory*, **13**(3), 414-427.

[31] Sen, A.K. (1971): Choice functions and revealed preference. *Review of Economic Studies*, **38**(3), 307-317.

[32] Suzumura, K. (1976): Remarks on the theory of collective choice. *Economica*, 43(172), 381-390.

[33] Varian, H.R. (1982): The nonparametric approach to demand analysis, *Econometrica*, 50 (4), 945-972.

[34] Walker, R.S.: An enumerative technique for a class of combinatorial problems. In: Combinatorial Analysis (Proceedings of the Symposium on Applied Mathematics, vol. X), pp. 91-94, American Mathematical Society, Providence, RI (1960).