DISCUSSION PAPER SERIES

Discussion paper No. 176

Limited consideration and limited data: revealed preference tests and observable restrictions

Yuta Inoue
(Graduate School of Economics, Waseda University)
Koji Shirai
(School of Economics, Kwansei Gakuin University)

March 2018

MASTERY FOR SERVICE

SCHOOL OF ECONOMICS

KWANSEI GAKUIN UNIVERSITY

1-155 Uegahara Ichiban-cho
Nishinomiya 662-8501, Japan

Limited consideration and limited data: revealed preference tests and observable restrictions

Yuta Inoue*and Koji Shirai†

March 2, 2018

Abstract

This paper develops revealed preference tests for choices under limited consideration, allowing a partially observed data set. Our tests are based on a common structure of various limited consideration models, and cover leading theories in the literature including the limited attention model, the rationalization model, the categorize-then-choose model, and the rational shortlist model. While tests involve combinatorial calculation, by applying the backtracking method, we perform simulations to numerically compare observable restrictions of various models. As a result, we find remarkable differences in observable restrictions across models.

Keywords: Revealed preference; Limited consideration; Limited attention; Rational shortlisting; Bounded rationality; Bronars test

JEL Classification numbers: C6 D1 D8

* Graduate School of Economics, Waseda University. Email address: y.inoue@toki.waseda.jp

† School of Economics, Kwansei Gakuin University. Email address: kshirai1985@kwansei.ac.jp

For helpful discussions and comments, the authors are grateful to Georgios Gerasimou, Felix Kübler, Joshua Lanier, SangMok Lee, Paola Manzini, Marco Mariotti, John Quah, John Rehbeck, Xinxi Song, and Bruno Strulovici as well as conference participants at the 2016 NFS/NBER/CEME Conference on Mathematical Economics (Johns Hopkins), Royal Economic Society Meeting 2017 (Bristol), Asian Meeting of Econometric Society 2017 (Hong Kong), China Meeting of Econometric Society 2017 (Wuhan), SAET Conference 2017 (Faro), and Singapore Economic Reivew Conference 2017, and the seminar audiences in Kyoto and Kwansei Gakuin.

1 Introduction

Let X be a set that is interpreted as the set of alternatives, and let $A \subset X$ be a set of feasible alternatives for an agent. Following the classical choice theory, an agent will choose the most preferable alternative according to her preference which is often assumed to be complete, asymmetric, and transitive. In testing if an agent's behavior can be accounted for by this standard framework, the theory of revealed preference is one of the most prevailing methods for economists. Typically, we collect finitely many observations of an agent's behavior $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$, where $\mathcal{T}$ is the set of indices of observations, A^t is the set of feasible alternatives at observation t, and a^t is the chosen alternative from A^t. It is well known that a data set $\mathcal{O}$ is consistent with the standard choice framework, if and only if it obeys the *strong axiom of revealed preference (SARP)*, which requires acyclicity of the direct revealed preference relation $\succ^R$ defined as $x'' \succ^R x'$, if $x'' = a^t$ for some $t \in \mathcal{T}$, $x'' \neq x'$, and $x' \in A^t$.

However, as pointed out in a number of experimental studies, violation of SARP is not rare at all, and various theories of bounded rationality have been proposed for systematic analyses of cyclical choices. Amongst others, a number of studies investigate decision procedures where some feasible alternatives are a priori excluded from an agent's consideration. Namely, for a given feasible set A, an agent maximizes her preference relation not necessarily on A itself, but on some subset $\Gamma(A) \subset A$. For example, Masatlioglu, Nakajima, and Ozbay (2012) and Lleras, Masatlioglu, Nakajima, and Ozbay (2017) consider a situation where an agent is overwhelmed by the number of alternatives offered to her. In this case, due to the limitation of recognition capacity, she has to maximize her preference on a subset of the feasible set. As another example, Manzini and Mariotti (2007, 2012) and Cherepanov, Feddersen, and Sandroni (2013) establish shortlisting decision models. There, an agent has some criteria possibly different from her preference (e.g. psychological restrictions, a preference on categories rather than alternatives, and others), and she makes a sequential decision: an agent firstly makes a shortlist which is "optimal" in terms of her criteria, and then she chooses an alternative to maximize her preference relation. In this case, $\Gamma(A)$ can be interpreted as a shortlist derived in the first step.

The primal objective of this paper is to develop a theory for testing these models from a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$. More specifically, we provide a necessary and sufficient condition under which $\mathcal{O}$ is consistent with a model as follows: for every feasible set $A \subset X$, an agent maximizes some complete, asymmetric, and transitive preference $\succ$ on her *consideration set*

$\Gamma(A) \subset A$. We emphasize that the tests in this paper do not require "full observation" of a choice function. After constructing revealed preference characterizations of below referred models, we compare the observable restrictions of them by using simulation.

It is clear that, without any restriction on a set mapping Γ, such a model is vacuous in that any choice behavior is accounted for by letting $\{a^t\} = \Gamma(A^t)$ for every $t \in \mathcal{T}$. Thus, we deal with models where some restrictions *are* imposed on an agent's *consideration mapping* $\Gamma : 2^X \to 2^X$, which specifies her consideration set for every $A \subset X$. In particular, we start from looking at the following three restrictions: (1) the *attention filter property (AFP)*, which requires that for every $A', A'' \subset X$, $\Gamma(A'') \subset A' \subset A'' \Longrightarrow \Gamma(A') = \Gamma(A'')$; (2) the *competition filter property (CFP)*, which requires that for every $A' \subset A''$, $\Gamma(A'') \cap A' \subset \Gamma(A')$; and (3) the joint of AFP and CFP, which we denote by AFP+CFP.[1] Loosely speaking, AFP requires that the removal of unrecognized alternatives does not change the set of recognized alternatives, while CFP requires that every alternative recognized at a larger feasible set must be recognized at a smaller feasible set.

A number of important decision procedures are covered by the above listed restrictions on a consideration mapping. First of all, the *limited attention* model in Masatlioglu, Nakajima, and Ozbay (2012) is nothing but a preference maximization model on a consideration mapping with AFP. Second of all, the *order rationalization* model in Cherepanov, Feddersen, and Sandroni (2013) can be characterized as a preference maximization model with a consideration mapping satisfying CFP. In addition, the *categorize-then-choose* model by Manzini and Mariotti (2012) also derives a consideration mapping that obeys CFP.[2] This property is also used in the *limited consideration with status quo* model in Dean, Kibris, and Masatlioglu (2017). If one admits that both AFP and CFP are reasonable restrictions on a consideration mapping, then it seems natural to require both of them, or AFP+CFP on a consideration mapping. Indeed, as shown in Lleras, Masatlioglu, Nakajima and Ozbay (2015), many real-world examples actually support both AFP and CFP (e.g. paying attention to n most advertised commodities).[3]

What is not covered by the above three types of restrictions is the *rational shortlist method (RSM)* in Manzini and Mariotti (2007). There, an agent makes a shortlist as the set of maximal elements of an asymmetric first step preference, and then she makes a choice to

[1] The definition of AFP can be rewritten as $x \in A$ and $x \notin \Gamma(A) \Longrightarrow \Gamma(A \backslash x) = \Gamma(A)$.

[2] In the original setting in Manzini and Mariotti (2012), an agent's preference is assumed to be just complete and asymmetric. However, throughout this paper, we shall require that an agent has a strict preference.

[3] Lleras, Masatlioglu, Nakajima, and Ozbay (2015) is a working paper version of Lleras, Masatlioglu, Nakajima, and Ozbay (2017).

maximize her preference relation. Regarding shortlists as consideration sets, a consideration mapping must obey CFP, but it has stronger observable restrictions. Indeed, even if a data set is rationalizable by a limited consideration model with AFP+CFP, it may not be supported as a result of a rational shortlisting model. In addition, under the *transitive rational shortlist method (TRSM)* where a first step preference is asymmetric and transitive, a consideration mapping obeys AFP+CFP, but again, such a model has stronger observable restrictions than a limited consideration model with AFP+CFP.[4] In this paper, we also cover a revealed preference characterization of these rational shortlisting type models.

As summarized above, amongst five models dealt with in this paper, there are several subclass/superclass relations, while some models are logically independent with each other. By simulation, we can numerically compare relative strength of observable restrictions across these models. Following Bronars (1987), we generate random choices on a sequence of feasible sets and apply our tests to see the fraction of data that are consistent with each model. Moreover, provided that observable restriction of each model depends on the structure of feasible sets, we repeat the above procedure over randomly generated profiles of feasible sets. This type of simulation is useful to evaluate and compare models based on actual or experimental data by using Selten index (see Selten, 1991 and Beatty and Crawford, 2011). In our simulation, we stick to the environment with 20 feasible sets each of which contains 2 - 8 alternatives out of 10 alternatives, which seems easily implementable in experiments.

The result of simulation is rather striking in that strength of observable restriction is quite different across models. To be specific, AFP model is very hard to reject with average pass rate of random data exceeds 99%, and CFP model is also permissive with average pass rate exceeds 60%. However, the joint of them, or AFP+CFP model, is far more restrictive with average pass rate being less than 4%. Thus, the joint of rather weak behavioral restrictions could result in strong observable restrictions. The rational shortlisting type models both have strong testing power: the averate pass rate of RSM is less than 3% and that of TRSM is less than 0.1% .

From a technical perspective, our revealed preference tests involve combinatorial calculations, which could be a challenge for practical use. We deal with it by suggesting a set of testing algorithms employing the *backtracking* method and actually apply it to simulation.

[4]Similar to the case of the categorize-then-choose model, the original setting in Manzini and Mariotti (2007) does not require the transitivity, while Au and Kawai (2011), which firstly investigates the transitive rational shortlist model, does require the transitivity also on a second step preference relation.

In that sense, one may regard our simulation also as the implementation of our algorithm, with which even 10,000 sets of random data can be calculated in acceptable time by using unexceptional computers.

Connection with existing studies: It is standard in the literature of bounded rationality that models are characterized by using an exhaustive data set, or a choice function. For example, Masatlioglu, Nakajima, and Ozbay (2012) and Lleras, Masatlioglu, Nakajima, and Ozbay (2015, 2017) characterize AFP, CFP, and AFP+CFP in terms of a restriction on a choice function. Regarding rational shortlisting models, Manzini and Mariotti (2007) and Au and Kawai (2011) provide a choice function based characterization. That is, these papers consider a data set where a choice is observed under every logically possible feasible set. However, these results are not extendable to partially observed data sets $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$, where choices on some subsets of X may not be observed. Indeed, as shown in simulation part, these full-observation tests cannot even "approximate" necessary and sufficient conditions, particularly for AFP+CFP, RSM, and TRSM models.

An important issue concerning revealed preference analysis with partially observed data is the "extendability" problem pointed out by De Clippel and Rozen (2014). The essence of this problem is described by using AFP model as follows. Suppose that for a data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$, there exists a pair $(\Gamma^*, \succ^*)$ so that $a^t \succ^* x$ for all $x \in \Gamma^*(A^t)\backslash x$ with Γ^* obeying AFP on all observed feasible sets, i.e., $\Gamma(A^t) \subset A^s \subset \Gamma(A^t) \implies \Gamma(A^s) = \Gamma(A^t)$ for every $s, t \in \mathcal{T}$. However, this does *not* ensure the existence of $\Gamma : 2^X \to 2^X$ that obeys AFP for *all* (including unobserved) feasible sets $A \in 2^X$. The same issue also applies to other models in this paper. Following De Clippel and Rozen, our definition of rationalizability requires for a consideration mapping to obey specific properties (AFP, CFP etc.) on entire domain, rather than observed feasible sets.[5] They also provide revealed preference tests for AFP and CFP by using a different method from ours. Roughly speaking, their approach focuses on finding out an acyclic binary relation on X that is interpreted as a "candidate" of a part of preference inferred from data and models, while our approach focuses on exploring some model-based structure of revealed preference cycles. As we will see in Section 6, the difference of approaches could raise quite different performance in computation.

Another important contribution of De Clippel and Rozen's paper is to show that testing

[5]Naturally, one could consider the rationalizability by AFP model (or other models) with imposing AFP only on observed feasible sets, which is done by Tyson (2013).

AFP model is computationally tricky: indeed, it is shown to be NP hard. Despite that, this paper shows that, including AFP, our revealed preference tests reasonably work for "not too large" data. Thus, this paper complements and extends De Clippel and Rozen's paper by proposing an alternative approach for testing limited consideration models, and adding tests that are not covered by that paper (AFP+CFP, RSM, and TRSM). In addition, we provide a practical algorithm and numerical comparison of observable restrictions of models, which shows that models that are newly treated in this paper have remarkably stronger observable restrictions compared to AFP and CFP models.

Organization of the paper: In Section 2, we introduce limited consideration models that are dealt with in this paper. We provide our basic idea for testing models in Section 3, followed by the revealed preference tests for AFP, CFP, and AFP+CFP in Section 4 and for rational shortlisting type models in Section 5. Simulation results and technical issues concerning it are stated in Section 6.

2 Choices under limited consideration

Consider a single-agent decision problem where X is a finite set of alternatives, and $\succ$ is a complete, asymmetric, and transitive preference of an agent, which we refer to as a strict preference.[6] If an agent obeys the rational choice model, then for every feasible set $A \subset X$, she maximizes her strict preference $\succ$ on A.

On the other hand, motivated by evidences contradicting the rational choice theory, a number of alternative decision procedures are proposed in the literature of bounded rationality. There, either consciously or unconsciously, an agent makes a shortlist of alternatives before she chooses an alternative. That is, there exists a *consideration mapping* $\Gamma : 2^X \to 2^X$ such that $\Gamma(A) \subset A$ for every $A \subset X$, and an agent maximizes her strict preference on $\Gamma(A)$, rather than A itself. In what follows, given a consideration mapping Γ, $\Gamma(A)$ is referred to as a *consideration set* on A. Furthermore, in general, we refer to a pair of a consideration mapping and a strict preference $(\Gamma, \succ)$ as a *limited consideration* model.

In Masatlioglu, Nakajima, and Ozbay (2012), they consider a situation in which an agent cannot recognize all feasible alternatives due to limitation of recognition capacity. There,

[6]For every $x \in X$, $x \not\succ x$, and for every distinct $x, y \in X$, either $x \succ y$ or $y \succ x$ holds, and for every distinct $x, y, z \in X$, $x \succ y$ and $y \succ z$ imply $x \succ z$.

following psychological literature, a consideration mapping Γ is supposed to have the *attention filter property (AFP)* defined as: for every $A \subset X$ and $x \in A$,

$$x \notin \Gamma(A) \Longrightarrow \Gamma(A \backslash x) = \Gamma(A). \tag{1}$$

In words, the consideration set is not affected when unrecognized elements are removed from a feasible set. Alternatively, (1) is rewritten as: for every $A \subset X$ and $B \subset A$,

$$\Gamma(A) \subset A \backslash B \Longrightarrow \Gamma(A \backslash B) = \Gamma(A). \tag{2}$$

In what follows, we refer to such model $(\Gamma, \succ)$ as a limited consideration model with AFP. When there is no confusion, we may simply refer to such model as an *AFP model*.

As an alternative structure of a consideration mapping, Lleras, Masatlioglu, Nakajima, and Ozbay (2017) consider the following restriction: for every $A' \subset A''$ and $x \in A'$,

$$x \notin \Gamma(A') \Longrightarrow x \notin \Gamma(A''). \tag{3}$$

In words, if an alternative is not recognized in a smaller feasible set, then it cannot be recognized in a larger feasible set. This seems plausible if an agent has limited capacity of recognition. Equivalently, (3) can be written as: for every $A' \subset A''$,

$$\Gamma(A'') \cap A' \subset \Gamma(A'). \tag{4}$$

This condition is equivalent to the monotonicity of the set of unrecognized alternatives. We say that Γ obeys the *competition filter property (CFP)* if it obeys (4), and $(\Gamma, \succ)$ is referred to as a limited consideration model with CFP, or in short a *CFP model*.

It is known that a consideration mapping that obeys CFP can be generated by conscious shortlisting. In Cherepanov, Feddersen, and Sandroni (2013), they consider a situation in which an agent has some criteria on alternatives, other than her strict preference. Each criterion is referred to as a *rationale*, which may be a psychological restriction or may be a social norm. A set of rationales of an agent is denoted by $\{R^k\}_{k=1}^K$, each of which is assumed to be just a binary relation, so it may not be complete, asymmetric, or transitive. An alternative $x \in X$ is said to be supported on $A \subset X$, if there exists some rationale R^k such that xR^kx' for all $x' \in A \backslash x$. Then, an agent is supposed to eliminate all unsupported alternatives from a feasible set, that is, a consideration mapping is defined such that for every $A \subset X$,

$$\Gamma(A) = \{x \in A : \exists R^k \text{ such that } xR^k x' \text{ for all } x' \in A \backslash x\}. \tag{5}$$

A pair $(\Gamma, \succ)$ as an *order rationalization* model, if Γ is represented as (5) for some set of rationales $\{R^k\}_{k=1}^K$. Cherepanov, Feddersen, and Sandroni (2013) showed that a consideration mapping obeys CFP, if and only if it can be represented as (5).

In addition, a *categorize-then-choose* model in Manzini and Mariotti (2012) also derives a consideration mapping with CFP. In their model, an agent has a *shading relation* $\succ^S$, which is assumed to be asymmetric on 2^X. In the first step, an agent makes a shortlist such that for every $A \subset X$,

$$\Gamma(A) = \{x \in A : \nexists B', B'' \subset A \text{ such that } B'' \succ^S B' \text{ and } x \in B'\}. \tag{6}$$

Loosely speaking, an alternative in a dominated category is eliminated from candidates of her choice, and then, in the second step, an agent maximizes her strict preference $\succ$ on $\Gamma(A)$. It is known that a consideration mapping defined as (6) obeys CFP and vice versa.[7]

If we admit that both AFP and CFP are reasonable, then it is natural to consider the joint of AFP and CFP. Indeed, as pointed out in Lleras, Masatlioglu, Nakajima, and Ozbay (2015), both AFP and CFP are plausible in a number of real-world examples. For example, consider the situations in which an agent 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. All of these decision procedures derive consideration mappings satisfying both AFP and CFP. A pair $(\Gamma, \succ)$ is referred to as a limited consideration model with AFP+CFP, or an *AFP+CFP model* in short, if Γ obeys AFP+CFP.

Limited consideration models with CFP and those with AFP+CFP can be related to Manzini and Mariotti (2007)'s two-step decision procedure called a *rational shortlist method*. There, an agent has a preference relation for each step, say $\succ'$ and $\succ$, and for every $A \subset X$,

[7]If some alternative x is eliminated from consideration at some A', then, by definition, there exist some $B', B'' \subset A'$ such that $B'' \succ^S B'$ and $x \in B'$. Clearly, by considering the same pair of subsets B' and B'', x must be excluded from $\Gamma(A'')$ for any A'' with $A' \subset A''$. See, for the other direction, Manzini and Mariotti (2012) as well as Cherepanov, Feddersen, and Sandroni (2013).

an agent firstly makes a shortlist $\Gamma(A)$ such that

$$\Gamma(A) = \{x \in A : \nexists x' \in A \text{ such that } x' \succ' x\}, \tag{7}$$

and then, in the second step, an agent maximizes her second step preference relation $\succ$ on $\Gamma(A)$. In Manzini and Mariotti (2007), the first step preference $\succ'$ is just assumed to be acyclic, while Au and Kawai (2011) deal with the case where $\succ'$ is asymmetric and transitive.[8] We say that Γ obeys the *(transitive) rational shortlist method*, or in short, RSM (TRSM), if it can be described as (7) by using an acyclic (asymmetric and transitive) binary relation $\succ'$. By abuse of terminology, we refer to $(\Gamma, \succ)$ as an *RSM (TRSM) model*, if Γ obeys RSM (TRSM).

By letting $x''Rx' \iff x' \not\succ' x''$, Γ defined as (7) is a special case of that in (5), and hence, it must obey CFP. Moreover, one can confirm that if $\succ'$ is asymmetric and transitive, Γ defined in (7) also obeys AFP, i.e. it obeys AFP+CFP. To see this, suppose that $x \in A$ and $x \notin \Gamma(A)$. If $z \in \Gamma(A)$, there exists no $x' \in A$ such that $x' \succ' z$, and, in particular, there is no such x' in $A\backslash x$. Hence, it holds that $\Gamma(A) \subset \Gamma(A\backslash x)$. To see the converse set inclusion, suppose that $z \in \Gamma(A\backslash x)$, or there exists no $x' \in A\backslash x$ such that $x' \succ' z$. If $z \notin \Gamma(A)$ were to hold, it must be that $x \succ' z$. Since $x \notin \Gamma(A)$, there exists some $x' \in A\backslash x$ such that $x' \succ' x$. However, by transitivity, this implies that $x' \succ' z$, contradicting the assumption that $z \in \Gamma(A\backslash x)$. Hence, it holds that $z \in \Gamma(A)$, which, in turn, implies that $\Gamma(A\backslash x) \subset \Gamma(A)$. Thus, every (transitive) rational shortlisting model obeys CFP (AFP+CFP), but not vice versa; there exists a consideration mapping with CFP (AFP+CFP) that cannot be represented as (7) for any acyclic (asymmetric and transitive) binary relation on X.

3 A basic idea for revealed preference tests

Before proceeding to our main results, in this section, we put forward a general observation that holds for all limited consideration models. It is well known that the rational choice theory can be easily tested from agent's observed choice behavior. Let $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ be a finite set of observed choices, where $\mathcal{T} = \{1, 2, ..., T\}$ is the set of indices of observations, $A^t \subset X$ is the feasible set at observation t, and $a^t \in A^t$ is the chosen alternative at $t \in \mathcal{T}$. We use x to

[8] In Manzini and Mariotti (2007), they assumed that both $\succ'$ and $\succ$ are just asymmetric. However, since they also assume that the choice function is nonempty for all $A \subset X$, it is clear that $\succ'$ must be acyclic (otherwise $\Gamma(A)$ would be empty for some A).

represent a generic alternative in X. A key for testing the rational choice model is the *direct revealed preference* relation $\succ^R$ defined as $x'' \succ^R x'$, if $x'' = a^t$ for some $t \in \mathcal{T}$, $x'' \neq x'$, and $x' \in A^t$. In the case of the rational choice theory, motivation of this terminology is obvious. Indeed, if an agent follows the rational choice model and $x'' \succ^R x'$ for some $x'', x' \in X$, then $\succ^R$ must be contained in the agent's "true" preference $\succ$, and hence, $\succ^R$ cannot have a cycle, that is,

$$x^1 \succ^R x^2 \succ^R \cdots \succ^R x^K \Longrightarrow x^K \not\succ^R x^1. \tag{8}$$

Actually, the acyclicity of $\succ^R$, which is referred to as the *strong axiom of revealed preference (SARP)*, fully characterizes the observable restrictions from the rational choice model.

The aim of this paper is to develop counterparts of SARP for testing limited consideration models presented in the previous section. We define the notion of rationalizability as follows.

Definition 1. A data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ is rationalizable by a limited consideration model $\mathbf{M} \in$ {AFP, CFP, AFP+CFP, RSM, TRSM}, if there exists a pair $(\succ, \Gamma)$, where $\succ$ is a strict preference and $\Gamma : 2^X \to 2^X$, such that for every $A \subset X$, $\Gamma(A) \subset A$ and obeys the property $\mathbf{M}$ on 2^X, and $a^t \succ x$ for every $t \in \mathcal{T}$ and $x \in \Gamma(A^t) \backslash a^t$.

Note that every $\mathbf{M} \in$ {AFP, CFP, AFP+CFP, RSM, TRSM} excludes the trivial rationalization of letting $\Gamma(A^t) = \{a^t\}$ for every $t \in \mathcal{T}$, even if a data set itself is consistent with the model (we will see this by an example in the next section). In addition, we require for Γ to satisfy the corresponding property on entire domain 2^X rather than the set of observed feasible sets. To be precise, we review an example by De Clippel and Rozen (2014) below. For the data set in the example, we can find a pair $(\Gamma, \succ)$ so that for every $t \in \mathcal{T}$, $a^t \succ x$ for every $x \in \Gamma(A^t) \backslash a^t$ and Γ obeys AFP on observed feasible sets. However, it is *impossible* to find any Γ that obeys AFP *on the entire domain* with satisfying $a^t \succ x$ for every $t \in \mathcal{T}$ and $x \in \Gamma(A^t) \backslash a^t$.

Example 1. *Let $X = \{x_1, x_2, x_3, x_4, x_5\}$, and consider a data set of six observations as below. In this case, we can construct a pair $(\succ, \Gamma)$ such that $a^t \succ x$ for every $x \in \Gamma(A^t) \backslash a^t$ and*

t	1	2	3	4	5	6
A^t	$\{x_1, x_4\}$	$\{x_4, x_5\}$	$\{x_1, x_2, x_3\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$	$\{x_2, x_4, x_5\}$
a^t	x_4	x_5	x_3	x_1	x_2	x_4

$t \in \mathcal{T}$ with Γ obeying AFP only on observed feasible sets. For example, let $\Gamma(A^1) = \{x_1, x_4\}$, $\Gamma(A^2) = \{x_4, x_5\}$, $\Gamma(A^3) = \{x_3\}$, $\Gamma(A^4) = \{x_1, x_3\}$, $\Gamma(A^5) = \{x_2, x_3\}$, $\Gamma(A^6) = \{x_2, x_4\}$, and let $x_5 \succ x_4 \succ x_2 \succ x_1 \succ x_3$. Then, each a^t maximizes $\succ$ on A^t and there exists no pair of $s, t \in \mathcal{T}$ such that $\Gamma(A^t) \subset A^s \subset A^t$, and hence, AFP is trivially satisfied on observed feasible sets. However, we cannot extend Γ to 2^X with AFP being satisfied on the entire domain. Indeed, for a unobserved set $A = \{x_2, x_3\}$, (i) $\Gamma(A^3) \subset A \subset A^3$ requires that $\Gamma(A) = \Gamma(A^3) = \{x_3\}$, while (ii) $\Gamma(A^5) \subset A \subset A^5$ requires that $\Gamma(A) = \Gamma(A^5) = \{x_2, x_3\}$. In fact, not only this specification, there is no $(\succ, \Gamma)$ that rationalizes the data set with AFP being satisfied on 2^X. We will come back to this point in Section 4.1.

It is easy to check that the rational choice model is a special case of AFP, CFP, AFP+CFP, RSM, and TRSM respectively: by letting Γ be the identity mapping, it obeys all these properties. Our theory becomes substantial when $\mathcal{O}$ contains revealed preference cycles. Formally, a profile of alternatives $(x^k)_{k=1}^K$ is a *cycle* with respect to $\succ^R$, if for every $k \leqslant K, x^k \succ^R x^{k+1}$ and $x^K = x^1$. A cycle $(x^k)_{k=1}^K$ is *minimal*, if it contains no other cycle than itself. 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.

Assuming that a data set $\mathcal{O}$ has $Q(> 0)$ minimal cycles in total, then for $1 \leqslant q \leqslant Q$, the q-th minimal cycle is represented as $(x^{k_q})_{k_q=1}^{K_q}$. Consider a profile of points $\left[x^{(q)}\right]_{q=1}^{Q}$ such that each $x^{(q)}$ is chosen from q-th cycle. We refer to such profile as a *selection profile*, and each $x^{(q)}$ is referred to as a *selection point* of each q-th cycle. As explained below, what is essential in our theory is that we must choose an "appropriate" selection profile from revealed preference cycles. Although required properties on a selection profile vary across models, the following discussion applies to all $\mathbf{M} \in$ {AFP, CFP, AFP+CFP, RSM, TRSM}.

Suppose that a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ is collected from an agent obeying a specific limited consideration model. Since we assume that an agent's preference $\succ$ is asymmetric and transitive, for every cycle $(x^{k_q})_{k_q=1}^{K_q}$ there exists at least one $\bar{k}_q$ such that $x^{\bar{k}_q} \succ^R x^{\bar{k}_q+1}$, but $x^{\bar{k}_q+1} \succ x^{\bar{k}_q}$. Although such points may not be unique, for each q-th cycle we can arbitrarily fix one of them, and set that point as a selection point, i.e. $x^{(q)} = x^{\bar{k}_q}$. Denoting by $y^{(q)}$ the alternative that succeeds the selection point in q-th cycle, we have $x^{(q)} = x^{\bar{k}_q}$, $y^{(q)} = x^{\bar{k}_q+1}$, and $y^{(q)} \succ x^{(q)}$. Since an agent is supposed to maximize her preference on $\Gamma(A)$ for every $A \subset X$, if there exist some q and t such that $a^t = x^{(q)}$, we have $y^{(q)} \succ a^t$, which in turn implies

$y^{(q)} \notin \Gamma(A^t)$. Then we can define for every $t \in \mathcal{T}$,

$$B^t\left(\left[x^{(q)}\right]_{q=1}^{Q}\right) = \left\{y^{(q)} \in A^t : a^t = x^{(q)}\right\}, \tag{9}$$

which is empty if there exists no selection point $x^{(q)}$ such that $a^t = x^{(q)}$. Although it is important to note that each B^t depends on the choice of a selection profile, we omit the argument from (9) for the sake of notational simplicity (there seems to be no crucial danger of confusion).

This set B^t plays crucial roles in our revealed preference tests. Since $x \succ a^t$ holds for every $x \in B^t$, we have that $x \notin \Gamma(A^t)$. Put otherwise, for every $t \in \mathcal{T}$, it holds that $\Gamma(A^t) \subset A^t \backslash B^t$. In other words, if an agent obeys a limited consideration model, we can find a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ such that the corresponding $\{B^t\}_{t\in\mathcal{T}}$ obeys $\Gamma(A^t) \subset A^t \backslash B^t$ for every $t \in \mathcal{T}$. We set this as a proposition for future references.

Proposition 1. *Suppose that $\mathcal{O}$ is generated by an agent obeying a limited consideration model $(\Gamma, \succ)$. Then there exists a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ such that the corresponding $\{B^t\}_{t\in\mathcal{T}}$ obeys $\Gamma(A^t) \subset A^t \backslash B^t$ for every $t \in \mathcal{T}$. In particular, such a selection profile can be constructed by letting $x^{(q)} = x^{\bar{k}_q}$ with $x^{\bar{k}_q+1} \succ x^{\bar{k}_q}$ for each q-th cycle.*

Example 2. *Let $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, and consider a data set of five observations as below:*

t	1	2	3	4	5
A^t	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4, x_6\}$	$\{x_1, x_3, x_5, x_7\}$	$\{x_2, x_4, x_6\}$	$\{x_3, x_5, x_7\}$
a^t	x_1	x_2	x_3	x_4	x_5

There are four cycles with respect to the direct revealed preference $\succ^R$: $x_1 \succ^R x_2 \succ^R x_1$; $x_1 \succ^R x_3 \succ^R x_1$; $x_2 \succ^R x_4 \succ^R x_2$; and $x_3 \succ^R x_5 \succ^R x_3$, and let us number the cycles in this order.[9] *Consider the case where we choose $[x_1, x_1, x_2, x_3]$ as a selection profile, i.e. $x^{(1)} = x_1$, $x^{(2)} = x_2$, $x^{(3)} = x_2$, and $x^{(4)} = x_3$. Then, in the first cycle we have $x^{(1)} = x_1$ and $y^{(1)} = x_2$. Since we have $a^1 = x^{(1)}$, it follows that $y^{(1)} = x_2 \in B^1$. Following a similar procedure, we have $\{B^t\}_{t\in\mathcal{T}}$ as follows.*

[9] Recall that we regard $x_2 \succ^R x_1 \succ^R x_2$ as the same with $x_1 \succ^R x_2 \succ^R x_1$ and we ignore any non-minimal cycle such as $x_1 \succ^R x_2 \succ^R x_4 \succ^R x_2 \succ^R x_1$.

t	1	2	3	4	5
B^t	$\{x_2, x_3\}$	$\{x_4\}$	$\{x_5\}$	$\varnothing$	$\varnothing$
$A^t \backslash B^t$	$\{x_1\}$	$\{x_1, x_2, x_6\}$	$\{x_1, x_3, x_7\}$	$\{x_2, x_4, x_6\}$	$\{x_3, x_5, x_7\}$

Table 1: The sets $\{B^t\}_{t\in\mathcal{T}}$ as a result of selection profile $[x_1, x_1, x_2, x_3]$.

4 Testing AFP, CFP, and AFP+CFP

4.1 AFP condition

Suppose that $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ is a data set collected from an agent obeying an AFP model, i.e. a limited consideration model $(\Gamma, \succ)$, where Γ obeys AFP defined in (2) on entire 2^X. With no loss of generality, we may assume that $\mathcal{O}$ contains cycles with respect to $\succ^R$. By choosing a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ as specified in Proposition 1, the corresponding $\{B^t\}_{t\in\mathcal{T}}$ satisfies $\Gamma(A^t) \subset A^t \backslash B^t$ for every $t \in \mathcal{T}$. Besides, the AFP model casts further restrictions; given $\Gamma(A^t) \subset A^t \backslash B^t$ for every $t \in \mathcal{T}$, since Γ must obey AFP, it holds that, for every $t \in \mathcal{T}$,

$$(A^t \backslash B^t) \subset A \subset A^t \Longrightarrow \Gamma(A) = \Gamma(A^t). \tag{10}$$

The above derives the following important restriction. If $(A^s \backslash B^s) \cup (A^t \backslash B^t) \subset (A^s \cap A^t)$ holds, then, by letting $A = (A^s \backslash B^s) \cup (A^t \backslash B^t)$, the LHS of (10) is satisfied both for s and t. As a result, it must hold that $\Gamma(A^s) = \Gamma((A^s \backslash B^s) \cup (A^t \backslash B^t)) = \Gamma(A^t)$, which implies that $a^s = a^t$. In fact, this property, which is summarized as the condition below, characterizes a data set that is rationalizable by an AFP model.

AFP CONDITION: A selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeys AFP condition, if for every $s, t \in \mathcal{T}$,

$$(A^s \backslash B^s) \cup \left(A^t \backslash B^t\right) \subset (A^s \cap A^t) \Longrightarrow a^s = a^t. \tag{11}$$

Recall that, as seen from its definition (9), B^t depends on the choice of a selection profile. It is already clear that for a data set to be rationalizable by an AFP model, there must exist a selection profile obeying AFP condition. Our more substantial claim is the converse: if there exists a selection profile obeying AFP condition, then the agent's behavior can be accounted for by an AFP model.

THEOREM 1. *A data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ is rationalizable by a limited consideration model* *with AFP, if and only if there exists a selection profile* $\left[x^{(q)}\right]_{q=1}^{Q}$ *obeying AFP condition.*

The procedure for the proof of sufficiency is as follows. Given a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeying AFP condition, we explicitly construct a consideration mapping that obeys AFP. Then by using it, a strict preference that rationalizes the data set is also constructed. Specifically, for $\{B^t\}_{t\in\mathcal{T}}$ corresponding to $\left[x^{(q)}\right]_{q=1}^{Q}$, we simply define a consideration mapping Γ such that for every $A \subset X$,

$$\begin{aligned}\Gamma(A) &= A\backslash B^t, \quad \text{if there exists some } t \in \mathcal{T} \text{ such that } A^t\backslash B^t \subset A \subset A^t \\ &= A \quad \text{otherwise.}\end{aligned} \tag{12}$$

In general, for a given $A \subset X$, there may be multiple observations that satisfy the condition in (12), i.e. for some $s, t \in \mathcal{T}$, $A^t\backslash B^t \subset A \subset A^t$ and $A^s\backslash B^s \subset A \subset A^s$. However, in that case, $A\backslash B^t = A\backslash B^s$ must hold, and hence, the above construction of Γ is well-defined, which is proved in Appendix I.

Lemma 1. *Suppose that for some* $s, t \in \mathcal{T}$, $A^t\backslash B^t \subset A \subset A^t$ *and* $A^s\backslash B^s \subset A \subset A^s$. *Then, it holds that* $A\backslash B^t = A\backslash B^s$.

Based on Γ defined as (12), the proof essentially completes with the help of the following two lemmas that are proved in Appendix I.

Lemma 2. *The consideration mapping* Γ *defined as (12) obeys AFP.*

Lemma 3. *Let* $\succ^*$ *be a binary relation such that* $x'' \succ^* x'$, *if* $x'' = a^t$ *for some* $t \in \mathcal{T}$, $x' \in \Gamma(A^t)$, *and* $x'' \neq x'$. *Then,* $\succ^*$ *is acyclic and for every* $t \in \mathcal{T}$, $a^t \in \Gamma(A^t)$.

The rest of the proof is somewhat routine work: by Lemma 3, the transitive closure of $\succ^*$ is an asymmetric and transitive ordering, and hence, by Szpilrajn's theorem, it can be extended to a strict preference $\succ$ on X. In addition, again by Lemma 3, it holds that for every $t \in \mathcal{T}$, $a^t \succ x$ for every $x \in \Gamma(A^t)\backslash a^t$. Then, together with Lemma 2, the data set is rationalizable by the limited consideration model $(\Gamma, \succ)$, where Γ obeys AFP.

Example 1 (continued). *Consider the data set in Example 1. In this data set, we can find seven minimal cycles, but here it suffices to focus on the following two cycles:* $x_1 \succ^R x_3 \succ^R x_1$ *and* $x_2 \succ^R x_4 \succ^R x_2$, *which we refer to as the first and second cycles. We show that for* *any combination of selection points of these two cycles, the corresponding selection profile* $\left[x^{(q)}\right]_{q=1}^{7}$ *cannot satisfy AFP condition. First we show that* $x^{(1)} = x_1$ *leads a violation of (11). If* $x^{(1)} = x_1$*, then regardless of selection points from other cycles,* $x_3 \in B^4$*. This, in turn, implies that* $(A^1 \backslash B^1) \cup (A^4 \backslash B^4) \subset \{x_1, x_4\} = (A^1 \cap A^4)$*, but* $x_4 = a^1 \neq a^4 = x_1$*, a violation of AFP condition. Similarly, any selection profile with* $x^{(2)} = x_4$ *fails to satisfy (11) at observations 2 and 6. Finally, we consider a selection profile with* $x^{(1)} = x_3$ *and* $x^{(2)} = x_2$*. This implies that* $x_1 \in B^3$ *and* $x_4 \in B^5$*, and hence,* $(A^3 \backslash B^3) \cup (A^5 \backslash B^5) \subset \{x_2, x_3\} = A^3 \cap A^5$*. However,* $x_3 = a^3 \neq a^5 = x_2$*, or the violation of (11). As a result, we cannot find any selection profile obeying (11), or equivalently, the data set in question is not rationalizable by AFP model.*

Example 2 (continued). *Consider the data set in Example 2, and recall that there are four cycles with respect to the direct revealed preference* $\succ^R$*:* $x_1 \succ^R x_2 \succ^R x_1$*;* $x_1 \succ^R x_3 \succ^R x_1$*;* $x_2 \succ^R x_4 \succ^R x_2$*;* $x_3 \succ^R x_5 \succ^R x_3$*. We show that the selection profile* $[x_1, x_1, x_2, x_3]$ *actually succeeds in satisfying AFP condition. In fact, the relevant sets:* $\{(B^t, A^t \backslash B^t)\}_{t \in \mathcal{T}}$ *are summarized in Table 1. Looking at Table 1, one can confirm that there is no pair* $s, t \in \mathcal{T}$ *such that* $(A^s \backslash B^s) \cup (A^t \backslash B^t) \subset (A^s \cap A^t)$*, and AFP condition is satisfied for this selection profile. Thus, this data set is rationalizable by AFP model. On the other hand, this data set cannot be rationalizable by AFP model with the trivial consideration mapping with* $\Gamma(A^t) = \{a^t\}$ *for all* $t \in \mathcal{T}$*. Indeed, if we set* Γ *like that,* $x_1 \notin \Gamma(A^3)$ *and* $\Gamma(A^3 \backslash x_1) = \Gamma(A^5) = \{x_5\} \neq \{x_3\} = \Gamma(A^3)$*, which is nothing but a violation of AFP condition.*

We conclude this subsection with referring to the connection with De Clippel and Rozen (2014). They showed that a data set $\mathcal{O}$ is rationalized by an AFP model, if and only if there exists an acyclic ordering $\succ^*$ such that for $s, t \in \mathcal{T}$ with $a^s, a^t \in A^s \cap A^t$ and $a^s \neq a^t$,

$$\exists x' \in A^s \backslash A^t : a^s \succ^* x' \text{ or } \exists x' \in A^t \backslash A^s : a^t \succ^* x'. \tag{13}$$

Note that an ordering $\succ^*$ here is not a priori related to the structure of data, and the existence of such an ordering is directly required. In the above statement, we consciously use the same notation $\succ^*$ with the ordering defined in Lemma 3, since the latter actually obeys the requirement in (13). To show this, note that under AFP condition, the consideration mapping Γ defined as in (12) obeys AFP, and $\succ^*$ defined in Lemma 3 is acyclic. Suppose by way of contradiction that (13) is violated by $\succ^*$. The construction of $\succ^*$ assures that $\Gamma(A^t) \subset A^s \cap A^t$

and $\Gamma(A^s) \subset A^s \cap A^t$. Since Γ obeys AFP, we have $\Gamma(A^t) = \Gamma(A^s)$, which in turn implies $a^t \succ^* a^s$ and $a^s \succ^* a^t$, a contradiction with the acyclicity of $\succ^*$. In this sense, Theorem 1 provides an equivalent condition with De Clippel and Rozen's test, and what is more, the condition is written in terms of the structure of observed choices. Actually, our test and the test by De Clippel and Rozen (2014) perform quite differently in computation as stated in Section 6. Loosely speaking, our method works better in most cases, but in the worst cases where our method takes too much time, De Clippel and Rozen's method works better. There, we applied De Clippel and Rozen's test by converting it as a simple 0-1 integer programming (see Appendix II for the formulation).

4.2 CFP condition

The issue in this subsection is to develop a revealed preference test for a limited consideration model $(\Gamma, \succ)$ where Γ obeys CFP on 2^X (once again, we require for Γ to have CFP on entire domain, rather than on observed feasible sets). Recall that, as stated in Section 2, this model is equivalent to the categorize-then-choose model in Manzini and Mariotti (2012) and the rationalization model by Cherepanov, Fedderson, and Sandroni (2013). Suppose that $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ is collected from an agent obeying a CFP model. Again, without loss of generality, we may assume that $\mathcal{O}$ contains cycles with respect to $\succ^R$. By letting $\left[x^{(q)}\right]_{q=1}^{Q}$ be a selection profile as specified in Proposition 1, the corresponding $\{B^t\}_{t\in\mathcal{T}}$ obeys $\Gamma(A^t) \subset A^t \backslash B^t$ for every $t \in \mathcal{T}$. Bearing this in mind, consider any $s, t \in \mathcal{T}$ such that $A^s \subset A^t$. Then, considering CFP defined in (4), it must hold that $\Gamma(A^t) \cap A^s \subset \Gamma(A^s)$. In addition, since $\Gamma(A^t) \subset A^t \backslash B^t$, this implies that $\Gamma(A^t) \cap B^s = \varnothing$, which in turn implies that $a^t \notin B^s$. In fact, this simple observation completely characterizes whether a data set is consistent with a CFP model.

CFP CONDITION: A selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeys CFP condition, if for every $s, t \in \mathcal{T}$,

$$A^s \subset A^t \Longrightarrow a^t \notin B^s. \tag{14}$$

THEOREM 2. *A data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ is rationalizable by a limited consideration model with CFP, if and only if there exists a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeying CFP condition.*

The proof of Theorem 2 is parallel to that of Theorem 1. The necessity of CFP condition has already been discussed, and the proof for sufficiency is constructive. First of all, given a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ that obeys CFP condition, we can construct $\{B^t\}_{t\in\mathcal{T}}$. Then we define a consideration mapping Γ such that for every $A \subset X$,

$$\Gamma(A) = A \Big\backslash \bigcup_{t: A^t \subset A} B^t. \tag{15}$$

The substantial parts of the proof are to show that Γ constructed as above obeys CFP, and that the binary relation $\succ^*$ defined as $x'' \succ^* x'$ if $x'' = a^t$, $x' \in \Gamma(A^t)$, and $x'' \neq x'$ is acyclic, which are proved in Appendix I.

Lemma 4. *The consideration mapping defined as (15) obeys CFP.*

Lemma 5. *Let $\succ^*$ be a binary relation such that $x'' \succ^* x'$, if $x'' = a^t$ for some $t \in \mathcal{T}$, $x' \in \Gamma(A^t)$, and $x'' \neq x'$. Then, $\succ^*$ is acyclic and for every $t \in \mathcal{T}$, $a^t \in \Gamma(A^t)$.*

The rest of the proof is again similar to the case of Theorem 1, just extending the transitive closure of $\succ^*$ to a strict preference $\succ$ by using Szpilrajn's theorem, which is easily proved to rationalize a data set by the limited consideration model $(\Gamma, \succ)$, where Γ obeys CFP.

Example 2 (continued). *Consider the data set in Example 2, and recall that there are four cycles with respect to the direct revealed preference $\succ^R$: $x_1 \succ^R x_2 \succ^R x_1; x_1 \succ^R x_3 \succ^R x_1; x_2 \succ^R x_4 \succ^R x_2$; and $x_3 \succ^R x_5 \succ^R x_3$. We show that selection profile $[x_1, x_1, x_2, x_3]$ actually succeeds in satisfying CFP condition. In fact, the relevant sets $\{B^t\}_{t\in\mathcal{T}}$ are summarized in Table 1. Looking at the data set and Table 1, one can confirm that CFP condition is satisfied. In particular, we have $A^4 \subset A^2$ and $A^5 \subset A^3$, but $a^2 = x_2 \notin B^4$ and $a^3 = x_3 \notin B^5$.*

Recall that the selection profile $[x_1, x_1, x_2, x_3]$ obeys both AFP condition and CFP condition. Then, a natural question would be: is this data set rationalizable by a limited consideration model with AFP+CFP? In fact, the answer is no, which motivates us to provide a revealed preference test for a model with AFP+CFP.

4.3 AFP+CFP condition

In the rest of this section, we deal with a revealed preference characterization of limited consideration models with AFP+CFP. Clearly, if a data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ is rationalizable by an AFP+CFP model, then it is also consistent with both AFP model and CFP model. Hence, by Theorems 1 and 2, such a data set must obey both AFP condition and CFP condition. However, as we shall show in the example at the end of this subsection, the joint of AFP condition and CFP condition is insufficient to characterize the observable restrictions of such models.

To clarify a necessary condition, let $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ be a data set collected from an agent obeying a limited consideration model with AFP+CFP. That is, an agent has a strict preference $\succ$ on X and a consideration mapping Γ that obeys AFP+CFP. Similar to the previous cases, we may assume that $\mathcal{O}$ contains cycles with respect to $\succ^R$, and let $\left[x^{(q)}\right]_{q=1}^{Q}$ be a selection profile as specified in Proposition 1. Corresponding to this selection profile, $\{B^t\}_{t\in\mathcal{T}}$ is determined as in (9). Since $\mathcal{O}$ must obey AFP, together with $\Gamma(A^t) \subset A^t\backslash B^t$ by Proposition 1, it holds that $\Gamma(A^t) = \Gamma(A^t\backslash B^t)$ for every $t \in \mathcal{T}$. In fact, as a slight extension of this, it holds for $s, t \in \mathcal{T}$ that

$$(A^s\backslash B^s) \subset A^t \Longrightarrow \Gamma(A^t) \subset A^t\backslash B^s. \tag{16}$$

To see (16), we employ both AFP and CFP. First, notice that if $(A^s\backslash B^s) \subset A^t$ holds, then there exist some sets $C \subset B^s$ and $D \subset X\backslash A^s$ such that $A^t = [(A^s\backslash B^s) \cup C \cup D]$.[10] Since, obviously, $(A^t \cap A^s) = [(A^s\backslash B^s) \cup C]$, it holds that $(A^s\backslash B^s) \subset (A^t \cap A^s) \subset A^s$. Then, gathering together with $\Gamma(A^s) = \Gamma(A^s\backslash B^s)$, AFP implies that $\Gamma(A^t \cap A^s) = \Gamma(A^s\backslash B^s)$. In addition, since $\left[\Gamma(A^t) \cap (A^t \cap A^s)\right] = (\Gamma(A^t) \cap A^s)$, it follows from CFP that $(\Gamma(A^t) \cap A^s) \subset \Gamma(A^t \cap A^s)$.[11] Gathering together with $\Gamma(A^t \cap A^s) = \Gamma(A^s\backslash B^s)$, we have $(\Gamma(A^t) \cap A^s) \subset A^s\backslash B^s$. Since $B^s \subset A^s$, we conclude that $\Gamma(A^t) \cap B^s = \varnothing$, which shows that (16) holds. As long as an agent obeys an AFP+CFP model, it must hold that $a^t \in \Gamma(A^t)$, and the relationship (16) impose a restriction on the relationship between chosen alternatives and selection points (or B^t's corresponding to them):

$$(A^s\backslash B^s) \subset A^t \Longrightarrow a^t \notin B^s. \tag{17}$$

As a matter of fact, the conclusions in (16) and (17) have further room for extension, which play a key role in characterizing AFP+CFP. We start from extending (16). Looking at the argument in the preceding paragraph, one can see that the facts of $\Gamma(A^s) \subset A^s\backslash B^s$ and $B^s \subset A^s$ are cornerstones, and once they are known, (16) follows from AFP and CFP. That is, even for general subsets $A', A'' \subset X$, if both $\Gamma(A') \subset A'\backslash V$ and $A'\backslash V \subset A''$ hold for some

[10] The sets C and/or D may be empty.

[11] This is seen by applying CFP for $(A^t \cap A^s) \subset A^t$, which results in $[\Gamma(A^t) \cap (A^t \cap A^s)] \subset \Gamma(A^t \cap A^s)$. Then we can combine this with $[\Gamma(A^t) \cap (A^t \cap A^s)] = (\Gamma(A^t) \cap A^s)$ to obtain the desired result.

$V \subset A'$, then $\Gamma(A'') \subset A''\backslash V$ must hold. We state this as a lemma for future reference. The lemma can be shown through the same logic as deriving (16) by letting $A^s = A'$, $B^s = V$, and $A^t = A''$

LEMMA 6. *Let $A', A'' \subset X$ and Γ be a consideration mapping satisfying AFP+CFP. If both $\Gamma(A') \subset A'\backslash V$ and $A'\backslash V \subset A''$ hold for some $V \subset A'$, then $\Gamma(A'') \subset A''\backslash V$.*

Now we turn to extending (17) with help of Lemma 6. We start from going one step further: consider the situation where for some $r, s, t \in \mathcal{T}$, it holds that $[(A^r \cup A^s)\backslash(B^r \cup B^s)] \subset A^t$. Note that we have $(A^r\backslash B^r) \subset (A^r \cup A^s)$. Then, we can apply Lemma 6 by setting $A' = A^r, V = B^r$, and $A'' = (A^r \cup A^s)$, which results in $\Gamma(A^r \cup A^s) \subset (A^r \cup A^s)\backslash B^r$.[12] Similarly, since we have $(A^s\backslash B^s) \subset (A^r \cup A^s)$, it follows that $\Gamma(A^r \cup A^s) \subset (A^r \cup A^s)\backslash B^s$. Combining, we have $\Gamma(A^r \cup A^s) \subset (A^r \cup A^s)\backslash(B^r \cup B^s)$. Then, since we assumed that $[(A^r \cup A^s)\backslash(B^r \cup B^s)] \subset A^t$, we can again apply Lemma 6 by setting $A' = (A^r \cup A^s), V = (B^r \cup B^s)$, and $A'' = A^t$. This yields $\Gamma(A^t) \subset A^t\backslash(B^r \cup B^s)$, and we conclude that

$$[(A^r \cup A^s)\backslash(B^r \cup B^s)] \subset A^t \Longrightarrow \Gamma(A^t) \subset A^t\backslash(B^r \cup B^s). \tag{18}$$

Clearly, (18) is an extension of (16), and the former derives an extenstion of (17) such that

$$[(A^r \cup A^s)\backslash(B^r \cup B^s)] \subset A^t \Longrightarrow a^t \notin B^r \cup B^s. \tag{19}$$

Then, by inductive argument, we can, in turn, extend (18) and (19) for any subset $\tau \subset \mathcal{T}$ such that $\left(\bigcup_{r\in\tau} A^r \backslash \bigcup_{r\in\tau} B^r\right) \subset A^t$. Namely, by the extension of (19), there must exist a selection profile obeying the following condition.

AFP+CFP CONDITION: A selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeys AFP+CFP condition, if for every $t \in \mathcal{T}$ and any set of indices $\tau \subset \mathcal{T}$,

$$\bigcup_{r\in\tau} A^r \Big\backslash \bigcup_{r\in\tau} B^r \subset A^t \Longrightarrow a^t \notin \bigcup_{r\in\tau} B^r. \tag{20}$$

THEOREM 3. *A data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ is rationalizable by a limited consideration model with AFP+CFP, if and only if there exists a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeying AFP+CFP condition.*

[12] Note that we have $\Gamma(A^r) \subset A^r\backslash B^r$ since the selection profile is chosen as such, $B^r \subset A^r$ by definition, and $(A^r\backslash B^r) \subset (A^r \cup A^s)$ is obvious. Thus the requirements in Lemma 6 are satisfied.

The substantial part of the proof is again the sufficiency part. Similar to Theorems 1 and 2, we construct a pair of a consideration mapping and a strict preference that rationalizes $\mathcal{O}$ based on a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeying the condition and the corresponding $\{B^t\}_{t\in\mathcal{T}}$. To define Γ, we need the following set of indices for every $A \subset X$:

$$\tau(A) = \max\left\{\tau \subset \mathcal{T} : \bigcup_{r\in\tau} A^r \Big\backslash \bigcup_{r\in\tau} B^r \subset A\right\}. \tag{21}$$

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

$$\Gamma(A) = A \Big\backslash \bigcup_{r\in\tau(A)} B^r. \tag{22}$$

Obviously, in order for the above definition to be well-defined, $\tau(A)$ must be uniquely determined for every $A \subset X$, which is actually the case as proved in Appendix I.

Lemma 7. *For every $A \subset X$, $\tau(A)$ is uniquely determined.*

Once we construct a consideration mapping as above, then the rest of proof follows a quite similar path to Theorems 1 and 2. The following two lemmas are proved in Appendix I, and the proof completes by extending $\succ^*$ using Szpilrajn's theorem.

Lemma 8. *The consideration mapping defined as (22) obeys AFP+CFP.*

Lemma 9. *Let $\succ^*$ be a binary relation such that $x'' \succ^* x'$, if $x'' = a^t$, $x' \in \Gamma(A^t)$, and $x'' \neq x'$. Then $\succ^*$ is acyclic and for every $t \in \mathcal{T}$, $a^t \in \Gamma(A^t)$.*

Finally, we point out that the joint of AFP condition and CFP condition does not work as a necessary and sufficient condition for a data set to be consistent with an AFP+CFP model. In the example below, a data set contains a selection profile obeying both AFP condition and CFP condition, and hence it is rationalizable respectively by an AFP model and a CFP model. However, it does *not* contain any selection profile obeying AFP+CFP condition, or equivalently, it is not rationalizable by any AFP+CFP model. This implies that, in general, the joint of two theoretical hypotheses is not necessarily tested by the joint of tests for each hypothesis.

Example 2 (continued). *Consider the data set given in Example 2, and recall that there are four cycles with respect to the direct revealed preference $\succ^R$: $x_1 \succ^R x_2 \succ^R x_1$; $x_1 \succ^R x_3 \succ^R x_1$;*

$x_2 \succ^R x_4 \succ^R x_2$; and $x_3 \succ^R x_5 \succ^R x_3$. It is shown in the examples above that selection profile $[x_1, x_1, x_2, x_3]$ actually succeed in satisfying both AFP condition and CFP condition. However, we claim that this profile violates AFP+CFP condition. The relevant sets for this selection profile are summarized in Table 1. Since $\{x_1\} = A^1 \backslash B^1 \subset A^2$ and $x_2 = a^2 \in B^1$, (17) is violated, let alone AFP+CFP condition. In addition, as shown below, selection profile $[x_1, x_1, x_2, x_3]$ is the only profile that obeys both AFP condition and CFP condition. For a selection profile to satisfy CFP condition, it can contain neither x_4 nor x_5. To see this, suppose that $x^{(3)} = x_4$. Then we have $y^{(3)} = x_2$, $B^4 = \{x_2\}$, $A^4 \subset A^2$, and $a^2 = x_2 \in B^4$, which violates CFP condition. Setting $x^{(4)} = x_5$ leads to a similar violation of CFP condition. Therefore, we must have $x^{(3)} = x_2$ and $x^{(4)} = x_3$ in the profile. Furthermore, if a selection profile satisfies AFP condition, it cannot have x_2 appear twice, or x_3 appear twice in the profile. To see this, consider profile $[x_2, x_1, x_2, x_3]$. Then we have $B^2 = \{x_1, x_4\}$, and thus

$$\{x_2, x_6\} = A^2 \backslash B^2 \subset A^4 \subset A^2 = \{x_1, x_2, x_4, x_6\},$$

but $x_2 = a^2 \neq a^4 = x_4$, which is a violation of AFP condition. The case of profile $[x_1, x_3, x_2, x_3]$ leads to a similar violation of AFP condition.[13]

5 Testing rational shortlist methods

If a data set $\mathcal{O} = \{(a^t, A^t)\}_{t \in \mathcal{T}}$ is collected from an agent obeying a (transitive) rational shortlist method $(\Gamma, \succ)$, then it is trivially consistent with a CFP model (AFP+CFP model). However, it is not difficult to find a data set that is consistent with CFP condition (AFP+CFP condition), but inconsistent with any RSM (TRSM) model. Indeed, for a data set to be rationalizable by an RSM model, it must hold that for every $r, s, t \in \mathcal{T}$ with $A^r = A^s \cup A^t$,

$$a^s = a^t \Longrightarrow a^r = a^s = a^t, \tag{23}$$

which is not guaranteed by the existence of a selection profile obeying CFP condition/AFP+CFP condition.[14] In this section, we provide a test for RSM (TRSM) model.

[13] One can confirm that this example is consistent with the straightforward adaptation of LC-WARP*, a revealed preference characterization of an AFP+CFP model for a choice function shown in Lleras et al. (2015).

[14] If an agent obeys rational shortlist method, $\Gamma(A^r) \subset \Gamma(A^s) \cup \Gamma(A^t)$ is obvious. In addition, $x = a^t = a^s$ implies that no element in $A^s \cup A^t = A^r$ can dominate x with respect to the first step preference, and x dominates with

Suppose that an agent has two preferences $\succ'$ and $\succ$, where the former is merely acyclic while the latter is a strict preference, and that a consideration mapping Γ is defined as (7). Similar to the previous models, we may assume that a data set $\mathcal{O}$ collected from such an agent contains cycles with respect to $\succ^R$, and let $\left[x^{(q)}\right]_{q=1}^{Q}$ be a selection profile as specified in Proposition 1. Corresponding to this selection profile, $\{B^t\}_{t\in\mathcal{T}}$ is determined as in (9). Recall, by the definition of selection points in Proposition 1, for every $x' \in B^t$ we have $x' \notin \Gamma(A^t)$, which means that there exists some $x'' \in A^t \backslash x'$ such that $x'' \succ' x'$. On the other hand, $x' \in B^t$ means that x' is a chosen alternative in some observed feasible set, say A^s. Then, it must follow that $x' \not\succ^R x''$; otherwise, since we have $x'' \succ' x'$, the definition of Γ will require $x' \notin \Gamma(A^s)$, which contradicts that x' is the chosen alternative at A^s.

Given the discussion above, we can define a binary relation $\rhd$ on X such that: $x'' \rhd x'$ if $x' \in B^t$ for some $t \in \mathcal{T}$, $x'' \in A^t \backslash x'$, and $x' \not\succ^R x''$. Note that for every $x' \in B^t$, 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 first step preference $\succ'$. In addition, the acyclicity of $\succ'$ requires that we can always find a selection $\rhd' \subset \rhd$ that is acyclic, and for every $t \in \mathcal{T}$ and $x' \in B^t$, there exists some $x'' \in A^t \backslash x'$ with $x'' \rhd' x'$. Furthermore, if the first step preference $\succ'$ is assumed to be transitive, a selection $\rhd'$ has to be chosen so that

$$\text{for every } x' \in B^t \text{ and } z^1, ..., z^k,\ x'' \rhd' z^1 \rhd' \cdots \rhd' z^k \rhd' x' \Longrightarrow x' \not\succ^R x''. \tag{24}$$

Now, $\rhd'$ is a "correct" guess of the first step preference, and if transitivity is imposed, the above implies that $x'' \succ' x'$. Hence, if $x' \succ^R x''$ were to hold, then it leads 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 an RSM (TRSM) model.

RSM condition: A selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeys RSM condition, if for the corresponding $\{B^t\}_{t\in\mathcal{T}}$, there exists an acyclic selection $\rhd'$ of $\rhd$, where for every $t \in \mathcal{T}$,

$$\text{for every } x' \in B^t, \text{ there exists } x'' \in A^t \text{ with } x'' \rhd' x'. \tag{25}$$

TRSM condition: A selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeys TRSM condition, if for the correspond-

respect to the second step preference all other elements in $\Gamma(A^r) \subset \Gamma(A^s) \cup \Gamma(A^t)$.

ing $\{B^t\}_{t\in\mathcal{T}}$, there exists an acyclic selection $\rhd'$ of $\rhd$ that obeys (24) and (25).

THEOREM 4. *A data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ is rationalizable by an RSM model, if and only if there exists a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeying RSM condition.*

THEOREM 5. *A data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ is rationalizable by a TRSM model, if and only if there exists a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ obeying TRSM condition.*

The proofs of the above theorems are almost identical and the necessity parts of them have been already discussed. Hence, we only prove the sufficient parts of them. Using an acyclic selection $\rhd'$ of $\rhd$, define Γ as

$$\Gamma(A) = \{x \in A : \nexists x' \in A \text{ such that } x' \rhd' x\}. \tag{26}$$

Note that the selection $\rhd'$ is acyclic, so we use it as a first step preference for the case of Theorem 4. If we can find $\rhd'$ so that it obeys (24) in addition to (25), then we use the transitive closure of it, say, $\rhd''$ as a first step preference and define Γ by using it instead of $\rhd'$. Note further that $\Gamma(A^t) \subset A^t \backslash B^t$, by the definition of $\rhd'$ 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, which are proved in Appendix I. In the following lemma, Γ is defined by (26) in testing RSM model, while $\rhd'$ in (26) should be replaced with $\rhd''$ in testing TRSM model.

LEMMA 10. *Let $\succ^*$ be a binary relation such that $x'' \succ^* x'$, if $x'' = a^t$ for some $t \in \mathcal{T}$, $x' \in \Gamma(A^t)$, and $x'' \neq x'$. Then $\succ^*$ is acyclic, and for every $t \in \mathcal{T}$, $a^t \in \Gamma(A^t)$.*

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 an RSM (TRSM) model $(\Gamma, \succ)$.

REMARK: In testing TRSM condition, the search for an acyclic selection $\rhd'$ of $\rhd$ that obeys (24) and (25) can be done by way of a simple 0-1 integer programming (see Appendix II for the formulation). In the simulation in Section 6, we actually use it, which very well works. In principle, the RSM model $\rhd'$ can also be searched using a similar 0-1 integer programming. However, requiring acyclicity of $\rhd'$ in the programming can be computationally heavy, so applying 0-1 integer programming for RSM may not be practical.[15]

It is shown by Manzini and Mariotti (2007) that an RSM model can be characterized by a combination of two axioms on a data set, namely, Weak WARP and Expansion (see Appendix III). The former is implied when the consideration mapping obeys CFP. The latter requires that for every $A', A'' \subset X$, if $x = f(A') = f(A'')$, then $x = f(A' \cup A'')$, where f is a choice function. Given this, one may be tempted to consider that an RSM model is tested by the joint of CFP condition and (23), a straightforward partial-observation version of Expansion. The following example shows that this is not the case, i.e. we present a data set that obeys CFP condition and (23), but violates RSM condition. A similar example can be found for the joint of AFP+CFP condition and (23).

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 below:*

t	1	2	3	4	5	6
A^t	$\{x_1, x_2, x_4\}$	$\{x_1, x_2\}$	$\{x_3, x_4, x_6\}$	$\{x_3, x_4\}$	$\{x_2, x_5, x_6\}$	$\{x_5, x_6\}$
a^t	x_1	x_2	x_3	x_4	x_5	x_6

It can be seen that Expansion is trivially satisfied, because the chosen alternatives are all different. Note that there are four cycles with respect to $\succ^R$*:* $x_1 \succ^R x_2 \succ^R x_1$*;* $x_3 \succ^R x_4 \succ^R x_3$*;* $x_5 \succ^R x_6 \succ^R x_5$*; and* $x_1 \succ^R x_4 \succ^R x_3 \succ^R x_6 \succ^R x_5 \succ^R x_2 \succ^R x_1$*. We first show that RSM condition cannot be satisfied by any selection profile. Consider the cycle* $x_1 \succ^R x_2 \succ^R x_1$*. If we choose* x_2 *to be a selection point, we will have* $a^1 = x_1 \in B^2$*. However, then, there does not exist any* $x \in A^2$ *such that* $x_1 \not\succ^R x$*, and we cannot define* $\rhd$ *for* x_1*. Therefore, we need to choose* x_1 *as the selection point for this cycle. By the same logic, we must choose* x_3 *and* x_5 *to be the selection points of the second and third cycles respectively. Then we must have* $x_4 \rhd x_2$*,* $x_6 \rhd x_4$*, and* $x_2 \rhd x_6$*, and it will be impossible to find an acyclic selection of* $\rhd$*. CFP condition is satisfied by the selection profile* $[x_1, x_3, x_5, x_3]$*. Note that the only set inclusions of feasible sets that we have are* $A^t \subset A^{t-1}$ *for* $t = 2, 4, 6$*. On the other hand, since* $B^t = \varnothing$ *for* $t = 2, 4, 6$*, CFP condition is trivially satisfied.*

[15] For RSM, we applied a different strategy to find a suitable selection $\rhd'$, of which the detail is available from the authors upon request.

6 Simulation: relative observable restrictions

6.1 Setting

In this section, the revealed preference tests in Sections 4 and 5 are applied to randomly generated data sets to compare relative strength of observable restrictions between models, not only between models that are theoretically interdependent, but also between models that are theoretically independent.. We generated 10,000 random data sets with $|X| = 10, |\mathcal{T}| = 20, \min |A^t| = 2$, and $\max |A^t| = 8$. Firstly, we randomly generated 100 variations of feasible sets $\mathbb{A}_n := \{A_n^t\}_{t\in\mathcal{T}}$ for $n = 1, \ldots, 100$: fixing n, in generating each A_n^t, we set $|A_n^t| \in \{2, \ldots, 8\}$ 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 $A_n^{t'} \neq A_n^{t''}$ for $t' \neq t''$. For each profile of feasible sets $\mathbb{A}_n = \{A_n^t\}_{t\in\mathcal{T}}$, a random choice data set $\{a_{i,n}^t\}_{t\in\mathcal{T}}$ is generated for $i = 1, \ldots, 100$: fixing n and i, $a_{i,n}^t$ is chosen following a uniform distribution over A_n^t for every $t \in \mathcal{T}$.

Consequently we have a random choice data set $\mathcal{O}_{i,n} = \{(a_{i,n}^t, A_n^t)\}_{t\in\mathcal{T}}$ for $i = 1, \ldots, 100$ for which we apply our revealed preference tests. Note that we randomize feasible sets, as well as choices over them, since, in general, observable restriction of a specific model depends on the structure of the feasible sets $\mathbb{A}_n = \{A_n^t\}_{t\in\mathcal{T}}$. For example, if $A^s \cap A^t = \emptyset$ for every $s, t \in \mathcal{T}$, then it is impossible to satisfy the LHS of (11) in AFP condition, and the AFP model is trivially nonrefutable. For each $\mathcal{O}_{i,n} = \{(a_{i,n}^t, A_n^t)\}_{t\in\mathcal{T}}$, we tested AFP, CFP, AFP+CFP, RSM, and TRSM, as well as SARP. We derived the pass rates for these tests under each profile of feasible sets, as well as the average pass rates of them over 100 profiles of feasible sets. In addition, we apply straightforward adaptations of existing full-observation based tests to our partially observed data sets to see if they could approximate necessary and sufficient conditions (see Appendix III for details of full observation based tests).

Our simulation can be regarded as Bronars' test in the context of limited consideration models, and one can measure observable restriction of each model by using its pass rate. Indeed, if we collect a sufficiently large number of random choices, then the pass rate approximates the proportion of choices that are model-consistent to 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 restriction is weak. As shown by Selten (1991) and Beatty and Crawford (2011), this measure of observable restriction plays a key role in considering the measure of plausibility of a model

test	our tests	full obsv. tests
SARP	0	0
AFP	0.9927	0.9954
CFP	0.6298	0.6298
AFP+CFP	0.0396	0.6176
RSM	0.0259	0.5083
TRSM	0.0006	0.5050

Table 2: Average pass rates.

based on empirical or experimental data sets. Given empirical or experimental data sets, Selten's index evaluates a model by the difference of the pass rate calculated from actual data sets and the proportion of model-consistent choices to all logically possible choices. Intuitively, a "nice" model in terms of Selten index is a model with higher pass rate and stronger observable restrictions.[16] Since the pass rate from uniformly generated data sets approximates observable restrictions of models, practically, one can calculate Selten's index as the difference of the pass rate of actual data sets and the pass rate of randomly generated data sets. We postpone the issue of comparing models by Selten indices using some empirical/experimental data as an interesting future research.

6.2 Results

In Table 2, the average pass rates of 100 different profiles of feasible sets (10,000 agents) are summarized. The left hand side column gives the pass rates of revealed preference tests presented in our paper, and the right column gives the pass rates of the corresponding straightforward adaption of full observation version tests existing in the literature.

The pass rate results show that the AFP model is extremely permissive, letting more than 99% of the random agents pass the test, and CFP model is also quite permissive. On the other hand, we can say that that for AFP+CFP, RSM, and TRSM models, observable restrictions are reasonably strong. What is striking is that while more than 60% of all agents passed both AFP and CFP, the pass rate of AFP+CFP model is significantly lower (lower than 0.04). This shows that the combination of the two models lead to a model that is much more restrictive, which is an observation that the theoretical results do not necessarily show. Our simulation also shows us that, in our setting, observable restriction of RSM model is slightly stronger

[16]It is known that this simple index satisfies several nice axioms. See Selten's original paper for details.

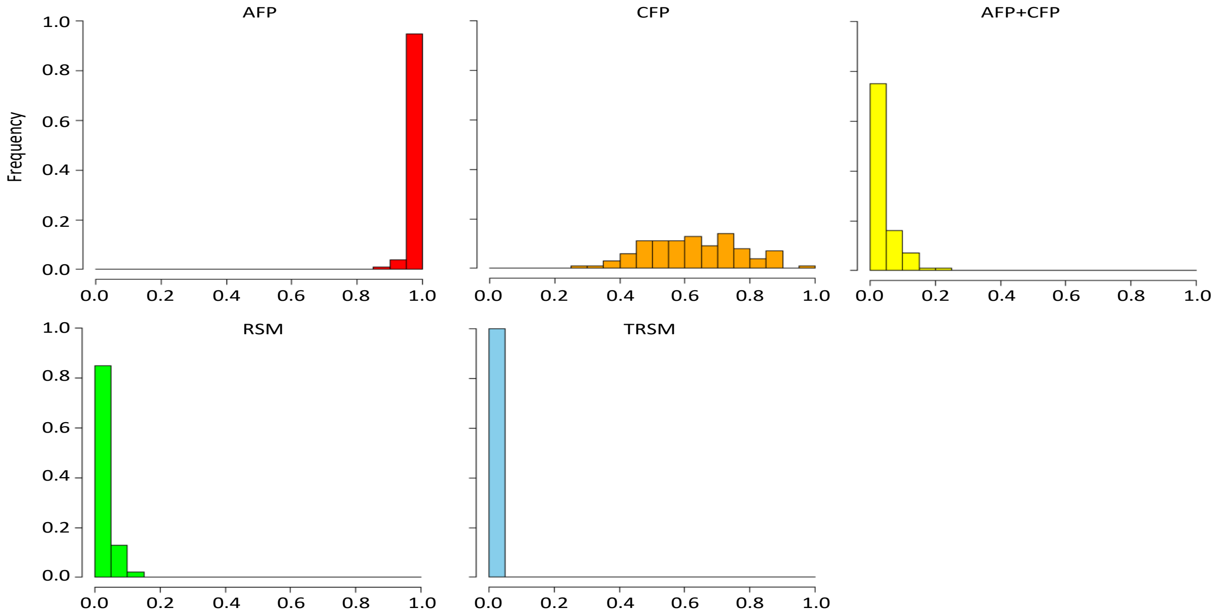

Figure 1: Histograms.

than that of AFP+CFP model. Also, we can loosely say that the full observation version tests may be a good approximation for AFP and CFP models, even when we deal with a partially observed data set. Meanwhile, for AFP+CFP, RSM, and TRSM models, the gaps between our tests and the full observation tests are large.

Note that the agents can be partitioned into eight types: agent obeys either (i) TRSM; (ii) RSM and AFP+CFP but not TRSM; (iii) RSM but not AFP+CFP; (iv) AFP+CFP but not RSM; (v) CFP and AFP but neither AFP+CFP nor RSM; (vi) only CFP; (vii) only AFP; (viii) none of the models. In our simulation, every type consists of a positive number of agents, which means that, under our setting, the theoretical independence between models is maintained.[17]

Figure 1 visually summarizes the distributions of pass rates for each test, where the horizontal axis is the pass rate given the profile of feasible sets, running from 0 to 1 with bin width 0.05. The vertical axis is the frequency of profiles of feasible sets of which the pass rates drop in each bin. It shows that the pass rate for CFP test has a large variance depending on the structure of feasible sets, while pass rates for other models are more accumulated to around

[17]Out of 10,000 agents, the distribution of agents' type is as follows: (i) 24 agents, (ii) 4 agents, (iii) 186 agents, (iv) 369 agents, (v) 5575 agents, (vi) 140 agents, (vii) 3576 agents, and (viii) 126 agents.

either 0 and 1.

We conclude this subsection with a remark concerning computation time. Though it varies across agents, loosely speaking, tests are reasonably fast except for AFP, and almost all tests for 10,000 agents finish within two weeks. For the AFP test, we first applied our test for all agents, and then for the agents that took too much time, we applied the 0-1 integer programming version of the test by De Clippel and Rozen (2014). For each agent, on average, it took 3.12 seconds to write out cycles with respect to $\succ^R$, 87.20 seconds for AFP, 0.96 seconds for CFP, 6.59 seconds for AFP+CFP, 0.85 seconds for RSM, and 0.01 seconds for TRSM. For the AFP test, firstly we first gave at most 1 second for each agent in applying our test. As a result, the calculation finished for 9,114 agents, which took an average of 0.84 seconds per agent. Then, for the remaining 886 agents we applied the method of De Clippel and Rozen (2014). The latter took an average of 974.65 seconds per agent, but faster than continuing our backgracking for these agents.[18]

6.3 Backtracking procedure

Here we explain how we use the backtracking method in searching for a selection profile obeying the conditions in our revealed preference tests. Suppose that we have a data set $\mathcal{O} = \{(a^t, A^t)\}_{t\in\mathcal{T}}$ which has $Q > 0$ cycles with respect to the direct revealed preference, and let us consider a limited consideration model $\mathbf{M} \in \{\text{AFP, CFP, AFP+CFP, RSM, TRSM}\}$. As already mentioned in Section 3, if the data set has no cycle, then it is rationalized by the rational choice model, which is a subclass of the models referred to above.

Before going into the description of our method, let us define some additional concepts. Fix $\overline{Q} \leqslant Q$ and consider a "partial" selection profile $\left[x^{(q)}\right]_{q=1}^{\overline{Q}}$. Define for every $t \in \mathcal{T}$,

$$B^t_{\overline{Q}}\left(\left[x^{(q)}\right]_{q=1}^{\overline{Q}}\right) = \left\{y^{(q)} \in A^t : a^t = x^{(q)},\ \text{where } 1 \leqslant q \leqslant \overline{Q}\right\}, \tag{27}$$

which are counterparts of (9) corresponding to a "partial" selection profile of length $\overline{Q} < Q$.[19] With a slight abuse of terminology, we say that a selection profile $\left[x^{(q)}\right]_{q=1}^{\overline{Q}}$ obeys (violates) $\mathbf{M}$ condition, if obeys (violates) $\mathbf{M}$ condition based on the corresponding $\{B^t_{\overline{Q}}\}_{t\in\mathcal{T}}$. For example,

[18] The computers that we used for calculation are standard computers: the processors vary from 1.20 GHz Intel Core i5 to 3.3 GHz Intel Core i7, and RAM vary from 4 GB to 16 GB.

[19] We omit the argument and denote these sets by $B^t_{\overline{Q}}$ when there is no risk of confusion.

for $\overline{Q}$, we say that a partial selection profile $\left[x^{(q)}\right]_{q=1}^{\overline{Q}}$ obeys AFP condition, if for every $s, t \in \mathcal{T}$,

$$\left(A^s \backslash B^s_{\overline{Q}}\right) \cup \left(A^t \backslash B^t_{\overline{Q}}\right) \subset (A^s \cap A^t) \Longrightarrow a^s = a^t. \tag{28}$$

Parallel terminology is used for other models as well.

An important fact is that a "longer" selection profile is harder to satisfy the conditions in revealed preference tests. To illustrate this more in detail, let us consider the case of AFP. Suppose that for some $\overline{Q} < Q$, there is a selection profile $\left[\bar{x}^{(q)}\right]_{q=1}^{\overline{Q}}$ that fails AFP condition. That is, there exist some $r, s \in \mathcal{T}$ such that $(A^r \backslash B^r_{\overline{Q}}) \cup (A^s \backslash B^s_{\overline{Q}}) \subset (A^r \cap A^s)$, but $a^r \neq a^s$. Now consider a selection profile $[\bar{x}^{(1)}, \ldots, \bar{x}^{(\overline{Q})}, x^{(\overline{Q}+1)}]$, and the corresponding sets $\{B^t_{\overline{Q}+1}\}_{t \in \mathcal{T}}$. Then, as seen from (27), it holds that $B^t_{\overline{Q}} \subset B^t_{\overline{Q}+1}$ for every $t \in \mathcal{T}$. This in turn implies that $(A^r \backslash B^r_{\overline{Q}+1}) \cup (A^s \backslash B^s_{\overline{Q}+1}) \subset (A^r \cap A^s)$ holds, and selection profile $[\bar{x}^{(1)}, \ldots, \bar{x}^{(\overline{Q})}, x^{(\overline{Q}+1)}]$ fails AFP condition. Hence, there is no hope for any partial or complete selection profile containing $\left[\bar{x}^{(q)}\right]_{q=1}^{\overline{Q}}$ to obey **M** condition.

In fact, the above remains to be true for other models. To see this, fix some $Q' < Q$ and selection profile $\left[\bar{x}^{(q)}\right]_{q=1}^{Q'}$, and consider any selection profile of $Q''(> Q')$ length $\left[x^{(q)}\right]_{q=1}^{Q''}$ that contains $\left[\bar{x}^{(q)}\right]_{q=1}^{Q'}$. Then by definition, $B^t_{Q'} \subset B^t_{Q''}$ for every $t \in \mathcal{T}$. This implies that (i) the LHS is more permissive in AFP condition; (ii) the RHS is more restrictive in CFP condition; (iii) the LHS is more permissive in AFP+CFP condition; (iv) $\rhd$ is stronger and thus more difficult to find an acyclic (asymmetric and transitive) selection of $\rhd$ in RSM (TRSM) condition. All of them imply that a longer selection profile makes each condition more restrictive. In particular, if a partial selection profile $\left[\bar{x}^{(q)}\right]_{q=1}^{\overline{Q}}$ fails to obey **M** condition, no complete selection profile containing it can obey **M** condition.

Due to the above observation, for every model **M**, we can search for a successful selection profile cycle-by-cycle. The big picture of our procedure is as follows. We start from searching the single-element selection profile $[x^{(1)}]$ from the first cycle so that it obeys **M** condition. If there is no such selection, then we can immediately conclude that the data set is inconsistent with the model in question. If we find a successful partial selection profile $\left[x^{(q)}\right]_{q=1}^{(\overline{Q}-1)}$ for $\overline{Q} \geqslant 2$, then we can concentrate on finding out a suitable selection point from $\overline{Q}$-th cycle. If there is no such selection, then we should discard the current $\left[x^{(q)}\right]_{q=1}^{(\overline{Q}-1)}$ immediately and go back to earlier cycles to update it. The detail of our procedure is summarized as the flowchart in Figure 2, and we provide an extended explanation below. The example in the end of this subsection may be helpful to follow it.

Recall that our ultimate goal is to find a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ that obeys $\mathbf{M}$ condition. In our procedure, we search for such a selection profile cycle-by-cycle using the following variables: $\overline{Q}$ $(1 \leqslant \overline{Q} \leqslant Q)$ indicating the cycle we are looking at, and k_q indicating for every $1 \leqslant q \leqslant Q$, which element in q-th cycle should be considered when $\overline{Q} = q$.

Initially, we set $k_q = 1$ for all q and $\overline{Q} = 1$. So, we start from choosing the first element of the first cycle as $x^{(1)}$ and check if the single-element selection profile $\left[x^{(1)}\right]$ obeys $\mathbf{M}$ condition. If it is the case, then we update $\overline{Q}$ from 1 to 2 and go to the next step, which is explained in the succeeding paragraph, with keeping $k_1 = 1$. If $[x^{(1)}]$ fails to obey $\mathbf{M}$ condition, then we update k_q from 1 to 2 and choose the second element of the first cycle as $x^{(1)}$ and test $\mathbf{M}$ condition for updated $[x^{(1)}]$. Repeating this, if we get $[x^{(1)}]$ obeying $\mathbf{M}$ condition at some k_1, then we go to $\overline{Q} = 2$ with keeping k_1 at that value. On the other hand, if the update of k_1 continues up to $k_1 = K_1$, which is the last element from the first cycle, and still we cannot get $[x^{(1)}]$ obeying $\mathbf{M}$ condition, then we set $\overline{Q} = 0$ and conclude that the data set in question is not rationalizable by $\mathbf{M}$ model.

If the procedure reaches some $2 \leqslant \overline{Q} \leqslant Q - 1$, then it means that we successfully found a partial selection profile $[x^{(1)}, x^{(2)}, ..., x^{(\overline{Q}-1)}]$ obeying $\mathbf{M}$ condition. Though, for every $1 \leqslant q \leqslant Q$, the value of k_q may vary depending on history, it does not affect the following argument. The procedure adds $x^{(\overline{Q})} = x^{k_{\overline{Q}}}$ to the existing selection profile to get $[x^{(1)}, x^{(2)}, ..., x^{(\overline{Q}-1)}, x^{\overline{Q}}]$. If this new selection profile obeys $\mathbf{M}$ condition, we go to $(\overline{Q}+1)$-th cycle with keeping the value of $k_{\overline{Q}}$ as it stands. Otherwise, we redefine $k_{\overline{Q}}$ by adding 1 to the previous $k_{\overline{Q}}$, and reexamine the selection profile with $x^{(\overline{Q})}$ being updated to the redefined $x^{k_{\overline{Q}}}$. Again, if we find a successful $x^{k_{\overline{Q}}}$ at some $k_{\overline{Q}}$, we proceed to $(\overline{Q}+1)$-th cycle with keeping its value. Suppose that we cannot find any successful selection point $x^{(\overline{Q})}$ even when $k_{\overline{Q}}$ reaches $K_{\overline{Q}}$, the last element from the $\overline{Q}$-th cycle. Then, we go back to the largest positive $Q' < \overline{Q}$ with $k_{Q'} < K_{Q'}$, i.e., $\overline{Q}$ is rewinded to Q'. In doing so, for all $q > Q'$, k_q is reset to 1 and $k_{Q'}$ is updated by adding 1 to its present value. When we cannot find any such Q', then, we just set $\overline{Q} = 0$ and conclude that the data set is inconsistent with $\mathbf{M}$ model.

At the conclusion of the algorithm, we have either $\overline{Q} = Q$ or $\overline{Q} = 0$. In the former case, we have a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ that obeys $\mathbf{M}$ condition. If the latter case is reached, we conclude that the data set is not rationalizable by $\mathbf{M}$ model. Note that, by the construction of our procedure, $\overline{Q} = 0$ is reached, if and only if there exists some $\overline{Q} > 0$ for which any

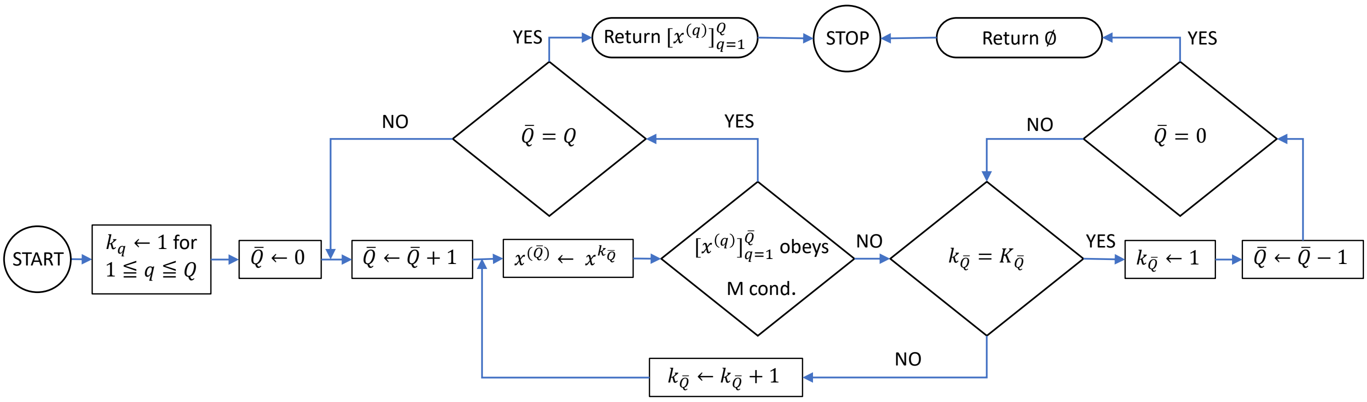

Figure 2: Flowchart of backtracking.

(partial) selection profile $\left[x^{(q)}\right]_{q=1}^{\overline{Q}}$ cannot satisfy $\mathbf{M}$ condition. Recalling the fact that a longer selection profile is harder to obey the conditions, this indeed implies that the data set cannot be rationalized by $\mathbf{M}$ model.

Remark: One advantage of the backtracking approach is that we may be able to determine, at an early stage of the process of searching for a selection profile, 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 shoretr cycles come first: whenever $q' < q''$, q'-th cycle is weakly shorter than q''-th cycle. The cycles in Example 3 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 profile fails when adding a selection point 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).

Example 3 (continued). *Consider the data set in Example 3. Table 3 shows the procedure of which we determine that the data set of Example 3 fails RSM condition. Recall that the data set has four cycles (we number cycles in the following order):*

1. $x_1 \succ^R x_2 \succ^R x_1$
2. $x_3 \succ^R x_4 \succ^R x_3$
3. $x_5 \succ^R x_6 \succ^R x_5$
4. $x_1 \succ^R x_4 \succ^R x_3 \succ^R x_6 \succ^R x_5 \succ^R x_2 \succ^R x_1$

$\overline{Q}$	(k_1, k_2, k_3, k_4)	selection profile	RSM condition
1	$(1, 1, 1, 1)$	$[x_1]$	PASS
2	$(1, 1, 1, 1)$	$[x_1, x_3]$	PASS
3	$(1, 1, 1, 1)$	$[x_1, x_3, x_5]$	FAIL
3	$(1, 1, 2, 1)$	$[x_1, x_3, x_6]$	FAIL
2	$(1, 2, 1, 1)$	$[x_1, x_4]$	FAIL
1	$(2, 1, 1, 1)$	$[x_2]$	FAIL
0		$\varnothing$	**STOP**

Table 3: Backtracking process applied to Example 3 for testing RSM.

Following our backtracking procedure, we first set $\overline{Q} = 1$ and $(k_1, k_2, k_3, k_4) = (1, 1, 1, 1)$ to see if $[x_1]$ passes RSM condition, which is actually affirmative. Then, we proceed to the second cycle by setting $\overline{Q} = 2$ and check if $[x_1, x_3]$ obeys RSM, which is again the case. Then, we set $\overline{Q} = 3$ and move to the third cycle to check if $[x_1, x_3, x_5]$ is consistent with RSM, which is, in turn, negative. In this case, we keep $\overline{Q} = 3$ and update k_3 from 1 to 2 in order to test if $[x_1, x_3, x_6]$ passes RSM, which is again negative.

Now, we cannot find any other element from the third cycle obeying RSM, as long as $[x_1, x_3]$ is selected from the first and second cycles. So we now rewind $\overline{Q}$ to 2 and change k_2 from 1 to 2 to change a selection from the second cycle (at this stage, k_3 is reset to 1). Looking at $[x_1, x_4]$, it fails RSM, and now we cannot find any successful selection profile as long as the selection from the first cycle is x_1.

Since there is no other possibility in the second cycle, we have to go back to $\overline{Q} = 1$ and reexamine the selection from the first cycle. The remaining possibility is $[x_2]$, but it fails to obey RSM. Then $\overline{Q}$ is set to 0, which means that the data set cannot be rationalized by RSM.

Appendix I: Proofs of Lemmas

Proof of Lemma 1

Suppose that for some $s, t \in \mathcal{T}$ both $A^t \backslash B^t \subset A \subset A^t$ and $A^s \backslash B^s \subset A \subset A^s$ simultaneously hold. Then, it follows that $(A^t \backslash B^t) \cup (A^s \backslash B^s) \subset (A^t \cap A^s)$. By AFP condition, we must have $a^t = a^s$, and then $B^t \cap A = B^s \cap A$ follows from the assumption that $A \subset (A^t \cap A^s)$ and the definition of sets B^t, B^s. Thus we conclude that $A \backslash B^t = A \backslash B^s$. □

Proof of Lemma 2

We show that Γ as defined in (12) obeys AFP. Consider $A \subset X$ and $x \in A$ such that $x \notin \Gamma(A)$. This implies that there exists some $t \in \mathcal{T}$ such that $A^t \backslash B^t \subset A \subset A^t$ and $x \in B^t$. Note that, by definition, $\Gamma(A) = A \backslash B^t$. Now consider the set $A \backslash x$. Since $x \in B^t$, it follows that $A^t \backslash B^t \subset A \backslash x \subset A^t$, and thus $\Gamma(A \backslash x) = (A \backslash x) \backslash B^t$. Recalling that $x \in B^t$, it follows that $\Gamma(A \backslash x) = (A \backslash x) \backslash B^t = A \backslash B^t = \Gamma(A)$. □

Proof of Lemma 3

To see that $\succ^*$ is acyclic, suppose to the contrary, that is, there exists a cycle with respect to $\succ^*$, expressed as: $x^1 \succ^* x^2 \succ^* \cdots \succ^* x^L \succ^* x^1$. Note that $x'' \succ^* x'$ implies $x'' \succ^R x'$, which follows by the way these binary relations are defined. Therefore, the cycle above implies $x^1 \succ^R x^2 \succ^R \cdots \succ^R x^L \succ^R x^1$. Then, there exists a selection point x^ℓ, i.e. $x^{(q)} = x^\ell$ and $y^{(q)} = x^{\ell+1}$ for some $q \in \{1, \ldots, Q\}$. Note that by definition of the direct revealed preference, there must exist some $t \in \mathcal{T}$ such that $a^t = x^{(q)}$ and $y^{(q)} \in A^t$. Moreover, definitions of a selection point and sets $\{B^t\}_{t \in \mathcal{T}}$ imply that $x^{\ell+1} = y^{(q)} \in B^t$ for such $t \in \mathcal{T}$, which in turn implies $x^{\ell+1} \notin \Gamma(A^t)$. Hence it is impossible to have $x^\ell = a^t \succ^* x^{\ell+1}$, and we conclude that a cycle with respect to $\succ^*$ cannot exist.

The fact that $a^t \in \Gamma(A^t)$ for every $t \in \mathcal{T}$ follows directly from AFP. To see this, suppose not. Then there exists some $s \in \mathcal{T}$ such that $A^s \backslash B^s \subset A^t \subset A^s$ and $a^t \in B^s$. However, this is impossible, since AFP requires $a^t = a^s$, which contradicts $a^t \in B^s$. Summarizing, we have shown that $\succ^*$ is acyclic and a^t maximizes $\succ^*$ within the set $\Gamma(A^t)$ for every $t \in \mathcal{T}$. □

Proof of Lemma 4

We show that Γ obeys CFP. Consider $A', A'' \subset X$ such that $A' \subset A''$, and $x \in A'$ with $x \notin \Gamma(A')$. Then it suffices to show $x \notin \Gamma(A'')$. Note that $x \notin \Gamma(A')$ implies that there exist some $t \in \mathcal{T}$ such that $A^t \subset A'$ and $x \in B^t$. Since $A' \subset A''$, we clearly have $A^t \subset A''$, and it follows that $x \notin \Gamma(A'')$. □

Proof of Lemma 5

To see that $\succ^*$ is acyclic, suppose to the contrary, that is, there exists a cycle with respect to $\succ^*$, expressed as: $x^1 \succ^* x^2 \succ^* \cdots \succ^* x^L \succ^* x^1$. Note that $x'' \succ^* x'$ implies $x'' \succ^R x'$, which follows by the way these binary relations are defined. Therefore, the cycle above implies $x^1 \succ^R x^2 \succ^R \cdots \succ^R x^L \succ^R x^1$. Then, there exists a selection point x^ℓ, i.e. $x^{(q)} = x^\ell$ and $y^{(q)} = x^{\ell+1}$ for some $q \in \{1, \ldots, Q\}$. Note that by definition of the direct revealed preference, there must exist some $t \in \mathcal{T}$ such that $a^t = x^{(q)}$ and $y^{(q)} \in A^t$. Moreover, definitions of a selection point and sets $\{B^t\}_{t\in\mathcal{T}}$ imply that $x^{\ell+1} = y^{(q)} \in B^t$ for such $t \in \mathcal{T}$, which in turn implies $x^{\ell+1} \notin \Gamma(A^t)$. Hence it is impossible to have $x^\ell = a^t \succ^* x^{\ell+1}$, and we conclude that a cycle with respect to $\succ^*$ cannot exist.

Next, we show that for every $t \in \mathcal{T}$, $a^t \in \Gamma(A^t)$. Suppose not. Then there exists some $s \in \mathcal{T}$ such that $A^s \subset A^t$ and $a^t \in B^s$. However, this is impossible, since CFP condition requires that $a^t \notin B^s$. Summarizing, we have shown that $\succ^*$ is acyclic and a^t maximizes $\succ^*$ within the set $\Gamma(A^t)$ for every $t \in \mathcal{T}$. □

Proof of Lemma 7

Suppose by way of contradiction that $\tau(A)$ is not unique, i.e. there exist $\tau_1(A) \neq \tau_2(A)$ that obey (21). Then $\left(\bigcup_{r\in\tau_1(A)} A^r \backslash \bigcup_{r\in\tau_1(A)} B^r\right) \subset A$ and $\left(\bigcup_{r\in\tau_2(A)} A^r \backslash \bigcup_{r\in\tau_2(A)} B^r\right) \subset A$. Hence $\left(\bigcup_{r\in\tau_1(A)} A^r \backslash \bigcup_{r\in\tau_1(A)} B^r\right) \cup \left(\bigcup_{r\in\tau_2(A)} A^r \backslash \bigcup_{r\in\tau_2(A)} B^r\right) \subset A$, which can be expressed as $\left[\bigcup_{r\in\tau_1(A)\cup\tau_2(A)} A^r \backslash \left(\bigcup_{r\in\tau_1(A)} B^r \cup \bigcup_{r\in\tau_2(A)} B^r\right)\right] \subset A$. Then, this implies

$$\left[\bigcup_{r\in\tau_1(A)\cup\tau_2(A)} A^r \Bigg\backslash \bigcup_{r\in\tau_1(A)\cup\tau_2(A)} B^r\right] \subset A.$$

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 $\tau_2(A)$. □

Proof of Lemma 8

To see that Γ obeys CFP, consider $A', A'' \subset X$ with $A' \subset A''$, and $x \in A'$ such that $x \notin \Gamma(A')$. This means that $x \in \bigcup_{r\in\tau(A')} B^r$. Since $\tau(\cdot)$ is clearly monotonic, it follows that $\tau(A') \subset \tau(A'')$, and hence, $x \in \bigcup_{r\in\tau(A'')} B^r$. This assures that $x \notin \Gamma(A'')$.

To see that Γ obeys AFP, take any $A \subset X$ and any $x \in A$ with $x \notin \Gamma(A)$. This means that

$x \in \bigcup_{r\in\tau(A)} B^r$, which in turn implies that

$$\left(\bigcup_{r\in\tau(A)} A^r \Big\backslash \bigcup_{r\in\tau(A)} B^r\right) \subset A\backslash x. \tag{29}$$

The maximality and uniqueness of $\tau(\cdot)$, combined with (29), imply $\tau(A) \subset \tau(A\backslash x)$. On the other hand, the monotonicity of $\tau(\cdot)$ implies $\tau(A\backslash x) \subset \tau(A)$. Hence we have $\tau(A) = \tau(A\backslash x)$. Then, we have $\Gamma(A\backslash x) = (A\backslash x)\backslash \bigcup_{r\in\tau(A\backslash x)} B^r = A\backslash \bigcup_{r\in\tau(A)} B^r = \Gamma(A)$. $\square$

Proof of Lemma 9

To see that $\succ^*$ is acyclic, suppose to the contrary, that is, there exists a cycle with respect to $\succ^*$, expressed as: $x^1 \succ^* x^2 \succ^* \cdots \succ^* x^L \succ^* x^1$. Note that $x'' \succ^* x'$ implies $x'' \succ^R x'$, which follows by the way these binary relations are defined. Therefore, the cycle above implies $x^1 \succ^R x^2 \succ^R \cdots \succ^R x^L \succ^R x^1$. Then, there exists a selection point x^ℓ, i.e. $x^{(q)} = x^\ell$ and $y^{(q)} = x^{\ell+1}$ for some $q \in \{1, \ldots, Q\}$. Note that by definition of the direct revealed preference, there must exist some $t \in \mathcal{T}$ such that $a^t = x^{(q)}$ and $y^{(q)} \in A^t$. Moreover, definitions of a selection point and sets $\{B^t\}_{t\in\mathcal{T}}$ imply that $x^{\ell+1} = y^{(q)} \in B^t$ for such $t \in \mathcal{T}$ and $A^t\backslash B^t \subset A^t$. Then we have $t \in \tau(A^t)$ and thus $x^{\ell+1} = y^{(q)} \in \bigcup_{r\in\tau(A^t)} B^r$, which in turn implies $x^{\ell+1} \notin \Gamma(A^t)$. Thus it is impossible to have $x^\ell = a^t \succ^* x^{\ell+1}$, and we conclude that a cycle with respect to $\succ^*$ cannot exist.

Next, we show that for every $t \in \mathcal{T}$, $a^t \in \Gamma(A^t)$. In fact, this follows immediately from AFP+CFP condition. For every $t \in \mathcal{T}$, we have $\left(\bigcup_{r\in\tau(A^t)} A^r \backslash \bigcup_{r\in\tau(A^t)} B^r\right) \subset A^t$. Then, AFP+CFP condition requires $a^t \notin \bigcup_{r\in\tau(A^t)} B^r$. Recalling the definition of Γ in (22), we have $a^t \in \Gamma(A)$ for every $t \in \mathcal{T}$. $\square$

Proof of Lemma 10

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 $\succ^* \subset \succ^R$, this cycle implies $x^1 \succ^R x^2 \succ^R \cdots \succ^R x^L \succ^R x^1$. Then there exists a selection point x^ℓ, and we have $x^{\ell+1} \in B^t$ for every $t \in \mathcal{T}$ with $a^t = x^\ell$ and $x^{\ell+1} \in A^t$. By RSM condition, there exists some $x \in A^t$ such that $x \rhd' x^{\ell+1}$, 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.

Now we show that $a^t \in \Gamma(A^t)$ for every $t \in \mathcal{T}$. Assume that a data set obeys RSM condition and Γ is defined as the set of maximal elements with respect to $\rhd'$. 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 $a^t \nsucc^R x$, while we have $a^t \succ^R x$. When a data set obeys TRSM 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 TRSM condition, $a^t \nsucc^R x$, which contradicts the assumption that $x \in A^t$. □

Appendix II: Formulation of integer programming

Here we describe how we can formulate AFP test in De Clippel and Rozen (2014) and our RSM/TRSM test as 0-1 integer programming problems. Let us denote the integer problems as $C \cdot \mathbf{x} \geqslant \mathbf{b}$, where matrix C and vector $\mathbf{b}$ are parameters determined from the data set and/or a selection profile, and vector $\mathbf{x}$ is the vector of interest. Throughout this appendix, $\mathbf{x}$ is restricted to be a 0-1 vector.

Before presenting the integer programming formulation of AFP test by De Clippel and Rozen, we note again the statement of their result.

Theorem (De Clippel and Rozen, 2014): *A data set $\mathcal{O} = \{(a^t, A^t)_{t \in \mathcal{T}}\}$ is rationalizable by AFP model if and only if there exists a binary relation $\succ^*$ on X such that*

(I) for every $s, t \in \mathcal{T}$ such that $a^s, a^t \in A^s \cap A^t$,

$$\exists x' \in A^s \backslash A^t : a^s \succ^* x' \text{ or } \exists x'' \in A^t \backslash A^s : a^t \succ^* x'', \tag{30}$$

(II) binary relation $\succ^*$ is acyclic.

In the problem $C \cdot \mathbf{x} \geqslant \mathbf{b}$, the matrix C and vector $\mathbf{b}$ are the factors for (I), and the acyclicity of $\succ^*$ is required through additional constraints on the solution vector $\mathbf{x}$. Specifically, vector $\mathbf{x} = (x_{11}, x_{21}, \ldots, x_{n1}, \ldots, x_{1n}, \ldots, x_{nn})$ is interpreted as a vector that represents binary relation $\succ^*$: $x_{ij} = 1$ if $x_i \succ^* x_j$ and $x_{ij} = 0$ otherwise. Matrix C and vector $\mathbf{b}$ are determined once data set $\mathcal{O}$ is observed. Let I be a set of index-pairs $(s, t) \in \mathcal{T} \times \mathcal{T}$ such that $a^s, a^t \in A^s \cap A^t$, and let $|I| = m$. Then, C is a matrix with m rows ($\mathbf{b}$ is an m dimensional vector), where each row (entry) represents the requirements that assure (30). Fix any row, say k-th row, and suppose that indices $s, t \in \mathcal{T}$ are the indices associated with this row. Note that the k-th row of C is an n^2 dimensional vector, which we denote as $\mathbf{c}_k = (c_{11}, c_{21}, \ldots, c_{n1}, \ldots, c_{1n}, \ldots, c_{nn})$, and b_k (k-th entry of $\mathbf{b}$) is a scalar. We omit the index k from entries of $\mathbf{c}_k$ for the sake of notational simplicity. Given a data set, $\mathbf{c}_k$ is defined so that $c_{ij} = 1$ if (i) $x_i = a^s$ and $x_j \in A^s \backslash A^t$, or (ii) $x_i = a^t$ and $x_j \in A^t \backslash A^s$; and $c_{ij} = 0$ otherwise. That is, x_i corresponds to a^s (resp. a^t), and x_j corresponds to x' (resp. x'') in (I). Then, for (30) to hold, we must have $\mathbf{c}_k \cdot \mathbf{x} \geqslant 1$, so we can set $b_k = 1$.

The additional constraints that require acyclicity of $\succ^*$ are straightforward: for every cyclic sequence of indices $J = (i, j, k, \ldots, \ell, i)$,

$$x_{ij} + x_{jk} + \cdots + x_{\ell i} \leqslant |J| - 2. \tag{31}$$

While these acyclicity constraints are easy to understand, since we must write a constraint for *every* cyclic sequence of indices, it may be computationally tough to list up: the number of constraints explodes as the number of alternatives gets larger.

Example 4. *Let $X = \{x_1, x_2, x_3, x_4\}$, and consider a data set of three observations as below.*

t	1	2	3
A^t	$\{x_1, x_2, x_3, x_4\}$	$\{x_1 x_2, x_3\}$	$\{x_2, x_3, x_4\}$
a^t	x_1	x_2	x_3

Note that we have $a^1, a^2 \in A^1 \cap A^2$ and $a^2, a^3 \in A^2 \cap A^3$. Hence the matrix C has two rows, where the first row corresponds to observations $(1, 2)$, and the second row corresponds to observations $(2, 3)$. As for observations $(1, 2)$, $A^1 \backslash A^2 = \{x_4\}$ and $A^2 \backslash A^1 = \varnothing$, so we must have $a^1 = x_1 \succ^* x_4$, and thus the c_{14} entry of $\mathbf{c}_1$ is 1. As for observations $(2, 3)$, $A^2 \backslash A^3 = \{x_1\}$ and $A^3 \backslash A^2 = \{x_4\}$, so we must have $a^2 = x_2 \succ^* x_1$ or $a^3 = x_3 \succ^* x_4$. Hence the entries c_{21} and c_{34} of $\mathbf{c}_2$ is 1.

	c_{11}	c_{21}	c_{31}	c_{41}	c_{12}	c_{22}	c_{32}	c_{42}	c_{13}	c_{23}	c_{33}	c_{43}	c_{14}	c_{24}	c_{34}	c_{44}	$\mathbf{b}$
$\mathbf{c}_1$	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
$\mathbf{c}_2$	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1

In this example with 4 alternatives, we only need to list up 25 constraints regarding acyclicity of $\succ^*$.[20] *This number will explode as the number of alternatives get larger.*

In testing RSM/TRSM, we search for a selection profile under which there exists an appropriate selection $\rhd'$ of binary relation $\rhd$. Recall that once a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ is determined, binary relation $\rhd$ is defined: $x'' \rhd x'$ if $x^t \in B^t$ for some $t \in \mathcal{T}$, $x'' \in A^t \backslash x'$, and $x' \nsucc^R x''$. We need to check if there exists an acyclic (or asymmetric and transitive) selection $\rhd'$ such that for every $x' \in B^t$, there exists $x'' \in A^t$ with $x'' \rhd' x'$. That is $\rhd'$ has to be chosen so that every alternative in B^t is dominated by some other alternative in A^t.

By nature of the problem, similar to the case of De Clippel and Rozen, it can be rephrased as the solvability of a 0-1 integer problem $C \cdot \mathbf{x} \geqslant \mathbf{b}$, and the restriction of acyclicity (asymmetry and transitivity) is required through some additional linear constraints. The solution vector $\mathbf{x} = (x_{11}, x_{21}, \ldots, x_{n1}, \ldots, x_{1n}, \ldots, x_{nn})$ is interpreted as a vector version of a selection $\rhd'$ from $\rhd$: $x_{ij} = 1$ if $x_i \rhd' x_j$, and $x_{ij} = 0$ otherwise, and matrix C is an $(nT \times n^2)$ matrix that tells us candidates of where to define $\rhd'$. More specifically, the matrix C consists of $(n \times n^2)$-matrices $\{C^t\}_{t=1}^{T}$, and the vector $\mathbf{b}$ consists of n-dimensional vectors $\{\mathbf{b}^t\}_{t=1}^{T}$. For $i \in \{1, \ldots, n\}$, i-th row of C^t and i-th coordinate of $\mathbf{b}^t$ correspond to information regarding alternative x_i at t-th observation. Denote them by $\mathbf{c}_i^t = (c_{11}, \ldots, c_{n1}, \ldots, c_{1i}, \ldots, c_{ni}, \ldots, c_{1n}, \ldots, c_{nn})$ and b_i^t. Though every entry c_{jk} of $\mathbf{c}_i^t$ depends on $t \in \mathcal{T}$ and $i \in \{1, 2, ..., n\}$, we omit them for the sake of notational simplicity. By using these notions, the problem $C \cdot \mathbf{x} \geqslant \mathbf{b}$ is equivalent to $\mathbf{c}_i^t \cdot \mathbf{x} \geqslant b_i^t$, or $\sum_{j=1}^{n} c_{ji} x_{ji} \geqslant b_i^t$ for every $t \in \mathcal{T}$ and $i \in \{1, 2, ..., n\}$.

For every $t \in \mathcal{T}$ and $i \in \{1, 2, ..., n\}$, the entries of $\mathbf{c}_i^t$ and b_i^t are set to 0 except for the following cases.

(I) Suppose that $x_j \ntriangleright x_k$. Since $\rhd'$ is defined as a selection from $\rhd$, we cannot have $x_j \rhd' x_k$, or equivalently $x_{jk} = 0$ must hold in such a case. To require this, for such a pair of indices (j, k), we let $c_{jk} = -n$.

(II) Suppose that $x_i \in B_i^t$, where B_i^t is specified by a given selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$. Then, $x_{ji} = 1$ must hold for at least one j such that $x_j \rhd x_i$. To require this, we set $c_{ji} = 1$ for all such j and $b_i^t = 1$.

Recall that for RSM model, this binary relation $\rhd'$ has to be acyclic, and for TRSM model

[20]There are 4 constraints that require asymmetry, 6 constraints regarding cycles involving two alternatives, 9 constraints regarding three-alternative cycles, and 6 regarding four-alternative cycles.

	c_{11}	c_{21}	c_{31}	c_{41}	c_{12}	c_{22}	c_{32}	c_{42}	c_{13}	c_{23}	c_{33}	c_{43}	c_{14}	c_{24}	c_{34}	c_{44}	$\mathbf{b}$
$\mathbf{c}_1^1$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0
$\mathbf{c}_2^1$	-4	-4	-4	-4	-4	-4	-4	1	0	-4	-4	-4	-4	-4	-4	-4	1
$\mathbf{c}_3^1$	-4	-4	-4	-4	-4	-4	-4	0	1	-4	-4	-4	-4	-4	-4	-4	1
$\mathbf{c}_4^1$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0
$\mathbf{c}_1^2$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0
$\mathbf{c}_2^2$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0
$\mathbf{c}_3^2$	-4	-4	-4	-4	-4	-4	-4	0	1	-4	-4	-4	-4	-4	-4	-4	1
$\mathbf{c}_4^2$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0
$\mathbf{c}_1^3$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0
$\mathbf{c}_2^3$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0
$\mathbf{c}_3^3$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0
$\mathbf{c}_4^3$	-4	-4	-4	-4	-4	-4	-4	0	0	-4	-4	-4	-4	-4	-4	-4	0

Table 4: Matrix C and vector b defined for Example 4

it has to be asymmetric and transitive. These requirements will be made as constraints on the solution vector $\mathbf{x}$. RSM model requires that binary relation $\rhd'$ is acyclic, which requires $\mathbf{x}$ to satisfy (31). TRSM model requires that binary relation $\rhd'$ is asymmetric and transitive. These two constraints are assured as follows: for every $i, j, k \in \{1, \ldots, n\}$,

$$1 - x_{ij} - x_{ji} \geqslant 0, \tag{32}$$

$$x_{ij} + x_{jk} \leqslant 2x_{ik} + 1. \tag{33}$$

Constraint (32) assures asymmetry of $\rhd'$ and (33) assures transitivity of $\rhd'$.

It is not difficult to check that, by constracting C and $\mathbf{b}$ as above, a data set is rationalizable by an RSM model if and only if there exists a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ such that the problem $C \cdot \mathbf{x} \geqslant \mathbf{b}$ has a solution $\mathbf{x}$ subject to constraint (31). A data set is rationalizable by an TRSM model if and only if there exists a selection profile $\left[x^{(q)}\right]_{q=1}^{Q}$ such that the problem $C \cdot \mathbf{x} \geqslant \mathbf{b}$ has a solution $\mathbf{x}$ subject to constraints (32) and (33).

Example 4 (continued). *Note that there are three cycles with respect to* $\succ^R$*:* $x_1 \succ^R x_2 \succ^R x_1$*;* $x_2 \succ^R x_3 \succ^R x_2$*; and* $x_1 \succ^R x_3 \succ^R x_2 \succ^R x_1$*. Let selection profile be* (x_1, x_2, x_1)*, which implies* $B^1 = \{x_2, x_3\}$*,* $B^2 = \{x_3\}$*,* $B^3 = \varnothing$*, and binary relation* $\rhd$ *is such that:* $x_4 \rhd x_2$ *and* $x_1 \rhd x_3$*. Then the* C *matrix and* b *vector is defined as in Table 4.*

This problem $C \cdot \mathbf{x} \geqslant \mathbf{b}$ *has a solution* $\mathbf{x} = (0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0)$*, where* $x_{42}, x_{13} = 1$*, and* 0 *elsewhere. This means that by setting* $x_4 \rhd' x_2$ *and* $x_1 \rhd' x_3$*, which is obviously asym-*

metric and transitive, this data set is consistent with an TRSM model.

Appendix III: Full-observation tests

Here we introduce full observation version characterizations of the limited consideration models, and describe how we adapt them to the limited data context in our simulation. The full observation characterizations are based on observation of a choice function $f : 2^X \to X$, where $f(A) \in A$ for every $A \subset X$.

AFP, CFP, and AFP+CFP models are characterized by acyclicity of a binary relation inferred from the choice function and the model: for AFP model, $x'' \succ^{AFP} x'$ if there exist $A, A' \subset X$ such that $x'' = f(A'), f(A') \neq f(A)$, and $A = A' \backslash x'$; for CFP model, $x'' \succ^{CFP} x'$ if there exist $A', A'' \subset X$ such that $f(A'') = x'', f(A') = x'$, and $\{x', x''\} \subset A'' \subset A'$; for AFP+CFP model, $x'' \succ^{AFP+CFP} x'$ if there exist $A, A', A'' \subset X$ such that $f(A'') = x'', f(A') = x', f(A') \neq f(A), A = A' \backslash x'$ and $\{x', x''\} \subset A'' \subset A'$. See Masatlioglu, Nakajima, and Ozbay (2012) for AFP, and Lleras, Masatlioglu, Nakajima, and Ozbay (2017) and its working paper version (2015) for CFP and AFP+CFP.

As shown in Manzini and Mariotti (2007), the choice function f is consistent with RSM model if and only if it satisfies,

- Weak WARP: for every $A, A', A'' \subset X$, $\{x', x''\} = A \subset A' \subset A''$ and $x'' = f(\{x', x''\}) = f(A'')$ implies $x' \neq f(A')$, and
- Expansion: for every $A, A', A'' \subset X$, $x = f(A') = f(A'')$ and $A = A' \cup A''$ implies $x = f(A)$.

Au and Kawai (2011) show that the choice function is consistent with TRSM model if and only if it satisfies Weak WARP, Expansion, and acyclicity of the following binary relation: $x'' \succ^{TRSM} x'$ if there exists $A', A'' \subset X$ such that $\{x', x''\} = A'' \subset A', x'' = f(A'')$, and $f(A') \neq f(A' \backslash x')$.

The above conditions are adapted to limited data environments as follows. Given a data set $\{(a^t, A^t)\}_{t \in \mathcal{T}}$, for model $\mathbf{M} \in \{\text{AFP, CFP, AFP+CFP}\}$, the binary relation $\succ^{\mathbf{M}}$ is defined in our context by rephrasing "there exists $A \subset X$" by "there exits $t \in \mathcal{T}$," and then we test acyclicity of this limited-data-based $\succ^{\mathbf{M}}$. For example, the binary relation in the AFP model is defined using a limited data set as follows: $x'' \succ^{AFP} x'$ if there exists $s, t \in \mathcal{T}$ such that $x'' = a^t, a^t \neq a^s$, and $A^s = A^t \backslash x''$. Similary, the conditions for testing RSM model can be molded into our context by rephrasing "for every $A \subset X$" by "for every $t \in \mathcal{T}$." For example, Weak WARP is expressed as: for every $r, s, t \in \mathcal{T}$, $\{x', x''\} = A^r \subset A^s \subset A^t$ and $x'' = a^r = a^t$ implies $x' \neq a^s$. The limited-data-based binary relation $\succ^{TRSM}$ of TRSM model is defined in a parallel fashion with AFP, CFP, and AFP+CFP models. Then we test TRSM by observing whether the data set obeys Weak WARP, Expansion, and acyclicity of this $\succ^{TRSM}$.

REMARK: For AFP, CFP, and AFP+CFP models, it is known that there are weak versions of WARP that characterize these limited consideration models. In particular, Masatligolu, Nakajima, and Ozbay (2012) show that there is an axiom WARP(LA) that is equivalent to acyclicity of $\succ^{AFP}$; Lleras, Masatlioglu, Nakajima, and Ozbay (2017) show that axiom WARP-CO is equivalent to the acyclicity of $\succ^{CFP}$; Lleras, Masatlioglu, Nakajima, and Ozbay (2015) show that axiom LC-WARP* is equivalent to the acyclicity of $\succ^{AFP+CFP}$. Since these equivalences break under a limited data set, we dealt only with the acyclicity conditions in testing these models.

References

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

[2] Beatty, T.K.M. and I. Crawford (2011): How demanding is the revealed preference approach to demand? *American Economic Review*, 101, 2782-2795.

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

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

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

[6] De Clippel, G. and K. Rozen, (2014): Bounded rationality and limited datasets. Mimeo.

[7] Lleras, J.S., Y. Masatlioglu, D. Nakajima, and E. Ozbay, (2015): When more is less: limited consideration. Mimeo (the working paper version of the reference below).

[8] Lleras, J.S., Y. Masatlioglu, D. Nakajima, and E. Ozbay, (2017): When more is less: limited consideration. *Journal of Economic Theory*, 170, 70-85.

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

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

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

[12] Selten, R. (1991): Properties of a measure of predictive success. *Mathematical Social Sciences*, 21, 153-167.

[13] Tyson, C.J. (2013): Behavioral implications of shortlisting procedures. *Social Choice and Welfare*, 41, 941-963.