Waseda Economics
Working Paper Series

Testable Implications of Decision-Making under Limited Attention

Yuta Inoue
(Graduate School of Economics, Waseda University)

Graduate School of Economics
Waseda University

1-6-1, Nishiwaseda, Shinjuku-ku, Tokyo
169-8050 JAPAN

Waseda Economics
Working Paper Series

Testable Implications of Decision-Making under Limited Attention

Yuta Inoue
(Graduate School of Economics, Waseda University)

Number 16-001
April 2016

Graduate School of Economics
Waseda University

1-6-1, Nishiwaseda, Shinjuku-ku, Tokyo

169-8050 JAPAN

Testable Implications of Decision-Making under Limited Attention

Yuta Inoue *

April 14, 2016

Abstract

This paper studies the limited attention model along the line of Masatlioglu et al. (2012). Based on observed choice behavior, our main theorem provides a necessary and sufficient condition to test the joint hypothesis that (i) the agent has limited attention, and (ii) the agent is maximizing her preference under her limited attention. As a main departure from existing studies, we allow the possibility that choices on some feasible action sets may not be observed. By applying the main theorem, we provide a robust inference on possible choices on such action sets, as well as the inference on agent's preference, attention, and inattention.

KEYWORDS: Revealed preference; Bounded rationality; Limited attention; Testable implications; Rationalizability

JEL CLASSIFICATION: D11, D81

1 Inroduction

Classical decision theory is based on the assumption that the agent is rational, in that she has a well-behaved preference, and she maximizes her preference extracting all the information available to her. While the classical theory may be a good illustration of how agents behave in general, when we focus on actual decision-making by individuals, violation of the classical theory is commonly observed.[1]

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

[1]Papers such as Huber et al. (1982), Loomes et al. (1991), and Roelofsma and Read (2000) report violations of the classical rationality assumption.

In order to account for such violations, various bounded rationality models give an illustration of how the agents are rational, but in a bounded fashion. For example, the agents may be maximizing her preference, but only within the elements that are not eliminated by some other rationale (Manzini and Mariotti, 2007), or she may be maximizing her preference only among elements in the constraint set that are superior to some reference point (Ok et al., 2015), or the agent may be maximizing her preference but her preference may be vulnerable to how the alternatives are framed (Salant and Rubinstein, 2008). Among various bounded rationality models, we focus in this paper on the model of decision-making under limited attention, due to Masatlioglu et al. (2012). In this model, it is assumed that an agent maximizes her preference, but she does not pay attention to all the alternatives that are in the constraint set.

In this paper, we provide testable implications of the model of decision-making under limited attention, that are actually empirically computable given a dataset generated by an economic agent. In particular, we provide necessary and sufficient conditions under which we can treat the agent *as if* she is maximizing her preference within her limited capacity of attention. Furthermore, we give conditions on when inference of preference, attention, and inattention can be made within the observed constraint sets. Finally, we provide conditions under which predictions of choices, attention, and inattention can be made on constraint sets that are not yet observed.

In the model of limited attention, which is due to Masatlioglu et al. (2012), we consider a consumption space with a finite number of alternatives. The agent is assumed to be boundedly rational in the sense that she has a irreflexive, complete, and transitive preference, and she aims to maximize her preference, but does not pay attention (consciously of unconsciously) to all the alternatives in the constraint sets. Within a given constraint set, the set of alternatives that the agent does pay attention to is called the *consideration set*. We cast a requirement that the consideration set must satisfy, namely the *attention filter property*, which requires that removing an ignored alternative from the constraint set does not change the consideration set.

The essential difference between the model by Masatlioglu et al. (2012) and the model we consider in this paper, is the strength of the observational assumption. Masatlioglu et al. (2012) assume that choices are observed for all possible constraint sets. In this paper we assume that choices need not be observed on all of the constraint sets. The testable conditions given in this paper is thus applicable to datasets that can actually be obtained in reality. Furthermore, this assumption allows us to make out-of-sample predictions about the agent's choices and attention, which may be one of the core interests of economists. In fact, de Clippel and Rozen (2014) study the same model as this paper, and they show that the acyclicity of some binary relation is a necessary and sufficient condition to treat an agent as if she is maximizing her preference under limited attention. In this paper we explicitly construct such binary relation. Therefore the results presented in this paper are applicable to general forms of datasets, and the binary relations are explicitly defined from the dataset.

The rest of this paper is arranged as follows. In section 2, we introduce in detail the model of decision-making under limited attention, and in section 3 we provide necessary and sufficient condition for the rationalization of an observed dataset. In section 4, we study the issue of robust inference of preference, attention, and inattention, provided that the dataset is rationalizable. In section 5, we show how out-of-sample prediction of choices, attention, and inattention can be made, given a dataset that is rationalizable. The appendix contains some proofs and additional examples.

2 The model

Let X be finite set which we consider as the grand space of alternatives, and denote by $\Omega := 2^X \backslash \{\emptyset\}$ the collection of all nonempty subsets of X. An element $S \in \Omega$ is interpreted as a constraint set over which the agent makes her choice, and Ω is a collection of all possible constraint sets. In this paper, we assume that the researcher observed choices made by the agent only on some of the possible constraint sets. Denote by $\mathcal{D}$ the collection of constraint sets on which the choice of the agent is observed. That is, we assume that $\mathcal{D} \subseteq \Omega$. For each constraint set $S \in \mathcal{D}$, we observe the choice made by a single agent when she faces the constraint set S. That is, we observe the *choice function* of an agent, which is denoted by $c : \mathcal{D} \to X$ such that $c(S) \in S$ for all $S \in \mathcal{D}$.

Summarizing, *the observed dataset*, which is denoted by $(c, \mathcal{D})$, is the pair of constraint sets and the choices on those constraint sets that are actually observed by the researcher.

While we observe choices only over sets in $\mathcal{D}$, we assume that the agent has a complete, transitive, and irreflexive preference over all the alternatives in X, which is unobservable to the researcher.

Further assume that the agent is making decisions under *limited attention.* That is, when she faces a constraint set $S \in \mathcal{D}$, we allow that she may not be aware of all the alternatives in S. In order to formulate this, we assume that an agent has a mapping $\Gamma : \Omega \to \Omega$ such that $\Gamma(S) \subseteq S$ for all $S \in \Omega$, which we call *the consideration set mapping.* When facing a constraint set S, $\Gamma(S)$ is the set of alternatives that the agent takes into consideration when making a choice. The alternatives in $S \backslash \Gamma(S)$ may be ignored consciously or unconsciously. Note that the consideration set mapping is not observable to the researcher.

To avoid trivial cases, we cast a property that the consideration set mapping must obey, which we call the *attention filter* property.[2] This is defined as;

$$\Gamma(S) \subseteq T \subseteq S \;\Rightarrow\; \Gamma(S) = \Gamma(T) \text{ for all } S, T \in \Omega.$$

In words, taking away ignored alternatives must not alter the consideration set.

One of the main questions of this paper is: given a dataset of choices made by an agent, when can we say that the agent is rationally making decisions under limited attention? That is, when is the agent maximizing her preference within the limited alternatives that she pays attention to? In answering this question, we define what is meant in this model when we say that the dataset is *rationalizable under limited attention.*

Definition 1. The dataset $(c, \mathcal{D})$ is *rationalizable under limited attention (LA-rationalizable)* if there exist a consideration set mapping $\Gamma : \Omega \to \Omega$ and a complete, transitive, and irreflexive preference $\succ$ such that (i): Γ is an attention filter, and (ii): $c(S) \succ x$ for all $x \in \Gamma(S) \backslash \{c(S)\}$ for all $S \in \mathcal{D}$.

Given a dataset made by an agent, suppose we can find a strict preference and consideration set mapping such that the agent is observed to have chosen

[2] Without the attention filter property, it may be possible to assume that the agent pays attention only to the alternative that she chooses, i.e. $\Gamma(S) = \{c(S)\}$ for all constraint sets S. Then the problem is degenerate. The attention filter property excludes such trivial cases.

the most preferred alternatives within the consideration sets. This means that it is possible to treat the agent *as if* she is rationally making decisions under her limited attention. Then it will be possible to make inference about her preference, attention, and inattention within the observed constraint sets, and make predictions about her choices, attention, and inattention on constraint sets that are not yet observed.

Note that the concept of LA-rationalizability is weaker than the rationalizability in the classical sense. In Example 3 in the Appendix, we show that a dataset can be LA-rationalizable while it is not rationalizable in the classical sense. We also show in the example that there may be multiple pairs of preference and consideration set mapping that LA-rationalize the dataset.

In the remainder of this paper, we first start from providing the conditions under which the dataset is LA-rationalizable. Then, assuming that the dataset is LA-rationalizable, we consider the problem of robust inference of preference, attention, and inattention on the constraint sets that are observed. Finally, we consider how we can predict the alternatives that will be chosen, and the alternatives that attract (or does not attract) attention in constraint sets that are not observed in the dataset, provided that the dataset is LA-rationalizable.

3 LA-rationalization with partial observation

In this section, we answer the following questions: given a dataset $(c, \mathcal{D})$, (i) what conditions must the dataset satisfy when it is LA-rationalizable? (ii) under what conditions can we say that an agent is behaving as if she is rational subject to limitation of attention? That is, we provide conditions that are necessary and sufficient for the LA-rationalizability of the dataset.

In doing this, we first consider necessary conditions of LA-rationalizability. Therefore, the discussion in this section will be based on the assumption that the observed dataset $(c, \mathcal{D})$ is LA-rationalizable.

Now assume that the pair of preference and consideration set mapping $(\succ, \Gamma)$ LA-rationalizes the dataset $(c, \mathcal{D})$. Note that the consideration set mapping Γ is an attention filter, and the observed choices are maximizers of preference $\succ$ within the consideration sets. First we consider what is implied by this.

Assume that we observe choices on constraint sets $S, S\backslash\{a\} \in \mathcal{D}$, the choices are different in those constraint sets, and the choice in S is different from the removed alternative a, i.e. $c(S) \neq c(S\backslash\{a\})$ and $c(S) \neq a$. Then it follows that the alternative a must attract attention in S, and thus it follows that $c(S)$ is preferred to a. The logic is as follows. Suppose by way of contradiction that the alternative a does not attract attention in S, i.e. $a \notin \Gamma(S)$. Then by the attention filter property, we must have $\Gamma(S) = \Gamma(S\backslash\{a\})$. Since $(\succ, \Gamma)$ LA-rationalize the dataset and $c(S) \neq c(S\backslash\{a\})$, we must have $c(S) \succ c(S\backslash\{a\})$ and $c(S\backslash\{a\}) \succ c(S)$, which is a contradiction.[3] Therefore we conclude that $a \in \Gamma(S)$, and thus $c(S) \succ a$.

Generalizing the discussion, consider constraint sets $S, T \in \mathcal{D}$ such that the choices in the constraint sets differ and are both in the intersection of the constraint sets, i.e. $c(S) \neq c(T)$ and $c(S), c(T) \in S \cap T$. Then there must exist "$a \in S\backslash T$ such that $c(S) \succ a$" and / or "$b \in T\backslash S$ such that $c(T) \succ b$," following the same logic as the case considered in the paragraph above.

Since this is a condition that is implied by the LA-rationalizability of the dataset, a necessary condition of LA-rationalizability is that there exists a "virtual preference" of the agent, a binary relation that we denote by $\tilde{P}$ such that

$$\text{for every } S, T \in \mathcal{D} \text{ such that } c(S) \neq c(T) \text{ and } c(S), c(T) \in S \cap T,$$
$$\text{we have } \begin{cases} c(S)\tilde{P}y \text{ for some } y \in S\backslash T, \text{ or} \\ c(T)\tilde{P}z \text{ for some } z \in T\backslash S. \end{cases} \tag{1}$$

The interest of this paper lies on how we can actually test whether or not the choices made by an agent is in line with the theory on decision-making under limited attention. Hence we explicitly construct a binary relation $\tilde{P}$ that obeys (1), and show that the acyclicity of the relation in combination with an additional condition is necessary and sufficient for the LA-rationalizability of the dataset.[4]

[3] Note that we trivially have $c(S) \in \Gamma(S)$ and $c(S\backslash\{x\}) \in \Gamma(S\backslash\{x\})$.

[4] In fact, de Clippel and Rozen (2014) show that the acyclicity of such binary relation $\tilde{P}$ is a necessary and sufficient condition for the LA-rationalizability of the dataset.

In the following subsection, we construct a binary relation $\tilde{P}$ that obeys (1) and is acyclic, under the assumption that the dataset $(c, \mathcal{D})$ is LA-rationalizable. In addition to these two conditions, we cast an additional condition that the revealed preference relation must satisfy, provided that the dataset is LA-rationalizable. The construction will be conducted in three steps. We first define the "revealed preference" P, a binary relation that actually reflects the agent's preference when the dataset is LA-rationalizable. In the remaining two steps, we define binary relations P^1 and P^2. These binary relations are defined so that the union of all binary relations obeys the restriction (1).

Given any binary relation R, we denote by R_T the transitive closure of the binary relation R.

3.1 Constructing the binary relation $\tilde{P}$

In this subsection, we construct an acyclic binary relation that obeys restriction (1), under the assumption that the dataset is LA-rationalizable. Hence, throughout this subsection, we assume that the dataset $(c, \mathcal{D})$ is LA-rationalizable, i.e., we treat the agent as if she has a consideration set mapping that obeys the attention filter property, she has a strict preference, and the observed choices are a result of maximizing her preference over the alternatives within her consideration.

First of all, we consider how we can infer the agent's preference from the observed dataset. Take any consideration set mapping Γ and any strict preference $\succ$ that LA-rationalize the dataset. Assume that we observe choices made on constraint sets $S, S\backslash\{a\} \in \mathcal{D}$, and the choices made on those constraint sets are different; $c(S) \neq c(S\backslash\{a\})$ where $c(S) \neq a$. Then we can infer that $a \in \Gamma(S)$, and thus $c(S)$ is preferred to a.

Since the consideration set mapping and preference were arbitrary, we can surely say that $c(S)$ is preferred to a if there exist $S, S\backslash\{a\} \in \mathcal{D}$ such that $c(S) \neq c(S\backslash\{a\})$ and $c(S) \neq a$. Thus we define a "revealed preference" on X as follows:

$$\begin{aligned}
&\text{for any distinct } x, y \in X,\ x \textit{ is revealed preferred to } y \\
\iff\ & \text{there exist } S, S\backslash\{y\} \in \mathcal{D} \text{ such that } x = c(S) \neq c(S\backslash\{y\}). \qquad (2)
\end{aligned}$$

For the time being, denote xPy when there are $x, y \in X$ that obey the relation (2).[5] With a slight abuse of terminology, we call this a "revealed preference" of the agent. In the classical decision theory, where there is an implicit assumption that the agent is aware of all the alternatives in the constraint set, the revealed preference relation was straightforward. However, in the case of the limited attention model, the issue is subtle. The fact that an alternative x is chosen when y is available does not immediately reveal that x is preferred to y. This is because the agent may not have realized the existence of alternative y when choosing x. Nevertheless, whenever the restriction in (2) is satisfied, we can surely infer that the alternative x is preferred to y. Thus we use the term "revealed preference," while it is essentially a robust inference of preference. The issue of robust inference of preference (and attention / inattention) will be discussed in detail in the following section.

The revealed preference defined here is identical to that defined in Masatlioglu et al. (2012). However, we emphasize that the difference in the observational assumption cannot be ignored. In the model of Masatlioglu et al. (2012), it is assumed that the choices are observed for all possible constraint sets. This means that the researcher observes both $c(S)$ and $c(S\backslash\{x\})$ for all $S \in \Omega$ and $x \in S$. However, under the assumption of partial observation, we may observe the choice from some constraint set S, while the choice from the constraint set $S\backslash\{x\}$ may not be observed. Thus we can extract less information from the dataset under the partial observation assumption.

While the dataset reveals less information about the agent's preference in the partial observation model, we can extract a bit more information using the attention filter property. Here we expand the revealed preference relation P, by adding to it some more pairs of elements. Suppose that we observe choices from constraint sets $S, T \in \mathcal{D}$, where $c(S) \neq c(T)$ and $c(S), c(T) \in S \cap T$. Suppose further that there is no element in $T\backslash S$ that does not dominate $c(T)$ with respect to the transitive closure of revealed preference, denoted by P_T, and there is only one element $a \in S\backslash T$ that does not dominate $c(S)$. That is, $\{y \in T\backslash S : yP_Tc(T) \text{ does not hold}\} = \emptyset$ and $\{y \in S\backslash T : yP_Tc(S) \text{ does not hold}\} = \{a\}$. Since the revealed preference reflects the agent's preference under the LA-rationalizability of the dataset, it follows that none of the elements in $T\backslash S$ can attract attention in T, and a is

[5] A formal construction of the "revealed preference" P will be given in Algorithm 1 below.

the only element in $S\backslash T$ that can attract attention in S. Then, the attention filter property requires that a must attract attention in S, which in turn implies that the agent must prefer $c(S)$ to a. Thus, in a case as above, we can infer that the agent prefers $c(S)$ to a. Therefore, we can expand the revealed preference relation. A formal construction is given below.

ALGORITHM 1. *Construct revealed preference relation P*

Input: A finite set X and a collection of sets $\mathcal{D} \subseteq 2^X \backslash \{\emptyset\}$

Output: A binary relation P

1. Set $P = \emptyset$, and go to 2.
2. For all $B, T \in \mathcal{D}$, define sets;

$$E(S,T) = \{x \in S\backslash T : c(S) P_T x\}, \tag{3}$$

$$\bar{E}(S,T) = \{x \in S\backslash T : x P_T c(S) \text{ does not hold}\}. \tag{4}$$

3. Define a binary relation R such that xRy if;

$$\text{there exist } S, T \in \mathcal{D} \text{ such that } \begin{cases} x = c(S) \neq c(T), \\ c(S), c(T) \in S \cap T, \\ E(S,T) \cup E(T,S) = \emptyset, \\ \bar{E}(T,S) = \emptyset, \text{and} \\ \bar{E}(S,T) = \{y\}. \end{cases} \tag{5}$$

4. If $P = P \cup R$, stop. Otherwise, define $P = P \cup R$, and go to 2.

Note that the first round of the algorithm is searching for $x, y \in X$ that satisfy the relation (2). In the paragraph proceeding Algorithm 1, a part of the second round of the algorithm was illustrated. Given the revealed preference relation P after the first round, the algorithm will add $(c(S), a)$ to the revealed preference P. Under LA-rationalizability of the observed choice function, it follows that xP_Ty implies that alternative x is preferred to y. Then, given the expanded revealed preference relation P, we look for constraint sets that satisfy (5), and add pairs (x, y) to the P. Iterating this procedure, we have a revealed preference relation P.[6]

[6]Since X and $\mathcal{D}$ are finite, it is obvious that Algorithm 1 terminates after a finite number of

In the remaining parts of the paper, for constraint sets $S, T \in \mathcal{D}$, the sets $E(S,T)$ and $\bar{E}(S,T)$ expressed in (3) and (4) will be defined using the revealed preference P, which is the output of Algorithm 1. The set $E(S,T)$ is the set of alternatives in $S \backslash T$ that are dominated by the chosen alternative $c(S)$ with respect to the revealed preference P_T. The set $\bar{E}(S,T)$ is the set of alternatives that do not dominate $c(S)$ with respect to P_T.

Since the binary relation P reflects the agent's preference, it follows that P is acyclic under LA-rationalizability of the dataset. However, we show in Example 4 in the Appendix that the acyclicity of P is not sufficient for the LA-rationalizability of c. That is, even when the revealed preference relation P is acyclic, we may not find a preference and consideration set mapping that LA-rationalizes the dataset $(c, \mathcal{D})$.

Example 4 tells us that even if the revealed preference relation P is acyclic, there are cases where the dataset $(c, \mathcal{D})$ is not LA-rationalizable. Therefore, in addition to the acyclicity of P, we need a further requirement.

Definition 2 (Condition LA). The dataset $(c, \mathcal{D})$ satisfies *condition (LA)* if for all $S, T \in \mathcal{D}$

$$\bar{E}(S,T) \cup \bar{E}(T,S) \neq \emptyset \text{ whenever } \begin{cases} c(S) \neq c(T), \\ c(S), c(T) \in S \cap T, \text{ and} \\ E(S,T) \cup E(T,S) = \emptyset. \end{cases} \tag{6}$$

In words, this condition can be explained as follows. The first two lines of the right-hand side of (6) means that there is some kind of choice reversal between the choices in constraint sets S and T. Condition (LA) requires that this kind of choice reversal must be caused because there is some element in $S \backslash T$ or $T \backslash S$ that attracted attention, and as a result, the agent did not pay attention to the choice in T when she faced constraint set S (or vice versa, or both).

It can be shown that condition (LA) is satisfied whenever the dataset $(c, \mathcal{D})$ is LA-rationalizable.

Fact 1. *Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Then the dataset satisfies condition (LA).*

steps.

Now we know that the acyclicity of the revealed preference P and condition (LA) are necessary conditions for the LA-rationalizability of the dataset. However, the Example 5 in the appendix shows that these two requirements are not sufficient. Therefore we need additional restrictions on the set of alternatives X for the condition (1) to be satisfied. Below, we define two additional binary relations on X, namely P^1 and P^2, so that the union of these binary relations with P_T, denoted by $\tilde{P} := (P_T \cup P^1 \cup P^2)$, obeys condition (1). We emphasize here that the binary relations P^1 and P^2 defined below are merely "binary relations" and not "revealed preferences." That is, even if we have xP^1y for alternatives x and y, this does not imply that we can surely infer the agent prefers x to y. These binary relations are defined in order to make $\tilde{P}$ satisfy (1), and be acyclic under LA-rationalizability of the dataset.

In order to have condition (1) satisfied, we need to define additional binary relations over alternatives in constraint sets $S, T \in \mathcal{D}$ such that

$$\begin{cases} c(S) \neq c(T), \\ c(S), c(T) \in S \cap T, \\ E(S,T) \cup E(T,S) = \emptyset, \text{ and} \\ \bar{E}(S,T) \cup \bar{E}(T,S) \neq \emptyset. \end{cases} \tag{7}$$

On these sets, the condition (1) is not satisfied by the revealed preference P_T. Thus the binary relations defined below will focus on the constraint sets as above. Note that the set $\bar{E}(S,T)$ is the set of "candidate alternatives" to be dominated by the observed choice $c(S)$ in order to satisfy condition (1).

First we begin by defining some sets that we shall use in defining the binary relation P^1. We denote by $\mathcal{F}$ the set of pairs of constraint sets that are relevant in defining the binary relation P^1, i.e. the pairs that satisfy (7). Denote by F the set of alternatives that are elements of the constraint sets in $\mathcal{F}$. We further define sets F^0, F^1, and F^2, which are subsets of F. Details are given below;

$$\begin{aligned}
\mathcal{F} &:= \{\{S,T\} \subseteq \mathcal{D} : (7) \text{ holds for } S,T\},\\
F &:= \{a \in X : \text{ there exists } \{S,T\} \in \mathcal{F} \text{ with } c(S)=a \text{ or } a \in \bar{E}(S,T)\},\\
F^0 &:= \{a \in F : \text{ there does not exist } \{S,T\} \in \mathcal{F} \text{ with } a \in \bar{E}(S,T)\},\\
F^1 &:= \{a \in F : \text{ there exists } \{S,T\} \in \mathcal{F} \text{ with } c(S)=a\}, \text{ and}\\
F^2 &:= F \backslash F^1.
\end{aligned}$$

In words, the set F^1 consists of alternatives that are observed to be chosen in the sets that are relevant in $\mathcal{F}$. The set F^0 consists of elements that are not a candidate alternative to be dominated by any alternative in F^1. The set F^2, on the contrary, consists of the elements in F that are never observed to be chosen in the constraint sets relevant in $\mathcal{F}$.

In addition to these sets, we define for each alternative x in F, the set of alternatives that are revealed preferred to x, and the set of alternatives that are revealed worse than x with respect to the revealed preference relation P_T. They will be denoted by $RP(x)$ and $RW(x)$ respectively. Formally for all $x \in F$;

$$\begin{aligned}
RP(x) &:= \{a \in F : aP_T x\},\\
RW(x) &:= \{a \in F : xP_T a\}.
\end{aligned}$$

Recall that for $\{S,T\} \in \mathcal{F}$, we need either $c(S)P^1 a$ for some $a \in S\backslash T$, or $c(T)P^1 b$ for some $b \in T\backslash S$. However, there are subtle cases where we have to be careful which alternative we choose to be dominated. For example, suppose $\{S^1,S^2\},\{S^3,S^4\} \in \mathcal{F}$ are such that $a \in \bar{E}(S^1,S^2), c(S^1) \in \bar{E}(S^3,S^4)$, and $aP_T c(S^3)$. Then if we set $c(S^1)P^1 a$, it will follow that $c(S^1)P^1 aP_T c(S^3)$, i.e. $c(S^1)$ dominates $c(S^3)$ with respect to the binary relation $(P_T \cup P^1)$. This means that $c(S^3)$ loses a candidate alternative to dominate.

Hence we consider cases where we can define P^1 without such subtle issues. The first case is when there is some $x \in F^0$ such that $RP(x) \cap (F\backslash F^0) = \emptyset$. It follows from $x \in F^0$ that we do not need x to be dominated by any other element to satisfy (1), and $RP(x) \cap (F\backslash F^0) = \emptyset$ implies that x is not dominated by any other element in F apart from those in F^0. Thus even if we set $xP^1 y$ for all $y \in F\backslash F^0$, subtleties as discussed in the previous paragraph will not arise.

Note that $x \in F^0$ means that there exists $\{S, T\} \in \mathcal{F}$ such that $x = c(S)$, and any element in $\bar{E}(S, T)$ is not member of the set F^0, so $c(S)P^1a$ will hold for all $a \in \bar{E}(S, T)$.

Another case where we can set P^1 without complication is when there exists an element $a \in F^2$ with $RW(a) \cap F^1 = \emptyset$. It follows from $a \in F^2$ that a does not have to dominate any other element in F in order for (1) to be satisfied, and $RW(a) \cap F^1 = \emptyset$ implies that a does not dominate any element that may have to dominate another element for (1) to be satisfied. Hence even if we set $c(S)P^1a$ for $\{S, T\} \in \mathcal{F}$ with $a \in \bar{E}(S, T)$, subtle issued discussed above will not arise.

Formally, we define the binary relation P^1 on F as follows: xP^1y if there exists $\{S, T\} \in \mathcal{F}$ that obeys (α) or (β), where

$$(\alpha) \begin{cases} x = c(S) \in F^0, \\ RP(x) \cap (F \backslash F^0) = \emptyset, \text{ and} \\ y \in \bar{E}(S, T), \end{cases} \qquad (\beta) \begin{cases} x = c(S), \\ y \in \bar{E}(S, T) \cap F^2, \text{ and} \\ RW(y) \cap F^1 = \emptyset. \end{cases}$$

Note that after defining the binary relation P^1, if the condition (1) is satisfied by the binary relation $(P_T \cup P^1)$, then we have a relation that obeys (1). We show in the following fact that this binary relation is acyclic under the LA-rationalizability of the dataset $(c, \mathcal{D})$.

FACT 2. *If the dataset $(c, \mathcal{D})$ is LA-rationalizable, then the binary relation $(P \cup P^1)$ is acyclic on X.*

We still need to define an additional binary relation that satisfies condition (1) on the pairs of constraint sets $\{S, T\} \in \mathcal{F}$ that satisfy neither (α) nor (β). We denote by $\mathcal{J}$ the collection of such pairs, i.e. $\mathcal{J} := \{\{S, T\} \in \mathcal{F} :$ neither (α) nor (β) holds $\}$. These pairs are the pairs where there arise subtleties when attempting to define a binary relation so that (1) is satisfied. Therefore, we go through an algorithm that defines a binary relation that satisfies condition (1) within the pairs in $\mathcal{J}$, and does not contradict the existing binary relations. For simplicity, let us denote $\mathcal{J} = \{\{S^1, T^1\}, \dots, \{S^m, T^m\}\}$.

ALGORITHM 2. *Define binary relation P^2*

Input: A collection of sets $\mathcal{J}$ and the binary relations Q and P^1.

Output: A binary relation P^2.

1. Define sets for $i = 1, \dots, m$;

$$K(S^i, T^i) := \{(x, a) : x = c(S^i), a \in \bar{E}(S^i, T^i)\},$$
$$K(T^i, S^i) := \{(x, a) : x = c(T^i), a \in \bar{E}(T^i, S^i)\},$$
$$K^i := K(S^i, T^i) \cup K(T^i, S^i), \text{ and}$$
$$L = \emptyset.$$

2. Take $(x^i, a^i) \in K^i$ for $i = 1, \dots, m$ such that $\{(x^i, a^i)\}_{i=1,\dots,m} \notin L$. If this is not possible, go to 3. Otherwise, set $x^i P^2 a^i$ for $i = 1, \dots, m$, and go to 4.
3. Take any $\{(x^i, a^i)\}_{i=1,\dots,m} \in L$, set $x^i P^2 a^i$ for $i = 1, \dots, m$, and stop.
4. If the binary relation $(P \cup P^2)$ is acyclic in X, stop. Otherwise, set $L = L \cup \{(x^i, a^i)\}_{i=1,\dots,m}$, and go to 2.

Recall that for condition (1) to be satisfied, for any $\{S^i, T^i\} \in \mathcal{J}$, we need "$c(S^i)P^2 a$ for some $a \in S^i \backslash T^i$" or "$c(T^i)P^2 b$ for some $b \in T^i \backslash S^i$." The collection of all such pairs is K^i, for $i = 1, \dots, m$.[7] We take one pair (x^i, a^i) from each K^i, and check whether setting $x^i P^2 a^i$ for $i = 1, \dots, m$ contradicts the existing binary relation $(P \cup P^1)$. In the algorithm, we are essentially checking all the combinations of $\{(x^1, a^1), \dots, (x^m, a^m)\}$ until we find one that does not contradict $(P \cup P^1)$. Since the sets X and $\mathcal{J}$ are finite, it is obvious that the algorithm terminates after a finite number of steps.

Below we show that the binary relation $\tilde{P} = (P_T \cup P^1 \cup P^2)$ obeys condition (1) whenever the dataset satisfies condition (LA).

FACT 3. *If the dataset $(c, \mathcal{D})$ satisfies condition (LA), the binary relation $\tilde{P}$ satisfies condition (1).*

Below we show that Algorithm 2 generates a binary relation P^2 that does not contradict the binary relation $(P \cup P^1)$ whenever the dataset $(c, \mathcal{D})$ is LA-rationalizable. In doing this, we present the following fact.

[7]The set K^i will be nonempty for all i if the dataset obeys condition (LA), which will be satisfied under the LA-rationalizability of the dataset. On the other hand, if condition (LA) is violated, the dataset is not LA-rationalizable, so we do not have to go through Algorithm 2, in testing whether the dataset is LA-rationalizable or not. That is why we ignore the case when K^i is empty for some i.

Fact 4. *Let* $(c, \mathcal{D})$ *be LA-rationalizable. If the binary relation* $(P \cup P^2)$ *is acyclic, then the binary relation* $(P \cup P^1 \cup P^2)$ *is acyclic.*

Now we show that such P^2 can be defined whenever the dataset $(c, \mathcal{D})$ is LA-rationalizable. Let the dataset be LA-rationalizable. Then there exist a strict ordering and consideration set mapping $(\succ, \Gamma)$ which we can interpret to be the agent's virtual preference and attention. Note that under LA-rationalizability of the dataset, the revealed preference P actually reflects the agent's true preference, i.e. xPy implies that $x \succ y$. We know that the preference $\succ$ obeys condition (1), so for any $\{S, T\} \in \mathcal{J}$, "there exists $a \in S \backslash T$ such that $c(S) \succ a$" or "there exists $b \in T \backslash S$ such that $c(T) \succ b$". Now consider $\succ^* \subset \succ$ such that $\succ^* = \{(x^1, a^1), \ldots, (x^k, a^k)\}$, where $(x^i, a^i) \in K^i$ for $i = 1, \ldots, k$. Then the binary relation $(P \cup \succ^*)$ is obviously acyclic. Hence,when the algorithm sets $P^2 = \succ^*$, the fact above tells that the binary relation $(P \cup P^1 \cup P^2)$ is acyclic, and thus the algorithm succeeds.

Summarizing, under the LA-rationalizability of the dataset $(c, \mathcal{D})$, the revealed preference relation P obeys condition (LA), and the binary relation $\tilde{P}$ defined by $\tilde{P} = (P_T \cup P^1 \cup P^2)$ is acyclic. Moreover, by definition of these binary relations, $\tilde{P}$ satisfies condition (1). Therefore, condition (LA) and the acyclicity of $\tilde{P}$ is a necessary condition of the LA-rationalizability of the dataset $(c, \mathcal{D})$. In fact, this is a sufficient condition as well. This is the main theorem of this paper.

Theorem 1. *The dataset* $(c, \mathcal{D})$ *is LA-rationalizable if and only if the dataset* $(c, \mathcal{D})$ *obeys condition (LA) and the binary relation* $\tilde{P}$ *is acyclic.*

(Proof) The proof of necessity is already presented in the discussion preceding the theorem. While the proof of sufficiency is analogous to the proof in de Clippel and Rozen (2014), we will present it here, since the logic in the proof will be used in the proofs of propositions in the later sections.

Assume that the dataset $(c, \mathcal{D})$ obeys condition (LA) and the binary relation $\tilde{P}$ is acyclic. By Fact 3, condition (LA) implies that the binary relation $\tilde{P}$ satisfies condition (1). The acyclicity of $\tilde{P}$ implies that there exists a complete ordering extension, which we denote by $\succ$.[8] Note that the complete ordering $\succ$ satisfies condition (1) as well, since $\tilde{P} \subseteq \succ$. Now we define the consideration

[8] Andrikopoulos (2009) gives a comprehensive survey on extension theorems relevant to economics.

set mapping $\Gamma : \Omega \to \Omega$ as follows. For $S \in \mathcal{D}$;

$$\Gamma(S) = \{c(S)\} \cup \{x \in S : c(S) \succ x\}, \tag{8}$$

and for $S \notin \mathcal{D}$, the consideration set mapping is defined as;

$$\Gamma(S) = \begin{cases} \Gamma(T) & \text{if there exists } T \in \mathcal{D} \text{ such that } \Gamma(T) \subseteq S \subseteq T, \\ S & \text{otherwise.} \end{cases} \tag{9}$$

It is clear that $\Gamma(S) = \emptyset$ for all $S \in \Omega$. It is also clear that, for $S \in \mathcal{D}$, $c(S) \in \Gamma(S)$ and $c(S)$ is the unique preference maximizer in $\Gamma(S)$, under preference $\succ$.

Now we show that the consideration set mapping Γ is well-defined. Assume by way of contradiction that the consideration set mapping Γ is not well-defined for some constraint set $S \notin \mathcal{D}$. This means that there exist $T, T' \in \mathcal{D}$ such that $S \subseteq T, T'$ and $\Gamma(T), \Gamma(T') \subseteq S$, but $\Gamma(T) \neq \Gamma(T')$. This implies that $c(T) \neq c(T')$.[9] Now consider any $y \in T \backslash T'$. Since $S \subseteq T'$, we have $y \in T \backslash S$, which in turn implies $y \in T \backslash \Gamma(T)$. Hence $y \succ c(T)$ holds for all $y \in T \backslash T'$. Following an analogous logic, $z \succ c(T')$ for all $z \in T' \backslash T$. Hence condition (1) is violated, which is a contradiction.

Finally, we show that the consideration set mapping Γ obeys the attention filter property. That is, we prove in the four possible cases that $\Gamma(S) = \Gamma(S \backslash \{x\})$ for any $S \in \Omega$ and $x \in S \backslash \Gamma(S)$. In the proof below, we fix a constraint set $S \in \Omega$ and an alternative $x \in S \backslash \Gamma(S)$.

[*Case 1*: $S \backslash \{x\}, S \in \mathcal{D}$] Since we have $S \in \mathcal{D}$ and $x \notin \Gamma(S)$, it follows that $x \succ c(S)$. Now suppose that $\Gamma(S) \neq \Gamma(S \backslash \{x\})$. This implies that $c(S) \neq c(S \backslash \{x\})$. By focusing on condition (1) over the sets S and $S \backslash \{x\}$, we must have $c(S) \succ x$, which is a contradiction. Thus we conclude that $\Gamma(S) = \Gamma(S \backslash \{x\})$.

[*Case 2*: $S \backslash \{x\} \in \mathcal{D}$ and $S \notin \mathcal{D}$] The set $S \backslash \{x\} \in \mathcal{D}$ implies that $\Gamma(S \backslash \{x\}) = \{c(S \backslash \{x\})\} \cup \{y \in S \backslash \{x\} : c(S \backslash \{y\})\}$. Since $S \backslash \Gamma(S) \neq \emptyset$ (we have $x \in S \backslash \Gamma(S)$), there exists a constraint set $T \in \mathcal{D}$ such that $\Gamma(T) \subseteq S \subseteq T$.[10]

[9] If not, we have $\Gamma(T) = \{c(T)\} \cup \{x \in T : c(T) \succ x\}$. Since $\Gamma(T) \subseteq S$, $\{x \in T : c(T) \succ x\} = \{x \in S : c(T) \succ x\}$. Similarly, we have $\{x \in T' : c(T') \succ x\} = \{x \in S : c(T') \succ x\}$. Summarizing, it must follow that $\Gamma(T) = \Gamma(T')$.

[10] Recall from (9) that if there does not exist such $T \in \mathcal{D}$, we have $\Gamma(S) = S$, and $S \backslash \Gamma(S) = \emptyset$.

Note that we have $z \succ c(T)$ for all $z \in T\backslash S$, and that $\Gamma(S) = \Gamma(T)$ implies $x \in T\backslash\Gamma(T)$. This in turn implies that $x \succ c(T)$. Summarizing we have $y \succ c(T)$ for all $y \in T\backslash(S\backslash\{x\})$. If we suppose $\Gamma(S\backslash\{x\}) \neq \Gamma(S) = \Gamma(T)$, then it must follow that $c(S\backslash\{x\}) \neq c(T)$. This contradicts condition (1), and thus $\Gamma(S\backslash\{x\}) = \Gamma(S)$ must hold.

[*Case 3*: $S\backslash\{x\} \notin \mathcal{D}$ and $S \in \mathcal{D}$] Since $S \in \mathcal{D}$, by (8), $\Gamma(S) = \{c(S)\} \cup \{y \in S : c(S) \succ y\}$. Since $x \in S\backslash\Gamma(S)$, we have $\Gamma(S) \subseteq S\backslash\{x\}$. Therefore, we have by (9), $\Gamma(S\backslash\{x\}) = \Gamma(S)$.

[*Case 4*: $S\backslash\{x\}, S \notin \mathcal{D}$] Since $S\backslash\Gamma(S) \neq \emptyset$, there exists a constraint set $T \in \mathcal{D}$ such that $\Gamma(T) \subseteq S \subseteq T$. Since we have $x \in S\backslash\Gamma(S)$, it follows that $\Gamma(T) = \Gamma(S) \subseteq S\backslash\{x\}$. Therefore, by (9), we have $\Gamma(S\backslash\{x\}) = \Gamma(T) = \Gamma(S)$. QED

This theorem tells us that given a dataset of constraint sets and observed choices $(c, \mathcal{D})$, it suffices to check whether or not the revealed preference P obeys condition (LA) and the binary relation $\tilde{P}$ is acyclic. If the two conditions are satisfied, then we *may* say that the agent chose the observed elements as a result of preference maximization under limited attention. On the other hand, if there is a violation of either of the conditions, then we can definitely say that the agent is not a preference maximizer under limited attention.

The main difference between this result and the existing studies is that this result can be applied to data obtained in reality. We have assumed observation of the choice function on a general dataset, and we explicitly construct a binary relation from the observed dataset. Therefore, the result in Theorem 1 is meaningful, in that we provide a condition that is computable given any dataset.

4 Inference of preference, attention, and inattention

In this section, we discuss how inference of preference, attention, and inattention can be made under the LA-rationalizability of the dataset. More specifically, given a dataset that is LA-rationalizable, we give conditions on when we can surely say that an alternative (i): is preferred to another, (ii): is paid attention to in some constraint set, and (iii): is not paid attention to in some constraint set. Throughout this section, we assume that the dataset $(c, \mathcal{D})$ is LA-rationalizable. Note that the preference and consideration set mapping that LA-rationalize the dataset is not unique in general, as illustrated in Example 3 in the Appendix. Taking this non-uniqueness into account, we first define the core concepts in making robust inference of the agent's behavior.

DEFINITION 3. Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Then we say that;

- *it is robust to infer that an alternative x is preferred to y* if for all preference and consideration set mapping $(\succ, \Gamma)$ that LA-rationalize the dataset $(c, \mathcal{D})$, we have $x \succ y$.
- *it is robust to infer that an alternative x attracts attention in consideration set $B \in \Omega$* if $x \in \Gamma(B)$ for all $(\succ, \Gamma)$ that LA-rationalize the dataset.
- *it is robust to infer that an alternative x* **does not** *attract attention in consideration set $B \in \Omega$* if $x \notin \Gamma(B)$ for all $(\succ, \Gamma)$ that LA-rationalize the dataset.

The definition above says that we can robustly say that an alternative is preferred to another when it will be preferred under any selection of preference that LA-rationalizes the dataset. We can see in Example 3 that the dataset may be LA-rationalized in a way that x is preferred to y or vice versa.

The issue of robust inference is interesting from the viewpoint of marketing. Suppose that the is an item that is unpopular to an agent. The results in this section may allow us to clarify whether the item is unpopular because it is less preferred or it does not attract attention. If this is possible, the seller of the item can construct meaningful marketing strategies to sell the unpopular item.

In the remainder of this section, we consider the problem of inference of preference, attention, and inattention, that is robust to the non-uniqueness of the preference and consideration set mapping that LA-rationalize the dataset. The results presented in this section are similar to those presented in Masatlioglu et al. (2012), but weaker in general. This is due to the weaker observational assumption of the model in this paper.

4.1 Robust inference of preference

Here we consider robust inference of preference.

The results here are less sharp than the result given in Masatlioglu et al. (2012). This happens because we cast a weaker observational assumption. That is, we assume that choices are observed on constraint sets in $\mathcal{D}$, which is a subset of Ω, while Masatlioglu et al. (2012) assume that choices are observed on Ω.

Proposition 1. *Let the observed choice function c be LA-rationalizable. Then;*

(a) *it is robust to infer that an alternative x is preferred to y only if $x\tilde{P}_T y$.*

(b) *it is robust to infer that an alternative x is preferred to y if $xP_T y$.*

(Proof) We first prove (a) by showing the contrapositive. Assume that $x\tilde{P}_T y$ does not hold. Then there exists an extension of $\tilde{P}$ with $y \succ x$. Then following the procedure in the proof of Theorem 1, we obtain preference and consideration mapping $(\succ, \Gamma)$ that LA-rationalizes the dataset. Therefore it is not robust to infer that x is preferred to y.

Next we prove (b). Since the dataset is LA-rationalizable, let $(\succ, \Gamma)$ LA-rationalize $(c, \mathcal{D})$. We first show that $x, y \in X$ that obey the restriction (2) implies $x \succ y$. By the restriction (2), we have $x = c(S) \neq c(S\backslash\{y\})$ for some $S, S\backslash\{y\} \in \mathcal{D}$. By the attention filter property, we have $y \in \Gamma(S)$, and thus $x \succ y$ follows.

Since P_T is the transitive closure of P, it suffices to show that xPy implies $x \succ y$. Note that in the paragraph above, we have already shown that the revealed preference relation P defined in the first round of Algorithm 1 has the desired property. Now we assume by way of induction that xPy implies $x \succ y$ for P defined after $(n-1)$ rounds of the algorithm, and consider the n^{th} round.

Assume there exist $S, T \in \mathcal{D}$ such that;

1. $x = c(S) \neq c(T)$,
2. $c(S), c(T) \in S \cap T$,
3. $\{a \in S\backslash T : c(S)P_T a\} \cup \{a \in T\backslash S : c(T)P_T a\} = \emptyset$,
4. $\{a \in T\backslash S : \text{not } aP_T c(T)\} = \emptyset$, and
5. $\{a \in S\backslash T \mid \text{not } aP_T c(S)\} = \{y\}$.

Then Algorithm 1 adds the pair (x, y) to the revealed preference relation P.

Since $(\succ, \Gamma)$ LA-rationalizes the dataset, it follows from 1 and 2 above that one of the following must hold;

$$\begin{cases} \text{there exists } a \in S \backslash T \text{ such that } c(S) \succ a, \\ \text{there exists } b \in T \backslash S \text{ such that } c(T) \succ b. \end{cases}$$

Note that by statements 3,4, and 5, we must have $x \succ y$. Since the pair of preference and consideration set mapping $(\succ, \Gamma)$ that LA-rationalize the dataset was arbitrarily chosen, we conclude that it is robust to infer that x is preferred to y if xP_Ty. QED

While our result on the robust inference of preference is weaker than that in Masatlioglu et al. (2012), when we have $\mathcal{D} = \Omega$, the results will coincide. That is because if we have $\mathcal{D} = \Omega$, the binary relations P^1 and P^2 will be empty, and we will not even have to go through Algorithm 1; i.e. $\tilde{P} = P$ will hold. Hence it is robust to infer that an alternative x is preferred to y if and only if xP_Ty, which is identical to the result in Masatlioglu et al. (2012).

4.2 Robust inference of attention and inattention

Here we consider robust inference of attention and inattention. We first consider the inference of in attention. That is, we consider when we can surely say that the agent does not pay attention to an alternative, when facing some constraint set. Again, the result given here is weaker than the result in Masatlioglu et al. (2012), due to the weaker observational assumption.

PROPOSITION 2. *Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Then;*

(a) *it is robust to infer that an alternative x does not attract attention in constraint set $S \in \mathcal{D}$ only if $x\tilde{P}_Tc(S)$.*

(b) *it is robust to infer that an alternative x does not attract attention in constraint set $S \in \mathcal{D}$ if $xP_Tc(S)$.*

(PROOF) We first prove (a) by showing the contrapositive. Assume that we do not have $x\tilde{P}_Tc(S)$. The LA-rationalizability of the dataset implies that $\tilde{P}$ is acyclic. Then, since $x\tilde{P}_Tc(S)$ does not hold, there exists a strict ordering extension of $\tilde{P}$, with $c(S) \succ x$. Then, define Γ such that

$$\Gamma(T) = \{c(T)\} \cup \{y \in T : c(T) \succ y\},$$

for $T \in \mathcal{D}$. Following the proof of sufficiency of Theorem 1, the pair of preference and consideration set mapping $(\succ, \Gamma)$ LA-rationalizes the dataset, but we have $x \in \Gamma(S)$. Hence, it is not robust to infer that x does not attract attention in S.

Next we prove (b). Let $(\succ, \Gamma)$ LA-rationalize the dataset. Then by $xP_T c(S)$, we have $x \succ c(S)$. Then since $(\succ, \Gamma)$ LA-rationalizes the dataset, it follows that we cannot have $x \in \Gamma(S)$. Thus we conclude that it is robust to infer that x does not attract attention in S if $xP_T c(S)$. QED

As in the case of robust inference of preference, our result coincides that of Masatlioglu et al. (2012) when we have $\mathcal{D} = \Omega$.

Now we consider robust inference of attention. That is, we consider when we can surely say that the agent pays attention to an alternative facing some constraint set. As was the case for the previous results, we have a waker result than Masatlioglu et al. (2012). However, unlike the previous results, in the case of robust inference of attention, we have not yet succeeded in obtaining a result where the necessity and sufficiency coincide when we have $\mathcal{D} = \Omega$.

Proposition 3. *Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Then;*

(a) *it is robust to infer that an alternative x attracts attention in constraint set $S \in \mathcal{D}$ only if $c(S)\tilde{P}_T x$.*

(b) *it is robust to infer that an alternative x attracts attention in constraint set $S \in \mathcal{D}$ if there exists $T \in \mathcal{D}$ (possibly equal to S) such that;*

- **(i)** $c(T) \neq c(T \backslash \{x\})$,
- **(ii)** *$yP_T c(S)$ for all $y \in S \backslash T$, and*
  $zP_T c(T)$ for all $z \in T \backslash S$.

(Proof) First we prove (a) by showing the contrapositive. Assume that we do not have $c(S)\tilde{P}_T x$. The LA-rationalizability of the dataset implies that $\tilde{P}$ is acyclic. Then, since $c(S)\tilde{P}_T x$ does not hold, there exists a strict ordering extension of $\tilde{P}$, which we denote by $\succ$, with $x \succ c(S)$. Then, define Γ such that

$$\Gamma(T) = \{c(T)\} \cup \{y \in T : c(T) \succ y\},$$

for $T \in \mathcal{D}$. Following the proof of sufficiency of Theorem 1, the pair of preference and consideration set mapping $(\succ, \Gamma)$ LA-rationalizes the dataset, but we have $x \notin \Gamma(S)$. Hence, it is not robust to infer that x attracts attention in S.

Now we prove (b). Assume that there exists $T \in \mathcal{D}$ with (i) and (ii), and consider any $(\succ, \Gamma)$ that LA-rationalizes the dataset. If $T = S$, then (ii) is irrelevant, and by the attention filter property, statement (i) implies that $x \in \Gamma(S)$. Since the preference and considerations set mapping $(\succ, \Gamma)$ was selected arbitrarily, we conclude that it is robust to infer that the alternative x attracts attention in constraint set S.

Next consider the case where $S \neq T$. By the attention filter property, it follows from (i) that $x \in \Gamma(T)$. Next we focus on the statement (ii). Note that $yP_Tc(S)$ for all $y \in S \backslash T$ implies that $y \notin \Gamma(S)$ for all $y \in S \backslash T$. Then, by the attention filter property, we have $\Gamma(S) = \Gamma(S \cap T)$. Following an analogous logic, we have $\Gamma(T) = \Gamma(S \cap T)$. Hence we have $x \in \Gamma(T) = \Gamma(S)$, and we conclude that it is robust to infer that x attracts attention in S. QED

5 Out of sample prediction

Here we discuss how we can make out-of-sample predictions. That is, we consider the question: provided that the dataset is LA-rationalizable, what can we say about the agent's decisions that would be made on constraint sets that are not yet observed? Since in the model of Masatlioglu et al. (2014) it was assumed that the researcher observed the the choices of the agent made on all possible constraint sets, this issue was irrelevant. In our model, we observe the agent's choices only on a subset of all possible constraint sets. Therefore, studying what we can say about the agent's decision-making on out-of-sample constraint sets is a natural question once we know that an agent's observed choices are consistent with the model of decision-making under limited attention.

In particular, we provide conditions on when we can surely predict that an alternative attracts attention, does not attract attention, will be chosen, and will not be chosen. This issue is interesting from the viewpoint of marketing. For example, if a shop is considering a new lineup of products and she can predict that some products may not attract attention from some customer, then the shop can promote those products so that the customer takes them into consideration. If the shop can surely predict that some alternative will never be chosen, the shop can just exclude that item from the lineup.

5.1 Robust prediction of attention and inattention

We first consider prediction of attention and inattention. In particular, given a dataset $(c, \mathcal{D})$ that is LA-rationalizable, we focus on a constraint set S on which we have not observed the agent's choice, i.e. $S \notin \mathcal{D}$, and some alternative $x \in S$. We give sufficient conditions to predict that x surely attracts attention in S, and conditions to predict that x surely does not attract attention in S. We begin by introducing a precise definition of prediction of attention / inattention.

Definition 4. Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Then we say that;

- *it is robust to predict that an alternative x attracts attention in constraint set $S \notin \mathcal{D}$* if $x \in \Gamma(S)$ for all preference and consideration set mapping $(\succ, \Gamma)$ that LA-rationalize the dataset.
- *it is robust to predict that an alternative x* **does not** *attract attention in constraint set $S \notin \mathcal{D}$* if $x \notin \Gamma(S)$ for all preference and consideration set mapping $(\succ, \Gamma)$ that LA-rationalize the dataset.

Proposition 4. *Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Fix any constraint set $S \notin \mathcal{D}$, and an alternative $x \in S$. Then it is robust to predict that x* **does not** *attract attention in S if there exists a constraint set $T \in \mathcal{D}$ such that;*

1. *$S \subseteq T$,*
2. *$yP_Tc(T)$ for all $y \in T \backslash S$, and*
3. *$xP_Tc(T)$.*

(Proof) Suppose that there exists a constraint set $T \in \mathcal{D}$ such that 1-3 hold, and let $(\succ, \Gamma)$ be a pair of preference and consideration set mapping that

LA-rationalizes the dataset.

It follows from the attention filter property that $\Gamma(T) = \Gamma(S)$. Furthermore, by Proposition 2 and the restriction 3 above, we have $x \notin \Gamma(T)$. Therefore, $x \notin \Gamma(S)$, and we conclude that it is robust to predict that x does not attract attention in S.

For the completeness of the proof, we show that there may exist multiple constraint sets that satisfy the three restrictions listed above. Suppose that there exist $T^1, \ldots, T^m \in \mathcal{D}$ that satisfy the restrictions in the statement. Note that restriction 2 implies that $c(T^i) \in S$ for all $i = 1, \ldots, m$. Furthermore, 2 implies that $yP_T c(T^i)$ for all $y \in T^i \backslash T^j$, for all i, j. Therefore, by the attention filter property, it follows that $\Gamma(T^i) = \Gamma(T^j)$ for all i, j, and we conclude that the existence of multiple constraint sets that satisfy the restrictions will not cause a problem. QED

PROPOSITION 5. *Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Fix any constraint set $S \notin \mathcal{D}$, and an alternative $x \in S$. Then it is robust to predict that x attracts attention in S if there exists a constraint set $T \in \mathcal{D}$ such that;*

1. *$S \subseteq T$,*
2. *$yP_T c(T)$ for all $y \in T \backslash S$, and*
3. *it is robust to infer that x attracts attention in T.*

(PROOF) Suppose that there exists a constraint set $T \in \mathcal{D}$ that satisfies the three restrictions above, and let $(\succ, \Gamma)$ be a pair of preference and consideration set mapping that LA-rationalizes the dataset.

Note that from restriction 3, it follows that $x \in \Gamma(T)$.[11] The restriction 2 implies that $y \notin \Gamma(T)$ for all $y \in T \backslash S$, and therefore we have $\Gamma(T) \subseteq S$. Then, by the attention filter property, we have $\Gamma(S) = \Gamma(T)$, and it follows that $x \in \Gamma(S)$.

For the completeness of the proof, we consider the case where there are multiple constraint sets $T^1, \ldots, T^m \in \mathcal{D}$ that satisfy the three restrictions above. Then by 2, it follows that $c(T^i) \in S$ for all $i = 1, \ldots, m$. Moreover, restriction 2 implies that $yP_T c(T^i)$ for all $y \in T^i \backslash T^j$ for all i, j. Therefore, we have $\Gamma(T^i) = \Gamma(T^j)$ for all i, j, and we conclude that the existence of multiple constraint sets that satisfy the restrictions will not cause a problem. QED

[11] See Proposition 3 for sufficient conditions for the robust inference of attention.

5.2 Robust prediction of choices

We consider below robust prediction of choices. That is, given a dataset $(c, \mathcal{D})$ that is LA-rationalizable, a constraint set $S \notin \mathcal{D}$, and an alternative $x \in S$, we consider when we can surely predict that x will be the chosen element (or that x will not be the chosen element) when the agent faces the constraint set S. In doing this, we consider an *extended dataset* $(\bar{c}, \bar{\mathcal{D}})$, where we have added the constraint set S to the original dataset, i.e. $\bar{\mathcal{D}} = \mathcal{D} \cup \{S\}$. The extended choice function $\bar{c}$ is such that $\bar{c}(T) = c(T)$ for all constraint sets $T \in \mathcal{D}$. Now, consider any preference and consideration set mapping $(\succ, \Gamma)$ that LA-rationalize the dataset $(c, \mathcal{D})$. If it must hold that $\bar{c}(S) = x$ in order for $(\succ, \Gamma)$ to LA-rationalize $(\bar{c}, \bar{\mathcal{D}})$, and this holds for all $(\succ, \Gamma)$ that LA-rationalize the observed dataset, we can robustly say that x must be chosen in S. On the other hand, if it must hold that $\bar{c}(S) \neq x$, we can robustly say that x must *not* be chosen in S. A formal definition of robust prediction of choices is given below.

Definition 5. Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Given the extended dataset $(\bar{c}, \bar{\mathcal{D}})$ with $\bar{\mathcal{D}} = \mathcal{D} \cup \{S\}$ and $\bar{c}(T) = c(T)$ for all $T \in \mathcal{D}$, we say that;

- *it is robust to predict that an alternative x is chosen in constraint set* $S \notin \mathcal{D}$ if, for all $(\succ, \Gamma)$ that LA-rationalizes the dataset $(c, \mathcal{D})$, $\bar{c}(S) = x$ must hold in order for $(\succ, \Gamma)$ to LA-rationalize the extended dataset $(\bar{c}, \bar{\mathcal{D}})$.
- *it is robust to predict that an alternative x is* **not** *chosen in constraint set* $S \notin \mathcal{D}$ if, for all $(\succ, \Gamma)$ that LA-rationalizes the dataset $(c, \mathcal{D})$, $\bar{c}(S) \neq x$ must hold in order for $(\succ, \Gamma)$ to LA-rationalize the extended dataset $(\bar{c}, \bar{\mathcal{D}})$.

Firstly we consider the robust prediction of choice. The proposition below gives a sufficient condition to say that an alternative is robustly predicted to be chosen.

Proposition 6. *Let the dataset $(c, \mathcal{D})$ be LA-rationalizable. Fix any constraint set $S \notin \mathcal{D}$ and an alternative $x \in S$, and consider the extended dataset $(\bar{c}, \bar{\mathcal{D}})$. Then it is robust to predict that x is chosen in S if there exists an observed constraint set $T \in \mathcal{D}$ such that;*

1. $S \subseteq T$,
2. $yP_Tc(T)$ *for all* $y \in T \backslash S$, *and*
3. $x = c(T)$.

(PROOF) Suppose there exists a constraint set $T \in \mathcal{D}$ that satisfies the three restrictions, and let $(\succ, \Gamma)$ be a preference and consideration set mapping that LA-rationalizes the dataset.

Note that restriction 2 implies that $y \notin \Gamma(T)$ for all $y \in T \backslash S$, and we have $\Gamma(T) \subseteq S$. Then by the attention filter property, it follows that $\Gamma(S) = \Gamma(T)$. Then the choice in S must be the same as in T, i.e. $\bar{c}(S) = x$. If not, the preference and consideration set mapping $(\succ, \Gamma)$ cannot LA-rationalize the extended dataset. Hence we conclude that it is robust to predict that x is chosen in S. QED

This is a strong result in that we can pin down which alternative must be chosen, provided that the agent is behaving in line with the model of decision-making under limited attention. In Example 1, we illustrate how this prediction is made.

Next we consider the robust prediction of non-choice. The proposition below gives a sufficient condition to say that an alternative is robustly predicted not to be chosen.

PROPOSITION 7. *Let the dataset* $(c, \mathcal{D})$ *be LA-rationalizable. Fix any constraint set* $S \notin \mathcal{D}$ *and an alternative* $x \in S$, *and consider the extended dataset* $(\bar{c}, \bar{\mathcal{D}})$. *It is robust to predict that* x *is* **not** *chosen in* S *if there exists an observed constraint set* $T \in \mathcal{D}$ *such that;*

1. $x \neq c(T)$,
2. $x, c(T) \in S \cap T$,
3. yP_Tx *for all* $y \in S \backslash T$, *and*
4. $zP_Tc(T)$ *for all* $z \in T \backslash S$.

(PROOF) Let $(\succ, \Gamma)$ be a preference and consideration set mapping that LA-rationalize the dataset $(c, \mathcal{D})$. The proof will be done by showing the contrapositive. Assume that $(\succ, \Gamma)$ LA-rationalizes the extended dataset $(\bar{c}, \bar{\mathcal{D}})$

with $\bar{c}(S) = x$. Assume further that there exists a constraint set $T \in \mathcal{D}$ that satisfies 2 - 4.

Then, 3 and 4 imply "$y \notin \Gamma(S)$ for all $y \in S \backslash T$" and "$z \notin \Gamma(T)$ for all $z \in T \backslash S$" respectively, which in turn implies that $\Gamma(S) = \Gamma(T)$. Since we have assumed that $(\succ, \Gamma)$ LA-rationalizes the extended dataset $(\bar{c}, \bar{\mathcal{D}})$, we must have $c(T) = x$. This is a violation of 1. Therefore, we conclude that there does not exist a constraint set $T \in \mathcal{D}$ that satisfies all 1 - 4. QED

In making out-of-sample predictions of choices, it is ideal if we can robustly predict that some element will be chosen in some constraint set. However, this may not always be possible. Therefore, here we broaden the issue, and consider which alternative in a constraint set may be a candidate element to be chosen. In making this prediction, the result in Proposition 7 is very useful.

Assume that the observed dataset $(c, \mathcal{D})$ is LA-rationalizable, and fix any constraint set $S \notin \mathcal{D}$. We consider which alternative $x \in S$ can possibly be chosen in the extended dataset $(\bar{c}, \bar{\mathcal{D}})$. In particular, we consider a sufficient condition to say that there exists a preference and consideration set mapping $(\succ, \Gamma)$ such that the extended dataset $(\bar{c}, \bar{\mathcal{D}})$ with $\bar{\mathcal{D}} = \mathcal{D} \cup \{S\}$ and $\bar{c}(S) = x$, is LA-rationalized by $(\succ, \Gamma)$.

In doing this, we consider the contrapositive. Fix any alternative $x \in S$, and assume that x is not a candidate element to be chosen in S. This means that for any preference and consideration set mapping $(\succ, \Gamma)$ that LA-rationalize the observed dataset $(c, \mathcal{D})$, we cannot have $\bar{c}(S) = x$ in order for $(\succ, \mathcal{D})$ to LA-rationalize the extended dataset $(\bar{c}, \bar{\mathcal{D}})$ with $\bar{\mathcal{D}} = \mathcal{D} \cup \{S\}$. This is identical to saying that it is robust to predict that x is not chosen in S.

Therefore, we conclude that an alternative $x \in S$ is a candidate alternative to be chosen if it is **not** robust to predict that x is not chosen in S. An illustration of how this inference is made is given in Example 2.

Below we give two examples of how out-of-sample prediction can be made. The first one will illustrate robust prediction of choice, attention, and inattention, while the second one will illustrate robust prediction of non-choice and prediction of candidate alternatives.

Example 1 (Robust prediction of choice, attention, and inattention)**.** Consider the dataset $(c, \mathcal{D})$ with $T \in \mathcal{D}$ and $S \notin \mathcal{D}$, such that $T = \{a, b, x, y, z\}$

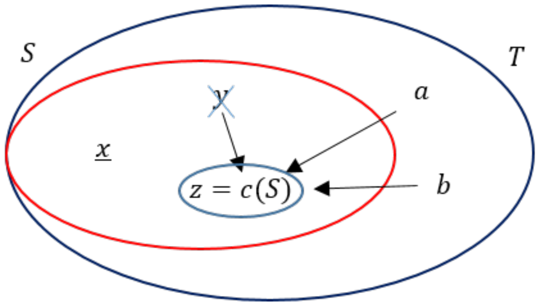

Figure 1: Robust prediction of choice, attention, and inattention

and $S = \{x, y, z\}$. Assume that the dataset is LA-rationalizable, and we know that;

1. $z = c(T)$,
2. $a, b, P_T z$,
3. it is robust to infer that x attracts attention in T, and
4. $y P_T z$.

Take any preference and consideration set mapping $(\succ, \Gamma)$ that LA-rationalize the dataset, and we consider predicting the choice, attention, and inattention on the constraint set S. By 1 and 2 we have $a, b \notin \Gamma(T)$, and thus $\Gamma(T) = \Gamma(S)$ follows from the attention filter property. Then by 3, we have $x \in \Gamma(T)$, and thus it is robust to predict that x attracts attention in S. It follows from 4 that $y \notin \Gamma(T)$, and thus it is robust to predict that y does not attract attention in S.

Finally, since we have $\Gamma(T) = \Gamma(S)$, in order for $(\succ, \Gamma)$ to LA-rationalize the extended dataset $(\bar{c}, \bar{\mathcal{D}})$ with $\bar{\mathcal{D}} = \mathcal{D} \cup \{S\}$, it must follow that $\bar{c}(S) = z$. □

In Figure 1, the alternatives with a $\times$ mark are those that are robustly predicted not to attract attention in S, the alternatives with an underline are the ones that are robustly predicted to attract attention in S, and the one that is circled is the one that is robustly predicted to be chosen. The arrows indicate that the alternative at the tail of the arrow is revealed preferred to the alternative at the head of the arrow.

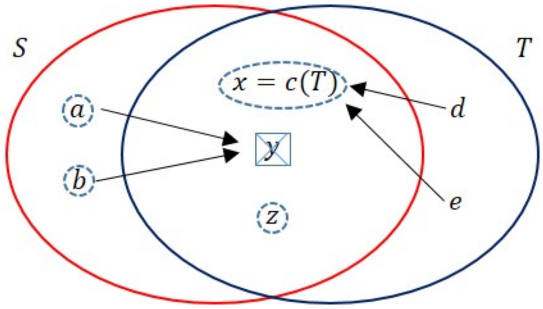

Figure 2: Robust prediction of non-choice and candidates

Example 2 (Robust prediction of non-choice and prediction of candidate alternatives)**.** Consider the dataset $(c, \mathcal{D})$ with $T \in \mathcal{D}$ and $S \notin \mathcal{D}$, such that $T = \{d, e, x, y, z\}$ and $S = \{a, b, x, y, z\}$. Assume that the dataset is LA-rationalizable, and we know that;

1. $x = c(T)$,
2. $a, b, P_T y$, and
3. $d, e, P_T x$.

Take any preference and consideration set mapping $(\succ, \Gamma)$ that LA-rationalize the dataset, and we consider the extended dataset $(\bar{c}, \bar{\mathcal{D}})$ with $\bar{\mathcal{D}} = \mathcal{D} \cup \{S\}$.

Note that 3 implies that $d, e \notin \Gamma(T)$. Now assume that we have $\bar{c}(S) = y$. Then, in order for $(\succ, \Gamma)$ to LA-rationalize the extended dataset $(\bar{c}, \bar{\mathcal{D}})$, we must have $a, b \notin \Gamma(S)$. Combining this with $d, e \notin \Gamma(T)$ and 1, it must follow that $x \succ y$ and $y \succ x$. This contradicts that $(\succ, \Gamma)$ LA-rationalizes the dataset. Hence we conclude that y cannot be chosen in S.

Since y is the only alternative in S that is robustly predicted not to be chosen, we predict that the alternatives a, b, x, z are candidate alternatives to be chosen in S. □

In Figure 2, the alternative in the box with the $\times$ mark is the one that is robustly predicted not to be chosen. The alternatives in the dashed circles are the candidate alternatives to be chosen.

Appendix

Proof of Fact 1

(Proof) Let preference and consideration set mapping $(\succ, \Gamma)$ LA-rationalize the dataset $(c, \mathcal{D})$. Assume by way of contradiction that the dataset violates condition (LA). This means that there exist constraint sets $S, T \in \mathcal{D}$ such that (i): $c(S) \neq c(T)$, (ii): $c(S), c(T) \in S \cap T$, (iii): $E(S,T) \cup E(T,S) = \emptyset$, and (iv): $\bar{E}(S,T) \cup \bar{E}(T,S) = \emptyset$. This implies that $aP_Tc(S)$ for all $a \in S \backslash T$, which in turn implies that $a \succ c(S)$ and $a \notin \Gamma(S)$. Then by the attention filter property, we have $\Gamma(S) = \Gamma(S \cap T)$. Analogously, we have $\Gamma(T) = \Gamma(S \cap T)$, and thus $\Gamma(S) = \Gamma(T)$ holds. However, this contradicts the assumption that $(\succ, \Gamma)$ LA-rationalizes the dataset, since $\Gamma(S) = \Gamma(T)$ and $c(S) \neq c(T)$ imply that $c(S) \succ c(T)$ and $c(T) \succ c(S)$. QED

Proof of Fact 2

(Proof) Assume by way of contradiction that the binary relation has a cycle as follows:

$$x^1(P \cup P^1)x^2(P \cup P^1)x^3 \cdots x^m(P \cup P^1)x^1.$$

Since the revealed preference P is acyclic under LA-rationalizability of $(c, \mathcal{D})$, there is at least one pair of elements in the cycle x^i, x^{i+1} such that $x^iP^1x^{i+1}$.

Note that the binary relation P^1 is defined only on F. Thus take a subset $\{y^1, \ldots, y^n\}$ of the elements in the cycle as follows; $\{y^1, \ldots, y^n\} \subset \{x^1, \ldots, x^m\}$, $\{y^1, \ldots, y^n\} \subseteq F$, and the elements in $\{y^1, \ldots, y^n\}$ are ordered in the same relative order as in $\{x^1, \ldots, x^m\}$. Then $\{y^1, \ldots, y^n\}$ is a cycle with elements of F such that

$$y^1(P_T \cup P^1)y^2(P_T \cup P^1) \cdots (P_T \cup P^1)y^n(P_T \cup P^1)y^1, \tag{10}$$

with at least one P^1 in the cycle above.

First assume that there is only one P^1 in the cycle, and let $y^1P^1y^2P_Ty^1$ without loss of generality. By definition of the binary relation P^1, $y^1P^1y^2$ implies $y^1 \in F^1$ and $y^2 \in F^1 \cup F^2$. Note that the cycle implies $RP(y^1) \cap (F \backslash F^0) \neq \emptyset$ and $RW(y^2) \cap F^1 \neq \emptyset$, and thus neither (α) nor (β) can be satisfied between the elements y^1 and y^2, and hence we must have two or more P^1s in the cycle (10).

Now assume that there are two pairs of alternatives that are connected via P^1 in cycle (10), and let $y^1P^1y^2$ and $y^iP^1y^{i+1}$, for some $i \in \{2, n-1\}$, without loss of generality.[12]

If $y^1P^1y^2$ is defined via (α), $y^iP^1y^{i+1}$ must be defined via (β), since we have $y^2 \in RP(y^i) \cap (F \backslash F^0)$. However, then we have $y^1 \in RW(y^{i+1}) \cap F^1$, so this case is impossible.

If $y^1P^1y^2$ is defined via (β), $y^i \in RW(y^2) \cap F^1$, so this case cannot hold as well.

Thus we conclude that we cannot have any P^1s in the cycle (10), which contradicts the LA-rationalizability of the dataset. QED

Proof of Fact 3

(Proof) Assume that the dataset satisfies condition (LA). Note from the definition of the binary relation P^1, for the constraint sets relevant to P^1, condition (1) will be satisfied.

Now consider the binary relation P^2. Since condition (LA) is satisfied, the set K^i will be nonempty for all pairs of constraint sets in $\mathcal{J}$. Then, by Algorithm 2, we see that the output, namely the binary relation P^2, will always satisfy (1) over the constraint sets relevant in $\mathcal{J}$. Hence we conclude that the binary relation $\tilde{P} = (P_T \cup P^1 \cup P^2)$ obeys condition (1). QED

Proof of Fact 4

(Proof) Note that Fact 2 tells us that the binary relation $(P \cup P^1)$ is acyclic under LA-rationalizability of the dataset $(c, \mathcal{D})$. Now assume by way of contradiction that there is a cycle with respect to the binary relation $(P \cup P^1 \cup P^2)$. That is, there exist elements $\{x^1, \dots, x^m\} \subseteq X$ such that

$$x^1(P \cup P^1 \cup P^2)x^2(P \cup P^1 \cup P^2) \cdots (P \cup P^1 \cup P^2)x^m(P \cup P^1 \cup P^2)x^1.$$

Since the binary relations $P, (P \cup P^1)$, and $(P \cup P^2)$ are acyclic by assumption, the cycle above must contain both P^1 and P^2. Assume without loss of

[12] It is impossible to have $y^1P^2y^2$ and $y^2P^2y^1$ by definition of the binary relation P^1.

generality that $x^1 P^1 x^2$.

Consider $x^i (P^1 \cup P^2) x^{i+1}$, where i is the smallest index greater than 1 with the pair (x^i, x^{i+1}) connected via the binary relation P^1 or P^2. Such a pair (x^i, x^{i+1}) exists because we must have both P^1 and P^2 in the cycle. Note that we cannot have $x^1 P^1 x^2$ be defined via the condition (β). This is because $x^i (P^1 \cup P^2) x^{i+1}$ implies that $x^i \in F^1$ and $x^2 P_T x^i$, which violates the condition required in (β). Thus we must have $x^1 P^1 x^2$ be defined via condition (α), which implies that $x^1 \in F^0$ and $RP(x^1) \cap (F \backslash F^0) = \emptyset$.

Next consider $x^j (P^1 \cup P^2) x^{j+1}$, where j is the greatest index with the pair (x^j, x^{j+1}) connected via the binary relation P^1 or P^2. That is, j is the index where we have $x^{j+1} P_T x^1$. Note that we cannot have $x^{j+1} = x^1$, since $x^j (P^1 \cup P^2) x^{j+1}$ implies $x^{j+1} \notin F^0$, which contradicts the assumption that $x^1 P^1 x^2$ is defined via condition (α). By definition of binary relations P^1 and P^2, we have $x^{j+1} \notin F^0$. This contradicts that $RP(x^1) \cap (F \backslash F^0) = \emptyset$. Hence we conclude that we cannot have both P^1 and P^2 in the cycle, which in turn implies that the binary relation $(P \cup P^1 \cup P^2)$ is acyclic. QED

Example 3 (Example of a dataset that is LA-rationalizable.)**.** Here we give an example of an dataset that is LA-rationalizable, while it may not be rationalizable in the classical sense. We also show that there may be multiple pairs of preference and consideration set mapping that LA-rationalize the dataset.

Let $X = \{x, y, z\}$, and consider the dataset $(c, \mathcal{D})$ as below;

	S^1	S^2	S^3
$\mathcal{D}$	$\{x, y, z\}$	$\{x, y\}$	$\{x, z\}$
c	x	y	x

Note that the observed choices in constraint sets S^1 and S^2 reveals that the dataset is not rationalizable in the classical sense, since $c(S^1) = x$ reveals that x is preferred to y, and $c(S^2) = y$ reveals that y is preferred to x, which is a contradiction.

Now we show that the dataset is LA-rationalizable. First of all, since we have $x = c(S^1) \neq c(S^2) = y$, we it is revealed that xPz, and that it is robust to infer that z attracts attention in S^1. Below we give two examples of pairs of preference and consideration set mapping that LA-rationalize the dataset.

(I) $(\succ_1, \Gamma_1)$ such that

$x \succ_1 y \succ_1 z$,

$\Gamma_1(S^1) = \{x, y, z\}, \Gamma_1(S^2) = \{y\}, \Gamma_1(S^3) = \{x, z\}, \Gamma_1(yz) = \{y, z\}$.

(II) $(\succ_2, \Gamma_2)$ such that

$y \succ_2 x \succ_2 z$,

$\Gamma_2(S^1) = \{x, z\}, \Gamma_2(S^2) = \{x, y\}, \Gamma_2(S^3) = \{x, z\}, \Gamma_2(yz) = \{z\}$.

□

Example 4 (The revealed preference P is acyclic, but the dataset is not LA-rationalizable.)**.** Let $X = \{v, w, x, y, z\}$, and consider the dataset $(c, \mathcal{D})$ as below;

	S^1	S^2	S^3	S^4	S^5	S^6	S^7	S^8
$\mathcal{D}$	$\{v, w, x, y\}$	$\{v, w\}$	$\{v, w, z\}$	$\{w, z\}$	$\{x, y, z\}$	$\{y, z\}$	$\{w, x, y, z\}$	$\{w, x\}$
c	v	w	z	w	y	z	x	w

Considering the first round of the algorithm, it follows from $c(S^3)$ and $c(S^4)$ that zPv, from $c(S^5)$ and $c(S^6)$ that yPx, and from $c(S^7)$ and $c(S^5)$ that xPw. Now we go through the algorithm again and add additional pairs of elements to the revealed preference P. Focus on constraint sets S^7 and S^8. We have (i) $x = c(S^7) \neq c(S^8) = w$, (ii) $c(S^7), c(S^8) \in S^7 \cap S^8$, (iii) $S^8 \backslash S^7 = \emptyset$ and $\{a \in S^7 \backslash S^8 \ : \ c(S^7) P_T a\} = \emptyset$, (iv) $\{a \in S^7 \backslash S^8 \ : \ a P_T c(S^7) \text{ does not hold}\} = \{z\}$. Then, by equation (5), we have xPz.

The full profile of the revealed preference P is as follows;

$$xPw, \ xPz, \ yPx, \ zPv,$$

which is clearly acyclic. Note that for any preference $\succ$ that LA-rationalizes $(c, \mathcal{D})$, we must have

$$y \succ x \succ z \succ v, \text{ and} \tag{11}$$

$$y \succ x \succ w. \tag{12}$$

Now we show that any preference that obeys (11) and (12) cannot satisfy condition (1). Consider constraint sets S^1 and S^2. We have $c(S^1) \neq c(S^2)$

and $c(S^1), c(S^2) \in S^1 \cap S^2$. By (11), we have $x, y \succ v$, which means that for all $a \in S^1 \backslash S^2$, we have $a \succ c(S^1)$, and (1) is violated. Thus we conclude that the dataset is not LA-rationalizable. □

Example 5 (Example where P is acyclic and obeys condition (LA), but not LA-rationalizable.)**.** In this example, we illustrate a case where the revealed preference relation P is acyclic and the dataset $(c, \mathcal{D})$ satisfies condition (LA), but the dataset is not LA-rationalizable. Let $X = \{a, b, d, e, f, g, h, i\}$, and consider the dataset $(c, \mathcal{D})$ as below;

	S^1	S^2	S^3	S^4	S^5	S^6
$\mathcal{D}$	$\{a, b, d, e\}$	$\{a, b\}$	$\{f, g, h, i\}$	$\{f, g\}$	$\{d, f, g\}$	$\{d, g\}$
c	a	b	f	g	d	g

	S^7	S^8	S^9	S^{10}	S^{11}	S^{12}
$\mathcal{D}$	$\{e, f, g\}$	$\{e, g\}$	$\{a, b, h\}$	$\{b, h\}$	$\{a, b, i\}$	$\{b, i\}$
c	e	g	h	b	i	b

The pairs of constraint sets $(S^5, S^6), (S^7, S^8), (S^9, S^{10})$, and (S^{11}, S^{12}) give dPf, ePf, hPa, and iPa respectively. These are the only revealed preference relations. It is cleat that the dataset satisfies condition (LA). It can also be seen that the binary relation P^1 is empty.

Now we consider the binary relation P^2. The set $\mathcal{J} = \{\{S^1, S^2\}, \{S^3, S^4\}\}$. Denote by $K^{12} = K(S^1, S^2) \cup K(S^2, S^1)$ and by $K^{34} = K(S^3, S^4) \cup K(S^4, S^3)$. Then $K^{12} = \{(a, d), (a, e)\}$ and $K^{34} = \{(f, h), (f, i)\}$. Then in each round of Algorithm 2, we will set one of the following;

$$aP^2d \text{ and } fP^2h,$$
$$aP^2d \text{ and } fP^2i,$$
$$aP^2e \text{ and } fP^2h,$$
$$aP^2e \text{ and } fP^2i.$$

In either case, the binary relation $(P \cup P^2)$ has a cycle. Therefore, this dataset is not LA-rationalizable. □

References

[1] Andrikopoulos, A. (2009): "Szpilrajn-type Theorems in Economics," MPRA paper, No. 14345.

[2] de Clippel, G., and K. Rozen (2014): "Bounded Rationality and Limited Datasets," Cowles Foundation Discussion Paper No. 1853.

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

[4] Loomes, G., C. Starmer, R. Sugden (1991): "Observing Violations of Transitivity by Experimental Methods," *Econometrica*, 59(2), 425-439.

[5] Manzini, P., and M. Mariotti (2007): "Sequentially Rationalizable Choice," *American Economic Review*, 97(5), 1824-1839.

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

[7] Ok, E. A., P. Ortoleva, and G. Riella (2015): "Revealed (P)Reference Theory," *American Economic Review*, 105(1), 299-321.

[8] Roelofsma, P. H. M. P, and D. Read (2000): "Intransitive Intertemporal Choice," *Journal of Behavioral Decision Making*, 13, 161-177.

[9] Salant, Y., and A. Rubinstein (2008): "(A, f): Choice with Frames," *Review of Economic Studies*, 75, 1287-1296.