Coarse Revealed Preference*

Preliminary and Incomplete.

Please do not circulate.

Gaoji Hu[†] Jiangtao Li[‡] John Quah[§] Rui Tang[¶]

April 24, 2018

Abstract

We identify necessary and sufficient conditions under which a coarse data set can be rationalized by a linear order (or weak order). The conditions are easy to check and algorithms are provided. We apply our theory to investigate the observable restrictions of economic models including (1) rational choice with imperfect observation; (2) multiple preferences; and (3) minimax regret.

*We thank David Dillenberger, Faruk Gul, Pietro Ortoleva, Wolfgang Pesendorfer, Satoru Takahashi, Can Urgun, and seminar participants at NUS, Singapore Management University, and Princeton Microeconomic Theory Student Lunch Seminar for helpful comments. All remaining errors are our own.

[†]Department of Economics, National University of Singapore, hu.gaoji@gmail.com

[‡]School of Economics, University of New South Wales, jiangtao.li@unsw.edu.au

[§]School of Economics, Johns Hopkins University, kquah1@jhu.edu

[¶]Department of Economics, Princeton University, ruit@princeton.edu

Contents

1 Introduction

Pioneered by Samuelson (1938), revealed preference is one of the most influential ideas in economics and has been applied to a number of areas of economics, including consumer theory (Afriat (1967)), general equilibrium theory (Brown and Matzkin (1996)), and industrial organization (Carvajal et al. (2013)), among many others.

In a typical revealed preference exercise, it is assumed that there is an observer who records the choice behavior of the decision maker (DM). Alternative x is revealed to be preferred to alternative y if and only if x is chosen when y is also available. Thus, the standard revealed preference argument hinges on an implicit assumption that the observer could perfectly observe the DM's choice in different alternative sets. What if the observer does not observe the DM's exact choice?

Consider the rational choice model with imperfect observation. We would like to test whether the DM has a strict preference $\succ$ over alternatives, and chooses the $\succ$-maximal alternative for all feasible sets from which the DM needs to make a choice. We have in mind an observer who might not have perfect observation of the DM's choice. Rather, he has a coarse observation (A, B), where A is a feasible set from which the DM needs to make a choice, and B is a nonempty subset of A. The interpretation of (A, B) is that the observer knows that, when facing alternative set A, the DM chooses some alternative in B, but she does not know the DM's exact choice. Suppose that the DM has a strict preference over alternatives, then the only inference that the observer can draw is that B contains an alternative that is maximal in A for the DM. If the observer has multiple coarse observations $\mathcal{O} = \{(A_i, B_i)\}_{i=1}^n$, termed a coarse data set, then for all A_i, the maximal alternative for the DM lies in B_i. A priori, it is not clear how to test whether the coarse data set is rationalizable by a strict preference. Consider the following example.

Example 1 (A coarse data set that is not rationalizable by a strict preference)**.** *Consider a coarse data set $\mathcal{O}$ consisting of the following four observations:*

1. $A_1 = \{x_1, x_2, x_3, x_4\}, B_1 = \{x_1, x_2\}$;
2. $A_2 = \{x_2, x_3, x_4, x_5\}, B_2 = \{x_2, x_3\}$;
3. $A_3 = \{x_3, x_4, x_5, x_1\}, B_3 = \{x_3, x_4\}$;
4. $A_4 = \{x_4, x_5, x_1, x_2\}, B_4 = \{x_4, x_5\}$.

We claim that $\mathcal{O}$ is not rationalizable by a strict preference. Indeed, suppose that there exists a strict preference $\succ$ that rationalizes $\mathcal{O}$. Each observation i reveals that there exists some alternative x in B_i such that $x \succ^* y$ for every $y \in A_i \setminus \{x\}$. That is,

1. *(1a) $x_1 \succ^* x_2$, $x_1 \succ^* x_3$, $x_1 \succ^* x_4$, or (1b) $x_2 \succ^* x_1$, $x_2 \succ^* x_3$, $x_2 \succ^* x_4$;*
2. *(2a) $x_2 \succ^* x_3$, $x_2 \succ^* x_4$, $x_2 \succ^* x_5$, or (2b) $x_3 \succ^* x_2$, $x_3 \succ^* x_4$, $x_3 \succ^* x_5$;*
3. *(3a) $x_3 \succ^* x_1$, $x_3 \succ^* x_4$, $x_3 \succ^* x_5$, or (3b) $x_4 \succ^* x_1$, $x_4 \succ^* x_3$, $x_4 \succ^* x_5$;*
4. *(4a) $x_4 \succ^* x_1$, $x_4 \succ^* x_2$, $x_4 \succ^* x_5$, or (4b) $x_5 \succ^* x_1$, $x_5 \succ^* x_2$, $x_5 \succ^* x_4$.*

One naive way to proceed is to consider all possible combinations case by case. For example, suppose that (1a), (2a), (3a), and (4a) hold simultaneously, we can then use the classical strong axiom of revealed preference (SARP) to check for acyclicity. Since (1a) requires that $x_1 \succ^* x_3$ and (3a) requires that $x_3 \succ^* x_1$, we arrive at a contradiction. Therefore, it cannot be the case that (1a), (2a), (3a), and (4a) hold simultaneously. By considering all possible combinations case by case, one eventually can show that $\mathcal{O}$ is not rationalizable by a strict preference.[1]

As illustrated by Example 1, the *or operator* prior to checking for acyclicity renders the use of the classical SARP more difficult. Nevertheless, we shall provide an easy-to-check condition for a coarse data set to be rationalizable by a strict preference.

In this paper, we provide a systematic analysis of revealed preference under coarse information. We identify necessary and sufficient conditions under which a coarse data set can be rationalized by a linear order (or weak order). These conditions are easy to check and algorithms are provided.

To explain our analysis in the most clear way, we start by discussing the rationalizability of a coarse data set by a linear order. A coarse data set $\mathcal{O}$ is set of data points $\{(A_i, B_i)\}_{i=1}^n$, where A_i is a feasible set from which the DM needs to make a choice, and B_i is nonempty

[1] The readers might object that, for this example, there is a more efficient way of deriving contradictions. (1a) cannot be true, since (1a) is inconsistent with the third observation. Suppose that (1b) holds, (4b) must hold, since (1b) is inconsistent with (4a). But (1b) and (4b) together imply that x_5 is the maximal element for any strict preference that rationalizes the coarse data set. However, this contradicts with the second observation. Therefore, the coarse data set is not rationalizable by a strict preference. Our objective in this paper is to provide a systematic test for the rationalizability of a coarse data set, going beyond the specific arguments that were tailor-made for each coarse data set.

subset of A_i for all i. We identify a necessary and sufficient condition, termed Coarse SARP, for the existence of a linear order P such that the P-maximal alternative within each alternative set A_i lies in B_i. Coarse SARP requires that

Coarse SARP. For any $\emptyset \neq \mathcal{O}' \subset \mathcal{O}$, $\cup_{(A_i,B_i)\in\mathcal{O}'} B_i \setminus \cup_{(A_i,B_i)\in\mathcal{O}'} (A_i \setminus B_i) \neq \emptyset$.

That Coarse SARP is a necessary condition should come as no surprise. Suppose that a coarse data set is rationalizable by a linear order. For any nonempty subcollection $\emptyset \neq \mathcal{O}' \subset \mathcal{O}$ and $(A_i, B_i) \in \mathcal{O}'$, since the maximal element in A_i is contained in B_i but not in $A_i \setminus B_i$, the maximal element in $\cup_{(A_i,B_i)\in\mathcal{O}'} A_i$ is necessarily contained in $\cup_{(A_i,B_i)\in\mathcal{O}'} B_i$ but not in $\cup_{(A_i,B_i)\in\mathcal{O}'} (A_i \setminus B_i)$. We show that Coarse SARP is also sufficient.

Notably, it is easy to check whether the Coarse SARP property holds. In the standard revealed preference exercise, the usual treatment is to derive preference relations between alternatives from each observation, and then check for acyclicity. Rather than deriving preference relations between alternatives from each observation, our approach offers for a more convenient way of testing the data by working with sets of observations and sets of alternatives. We provide a simple and efficient algorithm that drastically decreases the difficulty of checking whether the a coarse data set is rationalizable by a linear order. For an arbitrary coarse data set $\mathcal{O}$ with m alternatives and n data points, the number of steps required in the algorithm is at most $\min\{m, n\}$.

Our theory can be readily applied to test the observable restrictions of the rational choice model with imperfect observation. Note that the rational choice model with imperfect observation is suitable for instances in which an observer has only partial information about the DM's choice from a feasible set. This occurs in situations where the observer sees an agent committing to choosing from some subset of a feasible set, without observing her ultimate choice. For example, consider a two-stage choice of the DM. The DM is assumed to know all restaurants/ menus that he might select. The DM chooses a meal eventually, but his initial choice is of a restaurant/ menu from which he will later chooses his meal. If the observer only has knowledge of the restaurant/ menu that the DM chooses, but not the DM's ultimate choice, this could be modeled as rational choice under imperfect observation. For another example, consider $X = X_1 \times X_2$, where X_1 is the set of alternatives available at date 1, and X_2 is the set of available alternatives at data 2. The agent chooses (x_1, x_2) from a nonempty feasible set $A \subseteq X$. If the observer only has knowledge that the choice of the

DM at date 1, but not the choice at data 2, this could be modeled as rational choice under imperfect observation. Such a framework would be suitable in a variety of contexts, such as intertemporal choice problems, where information about the choice(s) of the agent is revealed in stages.[2]

We hasten to emphasize that our framework works beyond the rational choice model with imperfect observation. Besides the empirical perspective that the observer might not have perfect observation of the DM's choice, we also motivate our theory from a theoretical perspective. We show that our theory can also be applied to investigate the observable restrictions of a number of behavior models such as the multiple preferences model and the minimax regret model. While there is no imperfect observation of the DM's choice, imperfect information arises from these behavior models themselves. To the best of our knowledge, our paper provides the first study of the revealed preference implications of these two models. Our theory facilitates the study because we can transform the problem of checking rationalizability by multiple preferences and rationalizability under the minimax regret model into a problem of checking the rationalizability of certain coarse data sets by a linear order. Below, we illustrate via examples the relationship between our theory and the observable restrictions of economic models including the multiple preferences model and the minimax regret model. The formal analysis is contained in Section 4.

Example 2 (The multiple preferences model)**.** *In the multiple preferences model (see Kalai et al. (2002) and Salant and Rubinstein (2008)), the DM has a set of strict preferences over alternatives. In each feasible set A, the DM chooses all alternatives that are maximal according to some preference in the set of strict preferences.*

Suppose that the DM chooses $\{x_1, x_2\}$ when facing $A_1 = \{x_1, x_2, x_3, x_4\}$, and $\{x_2, x_3\}$ when facing $A_2 = \{x_2, x_3, x_4, x_5\}$. Intuitively, if the DM's choices are rationalizable by multiple preferences, then there must exist a strict preference that ranks x_1 above all other alternatives in A_1. But the preference also needs to be consistent with the DM's choice from A_2. Therefore, there must exists a preference that ranks x_1 above all other alternatives in A_1, and ranks either x_2 or x_3 above all other alternatives in A_2. In other words, the following coarse data set must be rationalizable by a linear order: The DM chooses x_1 in A_1, and chooses some alternative in $\{x_2, x_3\}$ from A_2. Building on this logic, Section 4 shows that,

[2]This has been briefly discussed in Nishimura et al. (2015) to point to the flexible nature of the revealed preference framework.

testing rationalizability by multiple preferences is equivalent to testing rationalizable of a family of coarse data sets by a linear order.

Example 3 (The minimax regret model)**.** *In the minimax regret model, the DM anticipates regrets and chooses an alternative that minimizes the worst-case regret. The DM has a finite set of utility functions $\mathcal{U}$ defined on X. Let $\phi(x, y) := \max_{u \in \mathcal{U}} \{u(y) - u(x)\}$ denote the worst-case regret of x towards y.*

To see how the minimax regret model is related to our theory, suppose that we observe that the DM chooses x from $A = \{x, y, z\}$. This reveals that x generates a lower worst-case regret than y does. Therefore,

$$\max\{\phi(y, x), \phi(y, z)\} > \max\{\phi(x, y), \phi(x, z)\},$$

but we don't know whether it is $\phi(y, x)$ or $\phi(y, z)$ that is larger than $\max\{\phi(x, y), \phi(x, z)\}$. The imperfect information arises from the minimax regret model itself. Section 4 shows that, testing rationalizability under the minimax regret model is equivalent to testing the rationalizability of a corresponding coarse data set by a linear order.

We then proceed to discuss the rationalizability of a coarse data set by a weak order. Note that without additional information, any behavior of the DM is rationalizable by a weak order, since the DM could be indifferent among all alternatives. For a meaningful discussion of rationalizability by a weak order, a data point is defined to be a tuple (A, B, D) where A is a feasible set, B and D are disjoint subsets of A. The interpretation is that when the DM is faced with a choice problem A, the observer knows that the exact choice of the DM is contained in B, and she also knows the alternatives in D are not optimal for the DM in A. Our analysis here parallels the analysis for the case of linear oder. We identify a necessary and sufficient condition, termed Coarse WARP, for the existence of a weak order such that the set of maximal elements has a common element with B_i and is disjoint with D_i. A simple and efficient algorithm is provided to test the Coarse WARP condition.

Before we move on to talk about the model and the results, we wish to discuss two closely related papers. Fishburn (1976) studies when a choice function, which maps each set of alternatives in a domain of feasible sets into a non-empty subset of itself (called the choice set), can be representable by a weak (resp. linear) order. A choice function is representable by a weak (resp. linear) order if some weak (resp. linear) order on the alternatives has maximal elements within each feasible set, all of which are in the choice set of the feasible set. Despite the resemblance between our paper and Fishburn (1976), there are important differences. First, while Fishburn (1976) provides a partial congruence axiom to characterize representable choice functions, he does not provide an algorithm to test the axiom. Our analysis is from a revealed preference perspective, and an algorithm is proposed to test the conditions. Second, we provide economic interpretations of the exercise, and apply the theory to investigate the observable restrictions of many economic models. Third, our analysis is completely different in the case of weak order, due to different definitions of rationalizability by a weak order.

In a thoughtful paper, de Clippel and Rozen (2014) point out a methodological pitfall when testing choice theories with limited data in the the recent bounded rationality literature. They then present testable implications for several bounded rationality theories, and an enumeration procedure is proposed to test these implications. While the motivations differ significantly, it turns out that our algorithm can be viewed as a generalization of their enumeration procedure. While de Clippel and Rozen (2014) addresses the important question of incomplete data in the the recent bounded rationality literature, the objective of our paper is provide a systematic analysis of the rationalizability of a coarse data set by a linear order (or weak order). We show that the theory can be used to investigate the observable restrictions of many economic models. Another important difference is that, we also propose a theory for rationalizability by a weak order, which is not covered in their analysis. Our paper and de Clippel and Rozen (2014) are clearly complementary.

Section 2 presents the basics of the model. Section 3 considers an arbitrary coarse data set and identifies a necessary and sufficient condition for the data to be rationalizable by a linear order. Section 4 applies our theory investigate the observable restrictions of economic models including (1) rational choice with imperfect observation; (2) multiple preferences; and (3) minimax regret. Section 5 studies rationalizability of a coarse data set by a weak order. Section 6 concludes the paper.

2 Preliminaries

We work with any arbitrarily fixed nonempty finite set X, which can be viewed as the universal set of available alternatives. Let $\mathcal{X}$ be the collection of all nonempty subsets of X.

A coarse data set $\mathcal{O}$ is a set of data points $\{(A_i, B_i)\}_{i=1}^n$, where $A_i \in \mathcal{X}$ and $\emptyset \neq B_i \subseteq A_i$ for all i. In the general model, each (A_i, B_i) need not have intrinsic meanings. In each of the applications that we consider, additional structure is imposed, and we shall specify the interpretation of (A_i, B_i). To simplify the statements below, we write C_i rather than $A_i \setminus B_i$. We also use the following notations in this paper:

$$A(\mathcal{O}') := \cup_{(A_i,B_i)\in\mathcal{O}'} A_i,$$
$$B(\mathcal{O}') := \cup_{(A_i,B_i)\in\mathcal{O}'} B_i,$$
$$C(\mathcal{O}') := \cup_{(A_i,B_i)\in\mathcal{O}'} C_i$$

for any $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$. Throughout the rest of the paper, unless it leads to confusion, we abuse the notation by suppressing the set delimiters, e.g., writing x rather than $\{x\}$.

A linear order is a complete, transitive, and antisymmetric binary relation, denoted by P. A strict preference is the asymmetric part of a linear order, denoted by $\succ$. We use $\max(A, P)$ (resp. $\max(A, \succ)$) to represent the maximal element in A according to P (resp. $\succ$).

Our objective is to find necessary and sufficient conditions under which a coarse data set is rationalizable by a linear order. We say that a coarse data set $\mathcal{O} = \{(A_i, B_i)\}_{i=1}^n$ is rationalized by some linear order P if

$$\max(A_i,\, P) \in B_i$$

for all i. If a coarse data set is rationalized by some linear order P, we say that the data set is rationalizable by a linear order.

3 Theory

This section contains our main results. We propose a property of coarse data sets that we term Coarse SARP and show that the Coarse SARP property is both necessary and sufficient for a coarse date set to be rationalizable by a linear order. Notably, the Coarse SARP property is easy to check. A simple and efficient algorithm to check the Coarse SARP property is provided. For an arbitrary coarse data set $\mathcal{O} = \{A_i, B_i\}_{i=1}^n$, the number of steps required in the algorithm is at most $\min\{|A(\mathcal{O})|, n\}$.[3] We report some identification results towards the end of this section.

To fix ideas, let us first consider the special case that B_i is a singleton for all i. Without loss of generality, we rewrite $\mathcal{O}$ as $\{(A_i, x_i)\}_{i=1}^n$. Suppose that the coarse data set $\mathcal{O}$ is rationalized by some linear order P, it must be that the maximal element in A_i according to P is x_i for all i. Thus, we have $x_i P^* y$ whenever $y \in A_i$ and $y \neq x_i$, where P^* is the binary relation that is revealed from the data. It follows from the the strong axiom of revealed preference (SARP) that the coarse data set $\mathcal{O}$ is rationalizable if and only if P^* is acyclic. However, when B_i contains more than one alternative for some i, the revealed preference analysis becomes less straightforward, as illustrated by Example 1.

Now consider the general case that B_i need not be a singleton for each i. Each data point $(A_i, B_i) \in \mathcal{O}$ reveals that the maximal element in A_i is contained in B_i but not in C_i for each i. For any nonempty subcollection $\mathcal{O}' = \{(A_{k_j}, B_{k_j})\}_{j=1}^m$ $(m \leq n)$ of $\mathcal{O}$, since the maximal element in A_{k_j} is contained in B_{k_j} but not in C_{k_j} for each j, the maximal element in $A(\mathcal{O}')$ is necessarily contained in $B(\mathcal{O}')$ but not in $C(\mathcal{O}')$. This simple logic suggests the following necessary condition for rationalizability by a linear order that we term Coarse SARP:

Coarse SARP. For any $\emptyset \neq \mathcal{O}' \subset \mathcal{O}$, $B(\mathcal{O}') \setminus C(\mathcal{O}') \neq \emptyset$.

We revisit Example 1 to illustrate how to use the Corase SARP property to show that the coarse data set is not rationalizable by a linear order.

Example 4 (Example 1 Revisited)**.** *The coarse data set $\mathcal{O}$ is the same as in Example 1. It is easy to see that $B(\mathcal{O}) \setminus C(\mathcal{O}) = \emptyset$, which is a violation of the Coarse SARP property. Thus, we conclude that the coarse data set $\mathcal{O}$ is not rationalizable by a linear order.*

We have argued above that the Coarse SARP property is a necessary condition for a coarse data set to be rationalizable by a linear order. Theorem 1 below shows that the Coarse SARP property is also a sufficient condition for a coarse data set to be rationalizable by a linear order.

Theorem 1. *A coarse data set $\mathcal{O} = \{(A_i, B_i)\}_{i=1}^n$ is rationalizable by a linear order if and only if it satisfies the Coarse SARP property.*

[3] For any set S, we denote by $|S|$ its cardinality.

We provide the intuition of the proof (the-if part) as follows. Our proof is constructive. That is, if a coarse data set $\mathcal{O}$ satisfies the Coarse SARP property, we show that $\mathcal{O}$ is rationalizable by explicitly constructing a linear order that rationalizes $\mathcal{O}$. We first aggregate all the data points in $\mathcal{O}$. Note that the set of alternatives that might be the globally maximal elements are alternatives in $B(\mathcal{O}) \setminus C(\mathcal{O})$, which is nonempty by the Coarse SARP property. We then construct an incomplete binary relation that ranks alternatives in $B(\mathcal{O}) \setminus C(\mathcal{O})$ in arbitrary ways and ranks x above y if $x \in B(\mathcal{O}) \setminus C(\mathcal{O})$ and $y \notin B(\mathcal{O}) \setminus C(\mathcal{O})$. It is easy to see that any incomplete binary relation constructed in this way will rationalize an observation (A_i, B_i) if B_i contains some element in $B(\mathcal{O}) \setminus C(\mathcal{O})$. Then, we can remove such data points from consideration, which leads us to a simpler problem (a coarse data set with fewer data points). We then repeat the above mentioned process. Note that in each step, we construct some incomplete binary relation and face a simpler problem. Since $\mathcal{O}$ is finite, after finitely many steps, there are no more data points. We end up with a linear order (that consists of all the incomplete binary relations we have constructed in each step) that rationalizes $\mathcal{O}$. The proof below formalizes this intuition.

Proof of Theorem 1. (The-if part) Suppose that $\mathcal{O}$ satisfies the Coarse SARP property. We show that $\mathcal{O}$ is rationalizable by a linear order by explicitly constructing a linear order that rationalizes $\mathcal{O}$. We start with $\mathcal{O}_1 := \mathcal{O}$. Let $P_1 := B(\mathcal{O}_1) \setminus C(\mathcal{O}_1)$. Since $\mathcal{O}$ satisfies the Coarse WARP property, $P_1 \neq \emptyset$. We construct $\mathcal{O}_{k+1}$ and P_{k+1} by induction. Let $k \geq 1$, and suppose that we have constructed $\mathcal{O}_k$ and P_k. If $\mathcal{O}_k \neq \emptyset$, let

$$\mathcal{O}_{k+1} := \{(A_i, B_i) \in \mathcal{O}_k : B_i \cap P_k = \emptyset\};$$
$$P_{k+1} := B(\mathcal{O}_{k+1}) \setminus C(\mathcal{O}_{k+1}).$$

We claim that for $\mathcal{O}_k \neq \emptyset$, $\mathcal{O}_{k+1}$ is a proper subset of $\mathcal{O}_k$. Since $\mathcal{O}$ satisfies the Coarse SARP property, if $\mathcal{O}_k \neq \emptyset$, we have that $P_k = B(\mathcal{O}_k) \setminus C(\mathcal{O}_k) \neq \emptyset$, and that $B(\mathcal{O}_k) \cap P_k \neq \emptyset$. Therefore, there exists some $(A_i, B_i) \in \mathcal{O}_k$ that is eliminated when constructing $\mathcal{O}_{k+1}$ from $\mathcal{O}_k$. Furthermore, for each (A_i, B_i) that is eliminated in this step,

$$A_i \cap P_k = (B_i \cup C_i) \cap P_k = (B_i \cap P_k) \cup (C_i \cap P_k) = B_i \cap P_k \neq \emptyset.$$

Since for $\mathcal{O}_k \neq \emptyset$, $\mathcal{O}_{k+1}$ is a proper subset of $\mathcal{O}_k$, the construction necessarily stops after finitely many steps ($t < \infty$), when $\mathcal{O}_t \neq \emptyset$ and $\mathcal{O}_{t+1} = \emptyset$. Let

$$P_{t+1} = A(\mathcal{O}) \setminus \cup_{k=1}^{t} P_k.$$

We claim that $\{P_k\}_{k=1}^{t+1}$ constitutes a partition of the set $A(\mathcal{O})$. To see this, consider P_j and $P_{j'}$ where $j \neq j'$. Without loss of generality, we assume that $j' > j$. Note that for any $(A_i, B_i) \in \mathcal{O}'_j$, it must be that $A_i \cap P_j = B_i \cap P_j = \emptyset$. Otherwise, (A_i, B_i) would have been eliminated in earlier steps. Therefore, we have $B(\mathcal{O}_{j'}) \cap P_j = \emptyset$. Since $P_{j'} = B(\mathcal{O}_{j'}) \setminus B(\mathcal{O}_{j'}) \subset B(\mathcal{O}_{j'})$, $P_{j'} \cap P_j = \emptyset$.

Next, we define the linear order P on $A(\mathcal{O})$ such that xPy if $x \in P_j$, $y \in P_{j'}$, $j' > j$. This is possible since P_j and $P_{j'}$ are disjoint whenever $j \neq j'$. We claim that any linear order P satisfying this property necessarily satisfies that $\max(A_i, P) \in B_i$ for all i. For any $(A_i, B_i) \in \mathcal{O}$, $(A_i, B_i) \in \mathcal{O}_k \setminus \mathcal{O}_{k+1}$ for some $k \geq 1$. Since $B_i \cap P_k \neq \emptyset$, and $A_i \cap P_{k'} = \emptyset$ whenever $k' < k$,

$$\max(A_i, P) \in A_i \cap P_k = B_i \cap P_k \subset B_i.$$

If $A(\mathcal{O}) \neq X$, we can extend P to be defined on X in arbitrary ways such that P rationalizes $\mathcal{O}$. This completes the proof. □

Remark 1. *The Coarse SARP property reduces to the SARP property in the special case that B_i is a singleton for all i. Without loss of generality, we rewrite $\mathcal{O}$ as $\{(A_i, x_i)\}_{i=1}^n$.*

(a) Suppose that the Coarse SARP property is violated, there exists $\mathcal{O}' = \{(A_{k_j}, x_{k_j})\}_{j=1}^m$ such that $B(\mathcal{O}') \setminus C(\mathcal{O}') = \emptyset$. Then for all x_{k_j}, we have $x_{k_j} \in C_{k_{j'}}$ for some j' and therefore, $x_{k_{j'}} P^* x_{k_j}$. Since this is true for all x_{k_j}, by the finiteness of $B(\mathcal{O}')$, P^* is cyclical and the SARP property is violated.

(a) Suppose that the SARP property is violated, we can find a sequence $\{x_{k_l}\}_{l=1}^L$ such that $x_{k_l} \in A_{k_{l+1}}$ with $l = 1, 2, \ldots, L-1$ and $x_{k_L} \in A_{k_1}$. Consider the subcollection of data points $\mathcal{O}' = \{(A_{k_l}, x_{k_l})\}_{l=1}^L$. It is easy to see that $B(\mathcal{O}') \setminus C(\mathcal{O}') = \emptyset$. Therefore, the Coarse SARP property is violated.

Next, we propose a simple algorithm to check whether a coarse data set is rationalizable or not (or equivalently, whether the Coarse SARP condition is satisfied).

Stratification Algorithm.

Step 1. Set $k := 1$ and $\mathcal{O}' := \mathcal{O}$.

Step 2. Define $\mathcal{O}_k := \mathcal{O}'$. If $\mathcal{O}_k = \emptyset$, stop and output *Rationalizable*; otherwise, proceed to Step 3.

STEP 3. Define $P_k := B(\mathcal{O}_k) \setminus C(\mathcal{O}_k)$. If $P_k = \emptyset$, stop and output *Not Rationalizable*. Otherwise, set $\mathcal{O}' := \{(A_i, B_i) \in \mathcal{O}_k : B_i \cap P_k = \emptyset\}$. Derive k' such that $k' = k + 1$. Set $k := k'$. Go to STEP 2.

It follows from the proof of Theorem 1 that the STRATIFICATION ALGORITHM can be used to determine whether a data set is rationalizable or not. Notably, the algorithm is efficient. Note that the algorithm removes at least one data point and one alternative at each step. Therefore, for an arbitrary coarse data set $\mathcal{O} = \{A_i, B_i\}_{i=1}^n$, the number of steps required in the algorithm is at most $\min\{|A(\mathcal{O})|, n\}$. In particular, this implies that the algorithm works well when the total number of elements are huge but the number of data point is relatively small.

When a coarse data set $\mathcal{O} = \{(A_i, B_i)\}_{i=1}^n$ is rationalizable by a linear order, the linear order that rationalizes $\mathcal{O}$ is not necessarily unique. As shown in the proof of Theorem 1, the multiplicity at least stems from two sources: (1) we can rank the alternatives in P_k for any k in arbitrary ways; and (2) in cases in which $A(\mathcal{O}) \neq X$, we can extend the linear order P defined on $A(\mathcal{O})$ to be defined on X in arbitrary ways. Given the multiplicity of rationalizing linear orders, we are interested to identify for a given pair of alternatives x and y, whether one alternative is ranked above the other for every linear order that rationalizes $\mathcal{O}$. We say that x is surely ranked above y, denoted by $xP^s y$, if for each linear order P that rationalizes $\mathcal{O}$, it holds that xPy. It is easy to see that P^s is transitive, since each linear order that rationalizes $\mathcal{O}$ is transitive. In what follows, we assume that the coarse data set $\mathcal{O}$ is rationalizable.

In the reminder of this section, we seek to identify whether $xP^s y$. This is particularly easy if $(\{x, y\}, y) \in \mathcal{O}$ or $(\{x, y\}, x) \in \mathcal{O}$, as $(\{x, y\}, x)$ reveals that x is surely ranked above y and $(\{x, y\}, y)$ reveals that y is surely ranked above x. In all other cases, we can apply the idea of Varian (1982) and add one additional data point $(\{x, y\}, y)$ into the original coarse data set. We then check whether the new coarse data set $\mathcal{O}^* := \mathcal{O} \cup \{(\{x, y\}, y)\}$ is rationalizable by a linear order. If $\mathcal{O}^*$ is rationalizable by a linear order, then there exists a linear order that ranks y above x and rationalizes the original coarse data set $\mathcal{O}$. This rejects the hypothesis that $xP^s y$. If $\mathcal{O}^*$ is not rationalizable by a linear order, then it must be that every linear order that rationalizes $\mathcal{O}$ ranks x above y. In other words, $xP^s y$. Building on Theorem 1, this observation leads us towards a sharp characterization result for $xP^s y$.

Theorem 2. *Suppose that a coarse data set $\mathcal{O} = \{(A_i, B_i)\}_{i=1}^n$ is rationalizable by a linear order, then xP^sy if and only if there exists $\emptyset \neq \mathcal{O}' \subset \mathcal{O}$ such that $y \in A(\mathcal{O}')$ and $B(\mathcal{O}') \setminus C(\mathcal{O}') = x$.*

Recall that for any nonempty subcollection $\mathcal{O}' = \{(A_{k_j}, B_{k_j})\}_{j=1}^m$ of $\mathcal{O}$, the maximal element in $A(\mathcal{O}')$ is necessarily contained in $B(\mathcal{O}')$ but not in $C(\mathcal{O}')$. If $B(\mathcal{O}') \setminus C(\mathcal{O}') = x$, then x is the unique candidate that the maximal element in $A(\mathcal{O}')$. Therefore, if $yinA(\mathcal{O}')$, then any linear order that rationalizes the coarse data set necessarily ranks x above y. Theorem 2 shows that the reverse is also true.

Proof of Theorem 2. (The only if-part) Suppose that xP^sy, we show that there exists $\emptyset \neq \mathcal{O}' \subset \mathcal{O}$ such that $y \in A(\mathcal{O}')$ and $B(\mathcal{O}') \setminus C(\mathcal{O}') = \{x\}$. First suppose that there is $(A_i, B_i) \in \mathcal{O}$ such that $A_i = \{x, y\}$ and B_i is singleton. Note that it cannot be the case that $B_i = y$. If $B_i = x$, we can show the desired result by setting $\mathcal{O}' = \{A_i, B_i\}$.

In all other cases, define a new coarse data set $\mathcal{O}^*$ by adding the following data point $(\{x, y\}, x)$ to the coarse data set $\mathcal{O}$. Since xP^sy, $\mathcal{O}^*$ is not rationalizable by a linear order. By Theorem 1, we could find a nonempty subcollection $\mathcal{O}^{**} \subset \mathcal{O}^*$ such that

$$B(\mathcal{O}^{**}) \setminus C(\mathcal{O}^{**}) = \emptyset. \tag{1}$$

Since $\mathcal{O}$ is rationalizable, it must be that $(\{x, y\}, y) \in \mathcal{O}^{**}$. Define $\mathcal{O}' := \mathcal{O}^{**} \setminus \{(\{x, y\}, y)\}$. We claim that $\mathcal{O}^*$ has the desired property. Since $\mathcal{O}' := \mathcal{O}^{**} \setminus \{(\{x, y\}, y)\} \subseteq \mathcal{O}$, by Theorem 1, we have

$$B(\mathcal{O}') \setminus C(\mathcal{O}') \neq \emptyset. \tag{2}$$

Since $\mathcal{O}' := \mathcal{O}^{**} \setminus \{(\{x, y\}, y)\}$, it follows from (1) and (2) that $y \in C(\mathcal{O}') \subseteq A(\mathcal{O})$ and $B(\mathcal{O}') \setminus C(\mathcal{O}') = x$. $\square$

4 Applications

In this section, we apply our theory in Section 3 to investigate the observable restrictions of specific economic models. The theory immediately carries over to the rational choice model with imperfect observation. We also apply our theory towards some classical economic models for which the revealed preference characterizations are not provided in the literature so far including the multiple preferences model and the minimax regret model.

4.1 Rational Choice with Imperfect Observation

In this subsection, we investigate the observable restrictions of the rational choice model, with a twist that the observer has imperfect observation of the choice made by the DM. The DM has a strict preference $\succ$ over X, and she chooses $max(A_i, \succ)$ in each $A_i \in \mathcal{X}$. Unlike the standard revealed preference theory, the observer does not observe the exact choice of the DM. Rather, he only observes that the DM chooses some alternative from a subset of the feasible set. We shall call this model rational choice with imperfect observation.

We represent the observed behavior of the DM by (Σ, f), where $\Sigma \subset \mathcal{X}$ and $f(A)$ is superset of the choice of the DM in $A \in \Sigma$. The interpretation is that, the observer does not observe the exact choice of the DM, but he knows that the exact choice happens in set $f(A)$ for each $A \in \Sigma$.

For each (Σ, f), we can construct a corresponding coarse data set $\mathcal{O}$ as follows: $\{(A_i, B_i)\}_{i=1}^n$ where A_i is a nonempty subset of $\mathcal{X}$, and $\emptyset \neq B_i \subset A_i$. It is easy to see that (Σ, f) is rationalizable under rational choice with imperfect observation if and only if $\mathcal{O}$ is rationalizable by a linear order. The results in Section 3 immediately carry over to this setting, and we shall not repeat the arguments here.

4.2 Multiple Preferences

In this subsection, we investigate the observable restrictions of the multiple preferences model; see, for example, Kalai et al. (2002) and Salant and Rubinstein (2008). In contrast with the single preference model, the choice behavior of the DM may be a result of utility maximizing under multiple rationales. Formally, the DM has a set of strict preferences $\{\succ^j\}_{j\in J}$ defined on X, and she chooses

$$\left\{x \in A : x = \max(A, \succ^j) \text{ for some } j \in J\right\}$$

for each feasible set A.

We represent the choice behavior of the DM by (Σ, f), where $\Sigma \subseteq \mathcal{X}$ and $f(A)$ is the set of all alternatives that the DM chooses in $A \in \Sigma$. For any set of strict preferences $\{\succ^j\}_{j\in J}$, we write

$$f_J(A) := \left\{x \in A : x = \max(A, \succ^j) \text{ for some } j \in J\right\}$$

for each $A \in \Sigma$. We say that (Σ, f) is rationalizable by multiple preferences if there exists a set of strict preferences $\{\succ^j\}_{j\in J}$ such that

$$f_J(A) = f(A)$$

for each $A \in \Sigma$. In what follows, we identify necessary and sufficient conditions under which the choice behavior of the DM can be rationalizable by multiple preferences.

For each (Σ, f), consider the corresponding coarse data set $\mathcal{O} = \{A, f(A)\}_{A\in\Sigma}$. If (Σ, f) is rationalizable by multiple preferences, then there must exist at least one preference $\succ$ defined on X such that $\max(A, \succ) \in f(A)$ for each $A \in \Sigma$. In other words, the corresponding coarse data set $\mathcal{O}$ must be rationalizable by a linear order, and the Coarse SARP property is a necessary condition for rationalizability by multiple preferences. The Coarse SARP property is not a sufficient condition, as the following example illustrates.

Example 5 (A data set that is not rationalizable by multiple preferences)**.** *Let $X = \{x, y, z\}$. Consider a data set (Σ, f) including the following three observations: $f(\{x, y\}) = \{x\}$, $f(\{y, z\}) = \{y\}$, and $f(\{x, y, z\}) = \{x, z\}$. It is easy to check that the corresponding coarse data set satisfies the Coarse SARP property. We claim that (Σ, f) is not rationalizable by multiple preferences. Suppose that (Σ, f) is rationalizable by multiple preferences, by definition, there exists a set of strict preferences $\{\succ^j\}_{j\in J}$ such that $f_J(A) = f(A)$. Then $f(\{x, y\}) = \{x\}$ reveals that $x \succ^j y$ for all $j \in J$, and $f(\{y, z\}) = \{y\}$ reveals that $y \succ^i z$ for all $j \in J$. Thus, it must be the case that $x \succ^j z$ for all $j \in J$, which contradicts with the third observation that $f(\{x, y, z\}) = \{x, z\}$.*

For rationalizability by multiple preferences, we need to ensure that, for each $A \in \Sigma$ and each $x \in f(A)$, x is chosen in A under some preference. Furthermore, the preference needs to ensure that the alternative chosen in any other set $B \in \Sigma$ lies in $f(B)$. This suggests the following divide-and-conquer approach for the observable restrictions of the multiple preferences model.

For each (Σ, f), we consider the following family of coarse data sets $\mathcal{D} := \{\mathcal{O}_{A,x}\}_{A\in\Sigma,\, x\in f(A)}$. Each coarse data set $\{\mathcal{O}_{x,A}\}$ is constructed as follows: $\mathcal{O}_{A,x} := \{(B, f_{A,x}(B))\}_{B\in\Sigma}$ with $f_{A,x}(A) = \{x\}$ and $f_{A,x}(B) = f(B)$ for $B \in \Sigma$ with $B \neq A$. That is, for each $A \in \Sigma$ and each $x \in f(A)$, $\mathcal{O}_{A,x}$ is derived from $\mathcal{O}$ by replacing $(A, f(A))$ with $(A, \{x\})$. The following theorem shows that (Σ, f) is rationalizable by multiple preferences if and only if each $\mathcal{O}_{A,x}$ in $\mathfrak{D}$ is rationalized by some linear order.

Theorem 3. (Σ, f) *is rationalizable by multiple preferences if and only if each* $\mathcal{O}_{A,x}$ *in* $\mathfrak{D}$ *is rationalizable by some linear order.*

Proof of Theorem 3. (The only if-part) Suppose that some $\mathcal{O}_{A,x}$ in $\mathcal{D}$ is not rationalizable by a linear order. By definition, there is no linear order P such that $\max(A, P) = x \in f(A)$ and $\max(B, P) \in f(B)$ for $B \in \Sigma$ and $B \neq A$. Equivalently, there is no strict preference $\succ$ defined on X such that $\max(A, \succ) = x \in f(A)$ and $\max(B, \succ) \in f(B)$ for $B \in \Sigma$ and $B \neq A$. This implies that (Σ, f) is not rationalizable by multiple preferences.

(The if-part) Suppose that $\mathcal{O}_{A,x}$ is rationalizable for each $A \in \Sigma$ and $x \in A$. Denote the linear order $P_{A,x}$ as a linear order that rationalizes $\mathcal{O}_{A,x}$. This implies that $\max(A, P) = x \in f(A)$ and $\max(B, P) \in f(B)$ for $B \in \Sigma$ and $B \neq A$. Denote the strict preference induced by $P_{A,x}$ defined on X by $\succ_{A,x}$. Consider the set of strict preferences $\{\succ_{A,x}\}_{A\in\Sigma,\, x\in f(A)}$. It is easy to see that for any $\succ_{A,x}$, $\max(B, \succ_{A,x}) \in f(B)$ for any $B \in \Sigma$. Furthermore, for any $A \in \Sigma$ and $x \in f(A)$, $\succ_{A,x}$ satisfies that $\max(A, \succ_{A,x}) = y$. Therefore, we have identified a set of strict preferences $\{\succ_{A,x}\}_{A\in\Sigma,\, x\in f(A)}$ that rationalizes (Σ, f) under the multiple preferences model. $\square$

Building upon our general theory in Section 3, Theorem 3 provides a tractable way to characterize the rationalizability of (Σ, f) by multiple preferences. For notational simplicity, we denote by $g(A) := A \setminus f(A)$ the set of alternatives are not chosen for each feasible set $A \in \Sigma$.

In what follows, we provide an axiomatization of the multiple preferences model. Building upon Theorem 1 and Theorem 3, we show that the multiple preferences model can be characterized by the Sen's α axiom (see Sen (1971)) and an additional axiom that we call betweenness. The Sen's α axiom says that if x is not chosen in set A, then x is not chosen in any superset of A. The betweenness axiom says that if B is a superset of A and the chosen alternatives in B coincide with those in A, then the alternatives chosen in any set that contains A but is contained in B must be the same alternatives chosen in A. Formally,

Sen's α Axiom. For $A, B \in \mathcal{X}$, $f(A \cup B) \cap A \subseteq f(A)$.

Betweenness Axiom. For $A, B \in \mathcal{X}$ with $A \subseteq B$, if $f(A) = f(B)$, then $f(C) = f(A)$ for any C such that $A \subseteq C \subseteq B$.

Theorem 4. $(\mathcal{X}, f)$ *is rationalizable by multiple preferences if and only if it satisfies the Sen's* α *axiom and the betweenness axiom.*

Proof of Theorem 4. (The only-if part) Suppose that $(\mathcal{X}, f)$ is rationalizable by multiple preferences. By definition, there exists a set of strict preferences $\{\succ^j\}_{j\in J}$ such that $f_J(A) = f(A)$ for each $A \in \mathcal{X}$. To see that the Sen's α axiom holds, suppose that $x \in f(A\cup B)\cap A$, then there exists $j \in J$ such that $x = \max(A \cup B, \succ^j)$. Therefore, $x = \max(A, \succ^j)$, and $x \in f(A)$. For the betweenness axiom, suppose that $A \subseteq C \subseteq B$ and $f(A) = f(B)$, then for any $x \in f(B)$, we have $x \in f(A) \subseteq A$. Therefore, for all $j \in J$, $\max(B, \succ^j) = \max(A, \succ^j)$. Since $A \subseteq C \subseteq B$, $\max(C, \succ^j) = \max(A, \succ^j)$ for all $j \in J$. Therefore, $f(C) = f(A)$.

(The if-part) Suppose that $(\mathcal{X}, f)$ satisfies the Sen's α axiom and the betweenness axiom, we show that $(\mathcal{X}, f)$ is rationalizable by multiple preferences. We proceed by establishing a series of claims.

Claim 1. *For any $A \in \mathcal{X}$, $f(f(A)) = f(A)$.*

Proof of Claim 1. For any $A \in \mathcal{X}$, we have $f(A) = f(f(A)\cup A)\cap f(A) \subseteq f(f(A))$, where the equality follows from the fact that $f(A) \subseteq A$, and the inclusion relation follows from the Sen's α axiom. Since $f(f(A)) \subseteq f(A)$, we have $f(f(A)) = f(A)$.

Claim 2. *For any $A \in \mathcal{X}$, if $x \in g(A)$, then $f(f(A)\cup\{x\}) = f(A)$.*

Proof of Claim 2. Note that $f(A) \subseteq f(A)\cup\{x\} \subseteq A$. It then follows from Claim 1 and the betweenness axiom that $f(f(A)\cup\{x\}) = f(A)$.

Claim 3. *For any $A, B \in \mathcal{X}$, if $f(A) \subseteq B$, then $f(B)\cap g(A) = \emptyset$.*

Proof of Claim 3. Suppose to the contrary, there exists some $x \in f(B)\cap g(A)$. Since x is not chosen in A, by Claim 2, x is not chosen in $f(A)\cup\{x\}$. By the Sen's α axiom, x is not chosen in a superset of $f(A)\cup\{x\}$. Therefore, $x \notin f(B)$. We have a contradiction.

Claim 4. *For any $B \in \mathcal{X}$ and $\emptyset \neq \mathcal{X}' \subseteq \mathcal{X}$, if $\cup_{A\in\mathcal{X}'} f(A) \setminus \cup_{A\in\mathcal{X}'} g(A) \subseteq B$, then $f(B) \cap \cup_{A\in\mathcal{X}'} g(A) = \emptyset$.*

Proof of Claim 4. Consider any $\emptyset \neq \mathcal{X}' \subseteq \mathcal{X}$, by the Sen's α axiom, $g(C) \subseteq g(\cup_{A\in\mathcal{X}'} A)$ for all $C \in \mathcal{X}'$. Therefore, $\cup_{A\in\mathcal{X}'} g(A) \subseteq g(\cup_{A\in\mathcal{X}'} A)$. We have

$$
\begin{aligned}
f(\cup_{A\in\mathcal{X}'} A) &= \cup_{A\in\mathcal{X}'} A \setminus g(\cup_{A\in\mathcal{X}'} A) \\
&\subseteq \cup_{A\in\mathcal{X}'} A \setminus \cup_{A\in\mathcal{X}'} g(A) \\
&= \cup_{A\in\mathcal{X}'} f(A) \setminus \cup_{A\in\mathcal{X}'} g(A).
\end{aligned}
$$

Thus, if $\cup_{A\in\mathcal{X}'} f(A) \setminus \cup_{A\in\mathcal{X}'} g(A) \subseteq B$, we must have that $f(\cup_{A\in\mathcal{X}'} A) \subseteq B$. By Lemma 3, $f(B) \cap g(\cup_{A\in\mathcal{X}'} A) = \emptyset$. Since $\cup_{A\in\mathcal{X}'} g(A) \subseteq g(\cup_{A\in\mathcal{X}'} A)$, we have $f(B) \cap \cup_{A\in\mathcal{X}'} g(A) = \emptyset$.

Claim 5. *The coarse data set $\mathcal{O} = \{A, f(A)\}_{A\in\Sigma}$ satisfies the Coarse SARP property.*

Proof of Claim 5. Suppose to the contrary, there exists $\emptyset \neq \mathcal{X}' \subseteq \mathcal{X}$ such that $\cup_{A\in\mathcal{X}'} f(A) \setminus \cup_{A\in\mathcal{X}'} g(A) = \emptyset$. Consider an arbitrary nonempty set $B \in \mathcal{X}'$, since $\cup_{A\in\mathcal{X}'} f(A) \setminus \cup_{A\in\mathcal{X}'} g(A) = \emptyset \subseteq B$, by Claim 4, we have $f(B) \cap \cup_{A\in\mathcal{X}'} g(A) = \emptyset$. We have arrived at a contradiction, since $f(B) \subseteq \cup_{A\in\mathcal{X}'} f(A) \subseteq \cup_{A\in\mathcal{X}'} g(A)$, where the second inclusion relation follows from that $\cup_{A\in\mathcal{X}'} f(A) \setminus \cup_{A\in\mathcal{X}'} g(A) = \emptyset$.

Finally, we proceed to show that $(\mathcal{X}, f)$ is rationalizable by multiple preferences. Suppose that $(\mathcal{X}, f)$ is not rationalizable by multiple preferences. By Theorem 3, there must exist some $A \in \Sigma$ and $x \in f(A)$ such that the coarse data set $\mathcal{O}_{x,A}$ is not rationalizable by a linear order. By Claim 5, $\mathcal{O} = \{A, f(A)\}_{A\in\Sigma}$ satisfies the Coarse SARP property. Since $\mathcal{O}_{x,A}$ is not rationalizable by a linear order, there exists $\emptyset \neq \mathcal{X}' \subset \mathcal{X}$ such that

$$\cup_{B\in\mathcal{X}'} f_{x,A}(B) \setminus \cup_{B\in\mathcal{X}'} g_{x,A}(B) = \emptyset. \tag{3}$$

Since $\mathcal{O} = \{A, f(A)\}_{A\in\mathcal{X}}$ satisfies the Coarse SARP property, it must be the case that $A \in \mathcal{X}'$. Let $\mathcal{X}'' := \mathcal{X}' \setminus \{A\}$. Again, since $\mathcal{O} = \{A, f(A)\}_{A\in\Sigma}$ satisfies the Coarse SARP property,

$$\cup_{B\in\mathcal{X}''} f_{x,A}(B) \setminus \cup_{B\in\mathcal{X}''} g_{x,A}(B) \neq \emptyset. \tag{4}$$

It follows from (3) and (4) that $\cup_{B\in\Sigma''} f_{x,A}(B) \setminus \cup_{B\in\Sigma''} g_{x,A}(B) \subseteq g_{x,A}(A) \subseteq A$, and $x \in \cup_{B\in\Sigma''} g_{x,A}(B)$. Thus, $x \in f(A) \cap \cup_{B\in\Sigma''} g_{x,A}(B)$, which contradicts with Claim 5. $\square$

4.3 Minimax Regret

The minimax regret model is first proposed by Savage (1951) to model a DM who might anticipate regret and thus incorporate in their choice the desire to minimize the worst-case regret. In this subsection, we show that, by applying our results in Section 3, one can easily obtain the observable restrictions of the minimax regret model.

We shall adopt the framework of the state space that has been discussed by, for example, Kreps (1979) and Dekel et al. (2001). Let $u : X \to R$ be a utility function for the DM. Under utility function u, the regret of choosing x towards y is represented by $u(y) - u(x)$. Given a finite set of utility functions $\mathcal{U}$, the worst-case regret of choosing x from a feasible set $A \in \mathcal{X}$ is

$$\max_{y \in A} \max_{u \in \mathcal{U}} \{u(y) - u(x)\}.$$

The DM chooses an alternative that minimizes the worst-case regret. Formally, the DM has a finite set of utility functions $\mathcal{U}$ defined on X such that she chooses

$$\min_{x \in A} \left\{ \max_{y \in A} \max_{u \in \mathcal{U}} \{u(y) - u(x)\} \right\}$$

for all $A \in \mathcal{X}$.

We represent the choice behavior of the DM by (Σ, f), where $\Sigma \subseteq \mathcal{X}$ and $f(A)$ is the alternative that the DM chooses in $A \in \Sigma$. We assume that f is a choice function. We say that (Σ, f) is rationalizable under the minimax regret model if there exists a finite set of utility functions $\mathcal{U}$ such that

$$f(A) = \arg\min_{x \in A} \left\{ \max_{y \in A} \max_{u \in \mathcal{U}} \{u(y) - u(x)\} \right\}.$$

for each $A \in \Sigma$.

Suppose that (Σ, f) is rationalizable under the minimax regret model with a finite set of utility functions $\mathcal{U}$. We define the relative regret of x towards y as

$$\phi(x, y) := \max_{u \in \mathcal{U}} \{u(y) - u(x)\}.$$

Thus, the choice function could be rewritten as

$$f(A) = \arg\min_{x \in A} \left\{ \max_{y \in A} \phi(x, y) \right\}. \tag{5}$$

Given the flexibility to construct the set of utility functions to make the choice behavior of the DM consistent with the minimax regret model, the readers might doubt, whether there are any observable restrictions of the minimax regret model. The following example presents a data set (Σ, f) that is not rationalizable under the minimax regret model.

Example 6 (A data set that is not rationalizable under the minimax regret model)**.** *Let* $X = \{x, y, z, w\}$ *and* $A = X$*. Consider a data set* (Σ, f) *including the following three observations:* $f(A) = x$*,* $f(A \setminus z) = y$*, and* $f(A \setminus w) = y$*.*

Suppose that (Σ, f) *is rationalizable under the minimax regret model. Since* $f(A) = x$ *and* $f(A \setminus z) = y$*,* x *generates a lower worst-case regret in set* A *than* y *does, but* y *generates* *a lower worst-case regret than x does if we remove the alternative z from the alternative set A. Therefore, the relative regret of y towards z is the worst-case regret of y in set A, and is higher than the worst case regret of x in set A. But then, it cannot be that $f(A \setminus w) = y$, as x generates a lower worst-case regret in set $A \setminus w$ than the relative regret of y towards z, which is the worst-case regret of y in set $A \setminus w$. We arrive at a contradiction.*

The next lemma shows that, without loss of generality, we can work with the following choice function:

$$f(A) = \arg\min_{x \in A} \left\{ \max_{y \in A \setminus x} \phi(x, y) \right\}. \tag{6}$$

This simplies the analysis below. The proof is purely algebraic, and is relegated to the Appendix.

Lemma 1. *There exists a set of utility function $\mathcal{U}$ such that (5) holds for each $A \in \Sigma$ if and only if (6) holds for each $A \in \Sigma$ under the same $\mathcal{U}$.*

By Lemma 1, the rationalizability of (Σ, f) under the minimax regret model could be transformed to the problem of finding a finite set of utility functions $\mathcal{U}$ such that

$$\max_{r \in A \setminus f(A)} \phi(f(A), r) < \max_{r \in A \setminus z} \phi(z, r)$$

for any $z \in A \setminus f(A)$.

Interestingly, the rationalizability of a data set under the minimax regret model is related to the rationalizability of a corresponding coarse data set by a linear order. Before we discuss the construction of the corresponding coarse data set, it is perhaps best to consider the following thought experiment, which illustrates the logic of the construction. Suppose that (Σ, f) includes the following observation $f(\{x, y, z\}) = x$. If (Σ, f) is rationalizable under the minimax regret model, then it must be the case that the worst-case regret of any alternative different from x in $\{x, y, z\}$ is higher than the worst-case regret of x in $\{x, y, z\}$. That is, $f(\{x, y, z\}) = x$ reveals that

$$\max\{\phi(y, x), \phi(y, z)\} > \max\{\phi(x, y), \phi(x, z)\}, \text{ and}$$
$$\max\{\phi(z, x), \phi(z, y)\} > \max\{\phi(x, y), \phi(x, z)\}.$$

For notational simplicity, in what follows, we shall use (x, y) to represent the maximal regret of x from y.

Following the logic in the last paragraph, for any (Σ, f), we are now ready construct the corresponding coarse data set. Let $\bar{X} := \{(x, y) \in X \times X : x \neq y\}$. For any (Σ, f), the corresponding coarse data set $\bar{\mathcal{O}}$ can be constructed as follows: $(\bar{A}_i, \bar{B}_i) \in \bar{\mathcal{O}}$ if and only if there exists $y \in A \in \Sigma$ with $y \neq f(A)$ such that

$$\bar{A}_i = \{y \times (A \setminus y)\} \cup \{f(A) \times (A \setminus f(A))\}, \text{ and}$$
$$\bar{B}_i = y \times (A \setminus y).$$

The interpretation of the data point $(\bar{A}_i, \bar{B}_i)$ is then that alternative y which not chosen in A generates a higher worst-case regret in A that $f(A)$ does. The following theorem shows that (Σ, f) is rationalizable under the minimax regret model if and only if the corresponding coarse data set $\bar{\mathcal{O}}$ is rationalizable by a linear order.

Theorem 5. *(Σ, f) is rationalizable under the minimax regret model if and only if the corresponding coarse data set $\bar{\mathcal{O}}$ is rationalizable by a linear order.*

Proof of Theorem 5. (The only if-part) Suppose that (Σ, f) is rationalizable under the minimax regret model, by definition, there exists a finite set of utility functions $\mathcal{U}$ such that

$$f(A) = \arg\min_{x \in A} \left\{ \max_{y \in A \setminus x} \phi(x, y) \right\}$$

holds for each $A \in \Sigma$. We claim that ϕ could represent a complete and transitive binary relation defined on $\bar{X}$. Formally, denote this binary relation as R^R which we call the regret relation. We have that $\phi(x, y) R^R \phi(z, w)$ if and only if $\phi(x, y) \geq \phi(z, w)$. For each $A \in \Sigma$,

$$f(A) = \arg\min_{x \in A}(max(x \times A \setminus z, R^R), R^R)$$

Thus, the relation R^* satisfies that for any $(\bar{A}, \bar{D}) \in \mathcal{O}$, there is some set $A \in \Sigma$ such that $\bar{A} = \{z\} \times A \backslash \{z\} \cup \{f(A)\} \times A \backslash \{f(A)\}$ with $z \neq f(A)$ and $\bar{B} = \{z\} \times A \backslash \{z\}$. As a result, it must be that there is some $y \neq z$ such that $(z, y) R^* (f(A), \hat{y})$ and not $(f(A), \hat{y}) R^* (z, y)$ for any $\hat{y} \in A \backslash \{f(A)\}$. Easy to see that if we simply assigns some tie-breaking rule to R^* and transfer it to R^{**} which is a linear order over $\bar{X}$, it still holds that there is some $y \neq z$ such that $(z, y) R^{**} (f(A), \hat{y})$ and not $(f(A), \hat{y}) R^{**} (z, y)$ for any $\hat{y} \in A \backslash \{f(A)\}$. As a result, the maximal element in each $\bar{A}$ must be contained in the set $\bar{B}$ according to the linear order R^{**}. We finish the sufficient part of the proof.

(The if-part) Suppose that $\bar{\mathcal{O}}$ is rationalizable by a linear order, by definition, there exists a linear order P defined on $\bar{X}$ such that for each $(\bar{A}, \bar{B}) \in \bar{\mathcal{O}}$, $\max(\bar{A}, P) \in \bar{B}$. Take the asymmetric part of P as $\triangleright$. It suffices to show that there exists a finite set of utility function $\mathcal{U}$ such that for any $(x, y), (z, w) \in \bar{X}$, $(x, y) \triangleright (z, w)$ if and only if $\phi(x, y) > \phi(z, w)$, where ϕ is constructed from $\mathcal{U}$.

As $\bar{X}$ is finite, we can find a function $\beta : \bar{X} \to (1, 2)$ such that $(x, y) \triangleright (z, w)$ if and only if $\beta(x, y) > \beta(z, w)$. Now consider a finite set of utility functions such that $\mathcal{U} := \{u_{x,y}\}_{(x,y)\in\bar{X}}$, where the utility functions in $\mathcal{U}$ are indexed by $(x, y) \in \bar{X}$. For each $(x, y) \in \bar{X}$, define $u_{x,y}$ as follows:

$$u_{x,y}(z) = \begin{cases} 0, & \text{if } z = x, \\ \beta(x, y), & \text{if } z = y, \\ \frac{\beta(z,w)}{2}, & \text{otherwise.} \end{cases}$$

Thus, $u_{x,y}(y) - u_{x,y}(x) = \beta(x, y) \in (1, 2)$. Furthermore, we claim that $u_{z,w}(y) - u_{z,w}(x) < 1$ if $(z, w) \neq (x, y)$. To see this, first consider the case in which $w \neq y$. By construction of the utility functions and the β function,

$$u_{z,w}(y) - u_{z,w}(x) \leq u_{z,w}(y) \leq \frac{\beta(z, w)}{2} \in (0, 1).$$

Now consider the case in which $w = y$. Since $(z, w) \neq (x, y)$, we have $z \neq x$. By construction of the utility functions, $u_{z,w}(x) = \frac{\beta(z,w)}{2}$ and $u_{z,w}(y) = \beta(z, w)$. Therefore, $u_{z,w}(y) - u_{z,w}(x) = \frac{\beta(z,w)}{2} \in (0, 1)$. This implies that

$$\phi(x, y) = \max_{u\in\mathcal{U}} [u(y) - u(x)] = u_{x,y}(y) - u_{x,y}(x) = \beta(x, y).$$

Therefore, we have explicitly constructed a finite set of utility functions $\mathcal{U}$ such that for any $(x, y), (z, w) \in \bar{X}$, $(x, y) \triangleright (z, w)$ if and only if $\phi(x, y) > \phi(z, w)$, where ϕ is constructed from $\mathcal{U}$. This completes the proof. □

We revisit Example 6 to illustrate how to use Theorem 5 to show that the data set is not rationalizable.

Example 7 (Example 6 Revisited)**.** *The data set is the same as in Example 6. We construct the corresponding coarse data set $\bar{\mathcal{O}}$ as follows:*

1. $\bar{A}_1 = \{(y, x), (y, z), (y, w), (x, y), (x, z), (x, w)\}, \bar{B}_1 = \{(y, x), (y, z), (y, w)\};$
2. $\bar{A}_2 = \{(z, x), (z, y), (z, w), (x, y), (x, z), (x, w)\}, \bar{B}_2 = \{(z, x), (z, y), (z, w)\};$
3. $\bar{A}_3 = \{(w, x), (w, y), (w, z), (x, y), (x, z), (x, w)\}, \bar{B}_3 = \{(w, x), (w, y), (w, z)\};$

4. $\bar{A}_4 = \{(x, y), (x, w), (y, x), (y, w)\}, \bar{B}_4 = \{(x, y), (x, w)\}$*;*

5. $\bar{A}_5 = \{(w, x), (w, y), (y, x), (y, w)\}, \bar{B}_5 = \{(w, x), (w, y)\}$*;*

6. $\bar{A}_6 = \{(x, y), (x, z), (y, x), (y, z)\}, \bar{B}_6 = \{(x, y), (x, z)\}$*;*

7. $\bar{A}_7 = \{(z, x), (z, y), (y, x), (y, z)\}, \bar{B}_7 = \{(z, x), (z, y)\}$*.*

By Theorem 1, the coarse data set $\bar{\mathcal{O}}$ is not rationalizable by a linear order. In particular, for $\bar{\mathcal{O}}' = \{(\bar{A}_1, \bar{B}_1), (\bar{A}_4, \bar{B}_4), (\bar{A}_6, \bar{B}_6)\}$, $B(\bar{\mathcal{O}}') \setminus C(\bar{\mathcal{O}}') = \emptyset$. By Theorem 5, (Σ, f) is not rationalizable under the minimax regret model.

5 Weak order

Up to now, our analysis has centered around the rationalizability of a coarse data set by a linear order. This section extends our analysis to the case of weak order. A weak order is a complete and transitive binary relation, denote by R. We abuse the notation slightly and denote the asymmetric part of R by P. We write $\max(A, R)$ to represent the best alternative(s) in A according to R. If xRy, we say that x is weakly ranked above y, and if not xRy, we say x is not weakly ranked above y. If xRy and not yRx, we say x is strictly ranked above y.

One way to define rationalizability of a coarse data set by a weak order is as follows: a coarse data set $\mathcal{O} = \{(A_i, B_i)\}_{i=1}^n$ is rationalized by a weak order R if

$$\max(A_i,\ R) \cap B_i \neq \emptyset$$

for all i. Note that every coarse data set is rationalizable by a weak order as the DM could be indifferent between all alternatives in X.

Reveal preference analysis starts with data and derives implications about the DM's preferences. In many settings, we have some prior knowledge of the DM's preferences with a typical example being that the DM has a monotone preference. Then, we should incorporate the prior knowledge of the DM's preferences into the revealed preference exercise. Importantly, we can utilize such prior knowledge to obtain additional information about preferences. For a simple example, suppose that the observer has prior knowledge that the DM ranks x strictly above y. Then if the DM chooses z from $\{x, y, z\}$, this reveals that the DM weakly ranks z over x and strictly ranks z over y.

Throughout the rest of this section, we shall model a data point to be a tuple (A, B, D) where A is the feasible set, B and D are disjoint subsets of A, and $B \cup D \neq \emptyset$. The interpretation of (A, B, D) is that when the DM is faced with a choice problem A, the observer knows that the exact choice of the DM is contained in B. Furthermore, she also knows the alternatives in D are not maximal for the DM in A.

A coarse data set $\mathcal{O}$ is a set of data points $\{(A_i, B_i, D_i)\}_{i=1}^n$, where B_i and D_i are disjoint subsets of A_i, and $B_i \cup D_i \neq \emptyset$ for all i. We say that $\mathcal{O}$ is rationalizable by a weak order if there exists a weak order R defined on X such that for all i,

$$\max(A_i,\ R) \cap B_i \neq \emptyset \text{ and} \max(A_i,\ R) \cap D_i = \emptyset.$$

As in the case of linear order, we identify a necessary and sufficient condition for the coarse data set to be rationalizable by a weak order, provide an efficient algorithm, and then discuss some identification issues. For ease of notation, we write

$$D(\mathcal{O}') = \cup_{(A_i,B_i,D_i)\in\mathcal{O}'} D_i.$$

A natural conjecture is that we can work with $B(\mathcal{O}') \setminus D(\mathcal{O}')$ and check whether it is empty for each subcollection $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$. Indeed, if a coarse data set $\mathcal{O} = \{A_i, B_i, D_i\}_{i=1}^n$ is rationalizable by a weak order, by definition, for all (A_i, B_i, D_i), B_i contains an alternative that is strictly ranked above all alternatives in D_i. Thus, for each subcollection $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$, $B(\mathcal{O}')$ contains an alternative that is strictly ranked above all alternatives in $D(\mathcal{O}')$, and $B(\mathcal{O}') \setminus D(\mathcal{O}') \neq \emptyset$. While this is necessary, this is not sufficient for the rationalizability of a coarse data set by a weak order, as the following example illustrates.

Example 8. *Let $X = \{x_1, x_2, x_3, x_4\}$. Consider a coarse data set $\mathcal{O}$ including the following two observations:*

1. *$A_1 = \{x_1, x_2, x_3\}, B_1 = x_1, D_1 = x_2$;*
2. *$A_2 = \{x_1, x_2, x_4\}, B_2 = x_2, D_2 = x_3$.*

It is clear that for each subcollection $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$, $B(\mathcal{O}') \setminus D(\mathcal{O}') \neq \emptyset$. Suppose the coarse data set $\mathcal{O}$ is rationalizable by a weak order R with asymmetric part P. The first observation reveals that $x_1 P x_2$, and the second observation reveals that $x_2 R x_1$. By transitivity of the weak order, $x_1 P x_1$. We arrive at a contradiction. Thus, the coarse data set $\mathcal{O}$ is not rationalizable by a weak order.

Example 8 demonstrates that simply working with $B(\mathcal{O}') \setminus D(\mathcal{O}')$ for each subcollection $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$ overlooks the information that for all (A_i, B_i, D_i), some alternative in B_i is ranked weakly above all alternatives in A_i. In what follows, we discuss an alternative procedure that incorporates such information, and identify a necessary and sufficient condition for the rationalizability of a coarse data set by a weak order.

Suppose that a coarse data set $\mathcal{O} = \{A_i, B_i, D_i\}_{i=1}^n$ is rationalizable by a weak order. By definition, for all (A_i, B_i, D_i), B_i contains an alternative that is strictly ranked above all alternatives in D_i and weakly ranked above all alternatives in A_i. As such, $B(\mathcal{O}')$ is directly revealed to contain an alternative that is strictly ranked above all alternatives in $D(\mathcal{O}')$. Suppose that for some $(A_i, B_i, D_i) \in \mathcal{O}'$, $B_i \subseteq D(\mathcal{O}')$. Since B_i contains an alternative that is weakly ranked above all alternatives in A_i, $D(\mathcal{O}')$ contains an alternative that is wealy ranked above all alternatives in A_i. By transitivity of weak order, $B(\mathcal{O}')$ contains an alternative that is strictly ranked above all alternatives in A_i.

We define the following operator $\mathcal{E}^1$ such that for all $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$ and $\emptyset \neq E \in \mathcal{X}$,

$$\mathcal{E}^1(\mathcal{O}',\, E) := (\cup_{(A_i,B_i,D_i)\in\mathcal{O}':B_i\subseteq E}\, A_i) \cup E.$$

It follows that E contains an alternative that is weakly ranked above all alternatives in $\mathcal{E}^1(\mathcal{O}', E)$. Recursively, we can define

$$\mathcal{E}^n(\mathcal{O}',\, E) := \mathcal{E}^1(\mathcal{O}',\, \mathcal{E}^{n-1}(\mathcal{O}', E)).$$

By the same reasoning as above, for all n, $\mathcal{E}^{n-1}(\mathcal{O}',\, E)$ contains an alternative that is weakly ranked above to all alternatives in $\mathcal{E}^n(\mathcal{O}', E)$. By transitivity of the weak order, we can conclude that E concludes at least one alternative that is weakly ranked above to every alternative in $\mathcal{E}^n(\mathcal{O}', E)$ for all n.

Note that $\mathcal{E}^1(\mathcal{O}', E)$ is increasing in E and $E \subseteq \mathcal{E}^1(\mathcal{O}', E)$. Therefore,

$$\mathcal{E}^n(\mathcal{O}', E) \subseteq \mathcal{E}^{n+1}(\mathcal{O}', E)$$

for all n. Let N be the smallest integer such that $\mathcal{E}^N(\mathcal{O}', E) = \mathcal{E}^{N+1}(\mathcal{O}', E)$. Since X is finite, $\mathcal{E}^n(\mathcal{O}', E)$ necessarily converges to $\mathcal{E}^N(\mathcal{O}', E)$ in finitely many steps. Since $D(\mathcal{O}')$ contains at least one alternative that is weakly ranked above every alternative in $\mathcal{E}(\mathcal{O}', D(\mathcal{O}'))$ and $B(\mathcal{O}')$ contains an alternative that is strictly ranked above all alternatives in $D(\mathcal{O}')$, by transitivity of weak order, $B(\mathcal{O}')$ necessarily contains an alternative that is strictly ranked above all alternatives in $\mathcal{E}(\mathcal{O}', D(\mathcal{O}'))$. As such, we have $B(\mathcal{O}') \setminus \mathcal{E}(\mathcal{O}', D(\mathcal{O}')) \neq \emptyset$. For notational simplicity, we write $\mathcal{E}(\mathcal{O}')$ rather than $\mathcal{E}(\mathcal{O}', D(\mathcal{O}'))$. This suggests the following necessary condition for rationalizability that we call Coarse WARP:

Coarse WARP. For any $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$, $B(\mathcal{O}') \setminus \mathcal{E}(\mathcal{O}') \neq \emptyset$.

The theorem below shows that Coarse WARP is also a sufficient condition for rationalizability of a coarse data set by a weak order.

Theorem 6. *A coarse data set is rationalizable by a weak order if and only if it satisfies the Coarse WARP property.*

Proof of Theorem 6. (The if part) Suppose that $\mathcal{O}$ satisfies the Coarse WARP property. We show that $\mathcal{O}$ is rationalizable by a weak order by explicitly constructing a weak order that rationalizes $\mathcal{O}$. We start with $\mathcal{O}_1 := \mathcal{O}$. Let $P_1 := B(\mathcal{O}_1) \setminus \mathcal{E}(\mathcal{O}_1)$. Since $\mathcal{O}$ satisfies the Coarse WARP property, $P_1 \neq \emptyset$. We construct $\mathcal{O}_{k+1}$ and P_{k+1} by induction. Let $k \geq 1$, and suppose that we have constructed $\mathcal{O}_k$ and P_k. If $\mathcal{O}_k \neq \emptyset$, let

$$\begin{aligned} \mathcal{O}_{k+1} &:= \{(A_i, B_i, D_i) \in \mathcal{O}_k : B_i \cap P_k = \emptyset\}; \\ P_{k+1} &:= B(\mathcal{O}_{k+1}) \setminus \mathcal{E}(\mathcal{O}_{k+1}). \end{aligned}$$

We claim that for $\mathcal{O}_i \neq \emptyset$, $\mathcal{O}_{k+1}$ is a proper subset of $\mathcal{O}_k$. Since $\mathcal{O}$ satisfies the Coarse WARP property, if $\mathcal{O}_k \neq \emptyset$, we have that $P_k = B(\mathcal{O}_k) \setminus \mathcal{E}(\mathcal{O}_k) \neq \emptyset$, and that $B(\mathcal{O}_k) \cap P_k \neq \emptyset$. Therefore, there exists some $(A_i, B_i, D_i) \in \mathcal{O}_k$ that is eliminated when constructing $\mathcal{O}_{k+1}$ from $\mathcal{O}_k$. Let K be the smallest integer such that $\mathcal{O}_K \neq \emptyset$ and $\mathcal{O}_{K+1} = \emptyset$. Thus, $(P_k, \mathcal{O}_k)_{k=1}^K$ are well-defined. Let $P_{K+1} := X \setminus (\cup_{k=1}^K P_k)$.

Claim 6. $B(\mathcal{O}_{k+1}) \subseteq \mathcal{E}(\mathcal{O}_k)$ *for* $k = 1, 2, ..., K-1$.

Proof of Claim 6. For any $(A_i, B_i, D_i) \in \mathcal{O}_{k+1}$, by construction of $\mathcal{O}_{k+1}$ and P_k, it must be that

$$B_i \cap P_k = B_i \cap (B(\mathcal{O}_k) \setminus \mathcal{E}(\mathcal{O}_k)) = \emptyset. \tag{7}$$

Otherwise, (A_i, B_i, D_i) would have been eliminated in earlier steps. Since $(A_i, B_i, D_i) \in \mathcal{O}_{k+1} \subseteq \mathcal{O}_k$, $B_i \subseteq B(\mathcal{O}_k)$. It then follows from (7) that $B_i \subseteq \mathcal{E}(\mathcal{O}_k)$. This is true for all $(A_i, B_i, D_i) \in \mathcal{O}_{k+1}$, and we have $B(\mathcal{O}_{k+1}) \subseteq \mathcal{E}(\mathcal{O}_k)$.

Claim 7. $\mathcal{E}(\mathcal{O}_{k+1}) \subseteq \mathcal{E}(\mathcal{O}_k)$ *for* $k = 1, 2, ..., K-1$.

Proof of Claim 7. Note that $\mathcal{E}^1(\mathcal{O}', E) = (\cup_{(A,B,D)\in\mathcal{O}':B\subset E} A) \cup E$ is increasing in both $\mathcal{O}'$ and E in the set inclusion sense. Therefore,

$$\mathcal{O}_{k+1} \subseteq \mathcal{O}_k \implies \mathcal{E}^1(\mathcal{O}_{k+1}, D(\mathcal{O}_{k+1})) \subset \mathcal{E}^1(\mathcal{O}_k, D(\mathcal{O}_k)).$$

By induction in n, $\mathcal{E}^n(\mathcal{O}_{k+1}, D(\mathcal{O}_{k+1})) \subseteq \mathcal{E}^n(\mathcal{O}_k, D(\mathcal{O}_k))$ for all n. Therefore, $\mathcal{E}(\mathcal{O}_{k+1}) \subseteq \mathcal{E}(\mathcal{O}_k)$.

Claim 8. *$\{P_k\}_{k=1}^{K+1}$ constitutes a partition of X.*

Proof of Claim 8. Recall that $P_{K+1} := X \setminus (\cup_{k=1}^{K} P_k)$. It suffices to show that $P_i \cap P_j = \emptyset$ for $1 \le i < j \le K$. This is because

$$P_i = B(\mathcal{O}_i) \setminus \mathcal{E}(\mathcal{O}_i), \text{ and}$$
$$P_j = B(\mathcal{O}_j) \setminus \mathcal{E}(\mathcal{O}_j) \subseteq B(\mathcal{O}_j) \subseteq \mathcal{E}(\mathcal{O}_{j-1}) \subseteq \mathcal{E}(\mathcal{O}_i),$$

where the second set inclusion follows from Claim 6 and the third set inclusion follows from Claim 7. Thus, $P_i \cap P_j = \emptyset$.

Next, we proceed to construct a weak order R defined on X such that (1) xRy and yRx if and only if $x, y \in P_k$ for some $k = 1, \dots, K+1$; and (2) xPy if and only if $x \in P_i$ and $y \in P_j$ with $j > i$. By Claim 8, R is well defined. Furthermore, R is a weak order. In what follows, we show that R rationalizes the coarse data set.

Claim 9. $\mathcal{E}^1(\mathcal{O}', \mathcal{E}(\mathcal{O}')) = \mathcal{E}(\mathcal{O}')$.

Proof of Claim 9. This follows from the definition of $\mathcal{E}(\mathcal{O}')$.

Claim 10. *Suppose that $(A, B, D) \in \mathcal{O}'$, $A \subseteq \mathcal{E}(\mathcal{O}')$ if and only if $B \subseteq \mathcal{E}(\mathcal{O}')$.*

Proof of Claim 10. The only-if part is trivial. For other direction, if $B \subseteq \mathcal{E}(\mathcal{O}')$, then $A \subseteq \mathcal{E}^1(\mathcal{O}', \mathcal{E}(\mathcal{O}')) = \mathcal{E}(\mathcal{O}')$, where the set inclusion follows from the definition of $\mathcal{E}^1$ and the equality follows from Claim 9.

Consider an arbitrary data point $(A, B, D) \in \mathcal{O}_k \setminus \mathcal{O}_{k+1}$. Since $(A, B, D) \in \mathcal{O}_k \subseteq \mathcal{O}_{k-1}$, $B \subseteq B(\mathcal{O}_{k-1})$. Note that $(A, B, D) \in \mathcal{O}_k$ implies $B \cap P_{k-1} = \emptyset$. Thus, $B \subseteq \mathcal{E}(\mathcal{O}_{k-1})$ by the construction of P_{k-1}. By Claim 10, $A \subseteq \mathcal{E}(\mathcal{O}_{k-1})$. Since $P_i = B(\mathcal{O}_i) \setminus \mathcal{E}(\mathcal{O}_i)$, we have $P_i \cap \mathcal{E}(\mathcal{O}_i) = \emptyset$ and by Claim 7 that $P_i \cap \mathcal{E}(\mathcal{O}_j) = \emptyset$ for all $j \ge i$, which implies $A \cap P_i = \emptyset$ for all $i \le k-1$. Therefore, $A \subseteq \cup_{j=k}^{K+1} P_j$ by Claim 8. Since $(A, B, D) \in \mathcal{O}_k \setminus \mathcal{O}_{k+1}$, we know that $B \cap P_k \neq \emptyset$ and thus $A \cap P_k \neq \emptyset$. Hence, $P_k = \max(\cup_{j=k}^{K+1} P_j, R)$ implies that $\max(A, R) \subseteq P_k$ and $\max(A, R) \cap B \neq \emptyset$. Since $P_k \cap \mathcal{E}(\mathcal{O}_k) = \emptyset$ and $D \subseteq \mathcal{E}(\mathcal{O}_k)$, we have $\max(A, R) \cap D = \emptyset$. □

Next, we propose a simple algorithm to check whether a coarse data set is rationalizable or not (or equivalently, whether the Coarse SARP condition is satisfied). The algorithm replicates our proof to check whether the coarse data set $\{(A_i, B_i, D_i)\}_{i=1}^n$ is rationalizable by a weak order. k and $\mathcal{O}'$ are variables.

Stratification Algorithm*.

Step 1. Set $k := 1$ and $\mathcal{O}' := \mathcal{O}$.

Step 2. Define $\mathcal{O}_k := \mathcal{O}'$. If $\mathcal{O}_k = \emptyset$, stop and output *Rationalizable*; otherwise, proceed to Step 3.

Step 3. Define $P_k := B(\mathcal{O}_k) \setminus \mathcal{E}(\mathcal{O}_k)$. If $P_k = \emptyset$, stop and output *Not Rationalizable*; otherwise, set $\mathcal{O}' := \{(A, B, D) \in \mathcal{O}_k : B \cap P_k = \emptyset\}$. Derive k' such that $k' = k + 1$. Set $k := k'$. Go to Step 2.

We now proceed to identify for a given pair of alternatives x and y, whether one alternative is ranked above in the other for every weak order that rationalizes $\mathcal{O}$. We say that x is surely weakly ranked above y, denoted by $xR^s y$, if for each weak order R that rationalizes $\mathcal{O}$, it holds that xRy. It is easy to see that R^s is transitive, since each weak order that $\mathcal{O}$ is transitive. The definition for P^s is similar. We seek to identify the relation of $xR^s y$ and $xP^s y$, i.e. $xR^s y$ means that for any weak order R that rationalizes the choice data, xRy. In what follows, we assume that the coarse data set $\mathcal{O}$ is rationalizable by a weak order.

To test $xR^s y$, we add the data point $(\{x, y\}, y, x)$ into the original coarse data set $\mathcal{O}$, and work with the new coarse data set $\mathcal{O}^* := \mathcal{O} \cup (\{x, y\}, y, x)$. If $\mathcal{O}^*$ is rationalizable by a weak order, then there exists a weak order that strictly ranks y above x and rationalizes the original coarse data set $\mathcal{O}$. The hypothesis that $xR^s y$ is rejected. If $\mathcal{O}^*$ is not rationalizable by a weak order, then every weak order R that rationalizes $\mathcal{O}$ necessarily satisfies that xRy. In other words, $xR^s y$. Theorem 7 below provides a necessary and sufficient condition for $xR^s y$. The intuition is that when $xR^s y$, we can find a nonempty subcollection $\mathcal{O}'$ such that $x, y \in A(\mathcal{O}')$ and x is among the maximal elements in $A(\mathcal{O}')$ for all weak order that rationalize $\mathcal{O}$. Then whenever we stratify $A(\mathcal{O}')$, x must be in the higher ranked strata. For notational simplicity, we define $\mathcal{O}'_{B^*} := \{(A, B, D) \in \mathcal{O}' : B \cap B^* = \emptyset\}$ for any $B^* \subseteq X$.

Theorem 7. *Suppose that a coarse data set $\mathcal{O} = \{(A_i, B_i, D_i)\}_{i=1}^n$ is rationalizable by a weak order, then $xR^s y$ if and only if there exists $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$ such that $x, y \in A(\mathcal{O}')$, and for all $B^* \subseteq B(\mathcal{O}') \setminus \mathcal{E}(\mathcal{O}')$ with B^* nonempty, $B^* \cap A(\mathcal{O}'_{B^*}) = \emptyset$ implies $x \in B^*$.*

Proof. (The if-part.) It suffices to show that $x \in \max(A(\mathcal{O}'), R^*)$ for any weak order R^* that rationalizes $\mathcal{O}$. Since $\emptyset \neq \mathcal{O}' \subseteq \mathcal{O}$, R^* also rationalizes $\mathcal{O}'$. Therefore,

$$\max(A(\mathcal{O}'), R^*) \cap \mathcal{E}(\mathcal{O}') = \emptyset \text{ and } \max(A(\mathcal{O}'), R^*) \cap B(\mathcal{O}') \neq \emptyset.$$

Suppose to the contrary, $x \notin \max(A(\mathcal{O}'), R^*)$.

Let $B^* = \max(A(\mathcal{O}'), R^*) \cap B(\mathcal{O}')$. Then $\emptyset \neq B^* \subseteq B(\mathcal{O}') \setminus \mathcal{E}(\mathcal{O}')$. We claim that $B^* \cap A(\mathcal{O}'_{B^*}) = \emptyset$, which contradicts $x \notin B^*$. To see the claim, suppose to the contrast that $B^* \cap A \neq \emptyset$ for some $(A, B, D) \in \mathcal{O}'_{B^*} \subset \mathcal{O}'$. Then $\max(A; R^*) \cap \max(A(\mathcal{O}'); R^*) \neq \emptyset$ and thus $\max(A, R^*) \subseteq \max(A(\mathcal{O}'), R^*)$. Therefore, $B^* \cap B = \max(A(\mathcal{O}'), R^*) \cap B(\mathcal{O}') \cap B = \max(A(\mathcal{O}'), R^*) \cap B \supseteq \max(A, R^*) \cap B \neq \emptyset$, a contradiction.

(The only if-part.) $xR^s y$ implies that we could find $\mathcal{O}'$ such that if we add the observation $(\{x, y\}, \{y\}, \{x\})$, the new constructed data $\mathcal{O}^* := \mathcal{O}' \cup \{(\{x, y\}, \{y\}, \{x\})\}$ violates coarse WARP, i.e. $B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*) = \emptyset$.

We first show that $x, y \in A(\mathcal{O}')$. Suppose that $y \notin A(\mathcal{O}')$. Note in addition that $y \notin D(\mathcal{O}^*)$. By the definition of $\mathcal{E}(\mathcal{O}^*)$, we know $y \notin \mathcal{E}(\mathcal{O}^*)$, which leads to $y \in B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*)$, a contradiction. Now suppose that $x \notin A(\mathcal{O}')$. Since $x \notin B(\mathcal{O}')$, we have that

$$\begin{aligned} B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*) &\supset [\{y\} \cup B(\mathcal{O}')] \backslash [\mathcal{E}(\mathcal{O}') \cup \{x, y\}] \\ &= [\{y\} \cup B(\mathcal{O}')] \backslash [\mathcal{E}(\mathcal{O}') \cup \{y\}] \\ &= B(\mathcal{O}') \backslash \mathcal{E}(\mathcal{O}') \\ &\neq \emptyset, \end{aligned}$$

which contradicts $B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*) = \emptyset$.

Next suppose that there exists $B^* \subset B(\mathcal{O}') \backslash \mathcal{E}(\mathcal{O}')$ with B^* nonempty and $B^* \cap A(\mathcal{O}'_{B^*}) = \emptyset$, but $x \notin B^*$. Consider the subcollection $\mathcal{O}' \backslash \mathcal{O}'_{B^*}$ of $\mathcal{O}'$ which is not empty since B^* is nonempty. Notice that for any $(A, B, D) \in \mathcal{O}' \backslash \mathcal{O}'_{B^*}$, there exists $z \in B \cap B^*$. Since $x \notin B^*$, $z \neq x$. Since $B^* \subset B(\mathcal{O}') \backslash \mathcal{E}(\mathcal{O}')$, $z \notin \mathcal{E}(\mathcal{O}')$. Therefore, $z \notin \{x\} \cup \mathcal{E}(\mathcal{O}')$, which implies that $B \not\subseteq \{x\} \cup \mathcal{E}(\mathcal{O}')$. Hence, $\{(A', B', D') \in \mathcal{O}^* : B' \subset \{x\} \cup \mathcal{E}(\mathcal{O}')\} \subset \mathcal{O}'_{B^*} \cup \{(\{x, y\}, \{y\}, \{x\})\}$.

By definition of $\mathcal{E}^1$,

$$\mathcal{E}^1(\mathcal{O}^*, \{x\} \cup \mathcal{E}(\mathcal{O}')) = \left(\bigcup_{(A',B',D') \in \mathcal{O}^* : B' \subset \{x\} \cup \mathcal{E}(\mathcal{O}')} A' \right) \cup \{x\} \cup \mathcal{E}(\mathcal{O}').$$

Since $B^* \cap A(\mathcal{O}'_{B^*}) = \emptyset$, $z \notin A(\mathcal{O}'_{B^*})$. If $z \neq y$, then $z \notin \mathcal{E}^1(\mathcal{O}^*, \{x\} \cup \mathcal{E}(\mathcal{O}'))$. If $z = y$, then $z \notin \mathcal{E}(\mathcal{O}')$ implies that $y \notin \mathcal{E}(\mathcal{O}')$. Thus, $z = y \notin \mathcal{E}^1(\mathcal{O}^*, \{x\} \cup \mathcal{E}(\mathcal{O}'))$. Induction shows that $z \notin \mathcal{E}(\mathcal{O}^*, \{x\} \cup \mathcal{E}(\mathcal{O}')) = \mathcal{E}(\mathcal{O}^*)$. Therefore, $z \in B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*)$. This is a contradiction to the fact that $B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*) = \emptyset$. □

Similarly, to test whether $xP^s y$, we add the observation $(\{x, y\}, y, \emptyset)$ into the original coarse data set $\mathcal{O}$, and work with the new coarse data set $\mathcal{O}' = \mathcal{O} \cup \{(\{x, y\}, y, \emptyset)\}$ is not rationalizable. The following theorem provides a necessary and sufficient condition for $xP^s y$. The intuition is that when $xP^s y$, we can find a nonempty subcollection $\mathcal{O}'$ such that x is (one of) the best alternative in $A(\mathcal{O}')$ for all weak order that rationalize $\mathcal{O}$, and that $y \in \mathcal{E}(\mathcal{O}')$. Then whenever we stratify $A(\mathcal{O}')$, x must be in the higher ranked strata while y in the "dominated" strata.

Theorem 8. *$xP^s y$ if and only if there exists $\mathcal{O}' \subset \mathcal{O}$ such that $x \in A(\mathcal{O}')$, $y \in \mathcal{E}(\mathcal{O}')$, and for all $B^* \subset B(\mathcal{O}') \backslash \mathcal{E}(\mathcal{O}')$ with B^* nonempty, $B^* \cap A(\mathcal{O}'_{B^*}) = \emptyset$ implies $x \in B^*$.*

Proof of Theorem 8. (The if-part.) Consider any weak order R^* that rationalizes $\mathcal{O}$, R^* also rationalizes $\mathcal{O}'$. As in the proof of Theorem 7, we can show that $x \in max(A(\mathcal{O}'); R^*)$. Since $max(A(\mathcal{O}'); R^*) \cap \mathcal{E}(\mathcal{O}')$, we have $xP^* y$.

(The only if-part.) Suppose $xP^s y$. Then we can find $\mathcal{O}'$ such that if we add the observation $(\{x, y\}, \{y\}, \emptyset)$, the new constructed data $\mathcal{O}^* = \mathcal{O}' \cup \{(\{x, y\}, \{y\}, \emptyset)\}$ violates coarse WARP. It implies that $B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*) = \emptyset$.

Since $xP^s y$ implies $xR^s y$, it suffices to show that $x \in A(\mathcal{O}')$ and $y \in \mathcal{E}(\mathcal{O}')$. Suppose to the contrary that $y \notin \mathcal{E}(\mathcal{O}')$. Then $\mathcal{E}(\mathcal{O}^*) = \mathcal{E}(\mathcal{O}')$, which leads to $B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*) \supset B(\mathcal{O}') \backslash \mathcal{E}(\mathcal{O}') \neq \emptyset$, a contradiction. Now suppose $y \in \mathcal{E}(\mathcal{O}')$ and $x \notin A(\mathcal{O}')$. Then $\mathcal{E}(\mathcal{O}^*) = \{x\} \cup \mathcal{E}(\mathcal{O}')$, which leads to $B(\mathcal{O}') \backslash \mathcal{E}(\mathcal{O}') = B(\mathcal{O}^*) \backslash \mathcal{E}(\mathcal{O}^*) = \emptyset$, a contradiction. □

6 Conclusion

References

Afriat, S. N. (1967): "The Construction of Utility Functions from Expenditure Data," *International Economic Review*, 8, 67–77.

Brown, D. J. and R. L. Matzkin (1996): "Testable Restrictions on the Equilibrium Manifold," *Econometrica*, 1249–1262.

Carvajal, A., R. Deb, J. Fenske, and J. K.-H. Quah (2013): "Revealed Preference Tests of the Cournot Model," *Econometrica*, 81, 2351–2379.

de Clippel, G. and K. Rozen (2014): "Bounded Rationality and Limited Datasets," Mimeo. Brown University.

Dekel, E., B. L. Lipman, and A. Rustichini (2001): "Representing Preferences with a Unique Subjective State Space," *Econometrica*, 69, 891–934.

Fishburn, P. (1976): "Representable Choice Functions," *Econometrica*, 44, 1033–1043.

Kalai, G., A. Rubinstein, and R. Spiegler (2002): "Rationalizing choice functions by multiple rationales," *Econometrica*, 70, 2481–2488.

Kreps, D. M. (1979): "A Representation Theorem for 'Preference for Flexibility'," *Econometrica*, 47, 565–577.

Nishimura, H., E. A. Ok, and J. Quah (2015): "A Comprehensive Approach to Revealed Preference Theory," Working paper version.

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

Samuelson, P. (1938): "A Note on the Pure Theory of Consumer's Behaviour," *Economica*, 5, 61–71.

Savage, L. J. (1951): "The Theory of Statistical Decision," *Journal of the American Statistical Association*, 46, 55–67.

Sen, A. K. (1971): "Choice Functions and Revealed Preference," *Review of Economic Studies*, 38, 307–317.

Varian, H. R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50, 945–973.