IFN Working Paper No. 1321, 2020

A Rationalization of the Weak Axiom of Revealed Preference

Victor H. Aguiar, Per Hjertstrand and Roberto Serrano

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

A Rationalization of the Weak Axiom of Revealed Preference[*]

Victor H. Aguiar[†] Per Hjertstrand[‡] Roberto Serrano[§]

This version: February 2020

Abstract We offer a rationalization of the weak axiom of revealed preference (WARP) and of the weak generalized axiom of revealed preference (WGARP) for both finite and infinite data sets of consumer choice. We call it maximin rationalization, in which each pairwise choice is associated with a local utility function. We develop its associated revealed-preference theory. We show that preference recoverability and welfare analysis à la Varian (1982) may not be informative enough when the weak axiom holds but when consumers are not utility maximizers. In addition, we show that counterfactual analysis may fail with WGARP/WARP. We clarify the reasons for these failures and provide new informative bounds for the consumer's true preferences, as well as a new way to perform counterfactual analysis. Finally, by adding the Gorman form and quasilinearity restrictions to the "local" utility functions, we obtain new characterizations of the choices of the Shafer (1974) nontransitive consumer and those choices satisfying the law of demand.

JEL Classification: C60, D10.
Keywords: consumer choice; revealed preference; maximin rationalization; nonconvex preferences; reference-dependent utility; law of demand.

---

[*]This paper supersedes our previous draft titled "The Theory of Weak Revealed Preference." We are grateful to Roy Allen, Federico Echenique, Charles Gauthier, Reinhard John, John Rehbeck, Al Slivinsky and Gerelt Tserenjigmid for useful comments. We thank Judith Levi for her great editing work. Hjertstrand thanks Jan Wallander och Tom Hedelius stiftelse and Marianne och Marcus Wallenberg stiftelse for funding.

[†]Department of Economics, The University of Western Ontario, London, Ontario, Canada, vaguiar@uwo.ca.

[‡]Research Institute of Industrial Economics (IFN), Stockholm, Sweden, per.hjertstrand@ifn.se.

[§]Department of Economics, Brown University, Providence, Rhode Island, U.S.A., roberto_serrano@brown.edu.

1. Introduction

Rooted in the seminal work of Samuelson (1938), the weak axiom of revealed preference (WARP) and the weak generalized axiom of revealed preference (WGARP) have been seen as minimal, normatively appealing, and potentially empirically robust consistency conditions of choice. WGARP/WARP states that, for any pair of observations x^t and x^s, when a consumer chooses x^t with x^s being affordable, then, when she chooses x^s, it must be that x^t is more expensive/at least as expensive. Samuelson's original 1938 paper focuses on demand functions, and studies WARP. Following Varian (1982), it is empirically more convenient to work with demand correspondences, which, allowing for indifferences, provide a natural justification for WGARP.

Standard utility maximization requires, in addition to being consistent with WGARP, transitivity of preferences. However, there is abundant experimental and field evidence against this property of preferences (Tversky 1969; Quah 2006). The potential lack of robustness of the transitivity requirement on preferences motivated the influential work of Kihlstrom et al. (1976), which essentially proposed to rewrite the entire theory of demand on the basis of WARP alone. Also, the seminal work of Shafer (1974) proposed a nontransitive consumer who nevertheless satisfies WARP. More recently, practitioners have recognized some difficulties surrounding the computational complexity of using standard utility maximization in setups of empirical interest (e.g., stochastic utility maximization, which is NP-hard to check; see Kitamura and Stoye 2018). In response, there has been a renewed interest in using WGARP as a minimalist version of the standard model of rationality. (See, for example, Blundell et al. (2008), Hoderlein and Stoye (2014), Cosaert and Demuynck (2018), and Cherchye et al. (2019).)[1] In addition, many results in general equilibrium, consumer theory, and measurement rely on this condition (Quah 2008).

Nonetheless, our motivation for the present analysis is to address the important questions regarding the nonparametric approach to demand analysis under WGARP/WARP that have remained open, namely:

(i) Can a behavioral characterization of WGARP/WARP be provided without imposing additional restrictions?

(ii) Is WGARP/WARP suitable for counterfactual demand analysis or equivalently, can this condition alone make predictions out-of-sample?

(iii) How can preferences be recovered from observing behavior consistent with WGARP/WARP?

The main contribution of this paper is to provide complete answers for all these questions, based on the *maximin* model. These questions follow the classical approach

[1] In all of them, WGARP is usually stated without indifference, because the object of interest is a demand function, not a demand correspondence. We cover both here.

to demand analysis in Varian (1982). In addition to answering these questions, we also shed additional light on well-known paradigms– such as the Shafer (1974) nontransitive consumer and choices satisfying the law of demand–as special cases of the maximin model (i.e., incorporating shape constraints). Our results also unify the treatment of finite and infinite data sets.

Indeed, this paper proposes the maximin model as an answer to question (i). To motivate the model, suppose the consumer is trying to figure out her true preferences. Let us say she is trying to buy a car, and her three possible choices are a Ford, a Toyota, and a Subaru. There are two main attributes that she focuses on: gas consumption and reliability. When she compares the Ford to the Toyota, she focuses especially on the gas consumption attribute; in doing so, she obtains a complete ranking of all cars according to this feature. However, when she compares the Toyota to the Subaru, the attribute on which she focuses is the reliability of the car during extreme winter conditions; she then comes up with a possibly different ranking of all cars according to this dimension. Finally, when comparing the Ford to the Subaru, she gives weight to both gas consumption and reliability, and this produces yet a third ranking of all cars. We refer to each of these rankings as an expression of her "local" preferences. Could she express her "global" preferences over all cars by aggregating these "local" preferences?[2]

As explained next, the proposed notion brings to the forefront the idea of endogenous reference points (Kőszegi and Rabin 2006), relevant to each pairwise comparison. Indeed, the consumer acts as if any pairwise comparison colors her preferences over all possible choices she makes. (Our interpretation of reference points as attributes or state of moods is analogous to that in Richter and Rubinstein (2019).)

We find that a data set is consistent with WGARP if and only if it can be rationalized by a *maximin preference function* (Theorem 1); the notion for WARP is similar, simply switching from weak to strict rationalization (Theorem 2). We say that a finite data set $O^T = \{p^t, x^t\}_{t\in\mathbb{T}}$ (a collection of prices and commodity bundles) is weakly rationalized by a preference function $r : X \times X \to \mathbb{R}$, where X is the consumption set, if, for all t, $r(x^t, y) \geq 0$ for any $y \in X$ that is affordable at price p^t (and wealth $p^t x^t$).

We now define a maximin preference function. Let U be an arbitrary finite set, and $\Delta(U)$ the probability simplex on U. These abstract elements may be viewed as reference points. Let $u_{ij} : X \mapsto \mathbb{R}$ denote a reference-dependent utility function based on the (i, j) reference pair. Note that each utility function has a double subscript. This means that the utility may depend on either one $(i = j)$ or two $(i \neq j)$ reference points. We require that the order of such reference points not matter (i.e., $u_{ij} = u_{ji}$). We say that a data set is weakly rationalized by a maximin preference function r if, for any $x, y \in X$, we can

[2]Although this example is stated for a discrete choice set, it also works for commodity bundles that are the object of interest in this paper.

write $r(x, y)$ as:

$$r(x, y) = \max_{\mu \in \Delta(U)} \min_{\lambda \in \Delta(U)} \sum_{i \in U} \sum_{j \in U} \lambda_i \mu_j (u_{ij}(x) - u_{ij}(y)).$$

Thus, maximin rationalization can be interpreted as an aggregation of preferences of an individual with multiple utility functions that are heterogeneous, due to reference dependence. If this consumer is asked to make up her mind about how to compare any pair of consumption bundles, how would she aggregate her different preferences if her behavior is consistent with WGARP? The answer is provided by the maximin preference model. That is, we show that this consumer has a preference function that is the maximum over the minimal difference among the "local" utilities of the two bundles. Hence, to figure out her preference between x and y, this consumer is cautious, in that she first looks at the smallest differences between utilities (attribute by attribute), and only then maximizes among them. This maximin aggregation of local utilities extends a partial, reflexive, and asymmetric order (the direct revealed-preference relation under WGARP)[3] to a complete, reflexive, and asymmetric order on the grand-commodity set.

In fact, we show that a finite data set satisfying WGARP/WARP can be rationalized by a maximin preference function that is skew-symmetric- a key property of nontransitive consumers, first proposed by Shafer (1974) (i.e., $r(x, y) = -r(y, x)$ for all $x, y \in X$). The skew-symmetric maximin preference function dispenses with transitivity, while still being compatible with WGARP/WARP.

The maximin preference function $r(x, y)$ admits a game-theoretic interpretation. That is, when one focuses on the bilateral comparison between bundles x and y, the maximin $r(x, y)$ can be interpreted as the outcome of the interaction of two adversarial selves within the decision maker. In order to learn her own preferences, the consumer runs a contest between her two selves. In this zero-sum game, (i) in trying to find reasons for choosing x over y, one self chooses reference point i at random (mixed strategy μ) in order to maximize a payoff function equal to $u_{ij}(x) - u_{ij}(y)$; and (ii) in trying to find reasons to rank y over x, the other self chooses reference point j at random (mixed strategy λ), and thus has the negative of those payoffs. As in any two-player zero-sum game, the maximin, minimax, and equilibrium logic coincide. We have chosen to describe our representation as the maximin–reflecting cautious behavior–but we can also interpret it as a Nash equilibrium. This interpretation would be similar to the self-equilibrium notion in Kőszegi and Rabin (2006), because reference points are chosen endogenously from the set U.

Perhaps surprisingly, we have a negative answer to question (ii) above. We show that there are finite data sets that satisfy WGARP/WARP, and out-of-sample prices at which no demand bundle can be chosen without exhibiting a violation of the axiom. This negative

[3] A bundle x is revealed preferred to y whenever x is chosen with y being affordable.

result calls into question the usefulness of Samuelson's weak axiom for nonparametric demand analysis. We find the reason behind this failure and propose a way out. We show that the failure to obtain out-of-sample predictions from the skew-symmetric maximin preference function arises from the lack of convexity of preferences. This finding refutes a classical conjecture in Kihlstrom et al. (1976) about the empirical equivalence of WARP and the Shafer (1974) nontransitive consumer model. Indeed, Shafer (1974) proposed a preference function r that, in addition to being skew-symmetric (and therefore compatible with WARP), is also concave in the first argument, which ensures that the preferences of this consumer are convex and always produce a solution to the maximization problem. We propose the following two possible ways out of this pitfall involving WGARP/WARP in counterfactual analysis:

[a] Interpreting maximin rationalization as a model of preference formation, one can allow consumers to *drop* reference points out-of-sample. This reduction in the set of reference points can be interpreted as a way for the consumer to learn her preferences by prioritizing attributes. This means that the consumer will solve the maximin problem with a constrained set of reference points that is possibly smaller than the original set rationalizing the finite data set. We focus our attention on the minimal number of reference points that must be dropped in order to re-establish the existence of a solution to the maximization problem (Theorem 3).

[b] Alternatively, one can take a convex closure of the WGARP/WARP consistency conditions, thereby effectively imposing convexity on the consumer's preferences. In practice, this takes the form of assuming some shape constraints on the preference function. Such an approach was first proposed in John (2001), which also shows the equivalence of this new condition (i.e., convex closure of WGARP/WARP) and the model in Shafer (1974). Convexity of preferences allows us to reestablish the existence of the consumer maximization problem without transitivity, as shown in Shafer (1974). Thus, one could ask: what explains the gap in the maximin preference-function properties that makes WGARP/WARP counterfactual analysis ill-behaved? In Theorem 4 we find that the Shafer preference function is empirically equivalent to a restriction on the local utilities of the maximin model. That is, the local utilities must have parallel straight-line income expansion paths (i.e., they must admit a Gorman Form indirect-utility representation). This tells us that WGARP and the equivalent maximin model may fail to produce a solution to the consumer preference-maximization problem whenever local (reference-dependent) utilities are heterogeneous in their marginal utility of income, this finding resembles classical results from aggregation theory for a population of consumers.

Our finding that some data sets cannot be rationalized by convex preferences also provides a counterexample to Samuelson's *eternal darkness*, which refers to the impossibility of testing the convexity of preferences in the case of utility maximization. In addition, following Brown and Calsamiglia (2007) and Allen and Rehbeck (2018), we can craft our restrictions on the local utilities in the maximin preference function to provide a representation for choices obeying the *law of demand*. This is done through a quasilinear restriction in Theorem 5, a strengthening of the Gorman form. This characterization may be of interest in its own right, due to the importance of the law of demand in both theoretical and applied literatures.[4]

Let us turn now to the third central question we address in this paper, namely, question (iii) above. We find that attempting to recover preferences using the classical nonparametric tools developed in Varian (1982) may possibly result in uninformative bounds to preferences. The reason for this failure is that preferences that generate data sets consistent with WGARP may fail to be convex. Yet whenever possible, our model allows us to recover preferences and to do welfare analysis on the basis of WGARP-consistent data sets that cannot be generated by standard utility maximization. Specifically, in Theorem 6 we provide new informative bounds based on the notion of maximin rationalization for data sets consistent with WGARP. Our key innovation is to consider subsets of the data, consisting of pairs of observations, to which we can apply the tools in Varian (1982) to recover local preferences that are combined to get bounds on the true global preferences.

The plan of the paper is as follows. After the central notions of revealed-preference theory are reviewed in section 2, section 3 presents our characterizations of WGARP and WARP based on the maximin rationalization model. Section 4 addresses counterfactual analysis out-of-sample, and section 5 studies different shape constraints on the preference function, showcasing new applications of the maximin model. In section 6 we tackle the issue of recoverability of preferences. Section 7 extends the analysis to infinite data sets. Section 8 is a brief review of related literature, and section 9 concludes. Proofs are collected in an appendix.

2. Preliminaries

Suppose that a consumer chooses bundles consisting of $L \geq 2$ goods in a market. We assume that we have access to a *finite* number of observations, denoted by T, on the prices and chosen quantities of these goods, where observations are indexed by $\mathbb{T} = \{1, \ldots, T\}$. Let $x^t \in X \equiv \mathbb{R}^L_+ \setminus \{0\}$ denote the bundle of goods at time $t \in \mathbb{T}$, which was bought at prices $p^t \in P \equiv \mathbb{R}^L_{++}$. We impose Walras' law throughout: wealth at time t is equivalent to $p^t x^t \in W \equiv \mathbb{R}_{++}$, for all $t \in \mathbb{T}$.[5] We write $O^T = \{p^t, x^t\}_{t \in \mathbb{T}}$ to denote all price-quantity

[4]Specifically, demand functions satisfying the law of demand have downward-sloping demand curves, and allow the measurement of welfare changes in terms of consumer's surplus for a given change in market prices (Brown and Calsamiglia 2007).

[5]We use the following notation: The inner product of two vectors $x, y \in \mathbb{R}^L$ is defined as $xy = \sum_{l=1}^{L} x_l y_l$. For all $(x, y) \in \mathbb{R}^L$, $x \geqq y$ if $x_i \geq y_i$ for all $i = 1, \ldots, L$; $x \geq y$ if $x \geqq y$ and $x \neq y$; and $x > y$ if $x_i > y_i$ for all $i = 1, \ldots, L$. We denote $\mathbb{R}^L_+ = \{x \in \mathbb{R}^L : x \geqq (0, \ldots, 0)\}$ and $\mathbb{R}^L_{++} = \{x \in \mathbb{R}^L : x > (0, \ldots, 0)\}$.

observations, and refer to O^T as the data. In practice, the data O^T describe a single consumer that is observed over time.

2.1. Revealed-Preference Axioms

We begin by recalling some key definitions in the revealed-preference literature.

Definition 1. *(Direct revealed preferred relations) We say that x^t is directly revealed preferred to x^s, written $x^t \succeq^{R,D} x^s$, when $p^t x^t \geq p^t x^s$. Also, x^t is strictly and directly revealed preferred to x^s, written $x^t \succ^{R,D} x^s$, when $p^t x^t > p^t x^s$.*

If x^t is directly revealed preferred to x^s, this means that the consumer chose x^t and not x^s, when both bundles were affordable. If x^t is strictly and directly revealed preferred to x^s, then she could also have saved money by choosing x^s. These definitions only compare pairs of bundles. We can extend them to compare any subset of bundles by using the transitive closure of the direct relation:

Definition 2. *(Revealed preferred relations) We say that x^t is revealed preferred to x^s, written $x^t \succeq^R x^s$, when there is a chain $(x^1, x^2, \ldots, x^n) \in X$ with $x^1 = x^t$ and $x^n = x^s$ such that $x^1 \succeq^{R,D} x^2 \succeq^{R,D} \ldots \succeq^{R,D} x^n$. Also, x^t is strictly revealed preferred to x^s, written $x^t \succ^R x^s$, when at least one of the directly revealed relations in the revealed preferred chain is strict.*

Hence, the revealed preferred relation $\succeq^R$ is the transitive closure of the direct revealed-preference relation $\succeq^{R,D}$. Next, we use these binary relations to define axioms that characterize different types of rational consumer behavior. We begin with Samuelson's (1938) weak axiom of revealed preference:

Axiom 1. *(WARP) The weak axiom of revealed preference (WARP) holds if there is no pair of observations $s, t \in \mathbb{T}$ such that $x^t \succeq^{R,D} x^s$, and $x^s \succeq^{R,D} x^t$, with $x^t \neq x^s$.*

Kihlstrom et al. (1976) introduces a generalized version of WARP:[6]

Axiom 2. *(WGARP) The weak generalized axiom of revealed preference (WGARP) holds if there is no pair of observations $s, t \in \mathbb{T}$ such that $x^t \succeq^{R,D} x^s$, and $x^s \succ^{R,D} x^t$.*

Samuelson (1948) shows how WARP can be used to construct a set of indifference curves in the two-dimensional ($L = 2$) case, but also recognizes that WARP is not enough to characterize rationality in the multidimensional ($L > 2$) case. Responding to this

[6] In contrast to only observing a finite number of prices and quantities, suppose that we knew the entire demand function. In this case, Kihlstrom et al. (1976) shows that if the demand function is differentiable and satisfies WGARP at every point in its domain, then the Slutsky substitution matrix derived from the demand function is negative semidefinite at every point.

challenge, Houthakker (1950) introduces the strong axiom of revealed preference (SARP), which makes use of transitive comparisons between bundles as implied by the revealed preferred relation:

Axiom 3. *(SARP) The strong axiom of revealed preference (SARP) holds if there is no pair of observations $s, t \in \mathbb{T}$ such that $x^t \succeq^R x^s$, and $x^s \succeq^{R,D} x^t$, with $x^t \neq x^s$.*

Varian (1982) notes that SARP requires single-valued demand functions, and argues that it is empirically more convenient to work with demand correspondences and "flat" indifference curves. To accomodate these properties, Varian introduces the generalized axiom of revealed preference (GARP):

Axiom 4. *(GARP) The generalized axiom of revealed preference (GARP) holds if there is no pair of observations $s, t \in \mathbb{T}$ such that $x^t \succeq^R x^s$ and $x^s \succ^{R,D} x^t$.*

In the two-dimensional case, the following equivalences are known:

Theorem A. (Equivalence of axioms) Let $L = 2$. Consider a finite data set $O^T = \{p^t, x^t\}_{t \in \mathbb{T}}$:

- The data O^T satisfies SARP if and only if O^T satisfies WARP (Rose 1958).
- The data O^T satisfies GARP if and only if O^T satisfies WGARP (Banerjee and Murphy 2006).

2.2. Revealed-Preference Characterizations

In this section, we recall the main results from the revealed-preference literature that are needed in order to introduce our contribution. Consider the following definitions of rationalization:[7]

Definition 3. *(Utility rationalization) Consider a finite data set $O^T = \{p^t, x^t\}_{t \in \mathbb{T}}$ and a utility function $u : X \mapsto \mathbb{R}$. For all $x \in X$ and all $t \in \mathbb{T}$ such that $p^t x \leq p^t x^t$,*

- *the data O^T is weakly rationalized by u if $u(x^t) \geq u(x)$.*
- *the data O^T is strictly rationalized by u if $u(x^t) > u(x)$ whenever $x \neq x^t$.*

[7] We say that a utility function $u : X \mapsto \mathbb{R}$ is: *(i) continuous* if for any sequence (x^n) for $n \in \mathbb{N}_+$ such that $x^n \in X$ and $\lim_{n \to \infty} x^n = x$ with $x \in X$ implies $\lim_{n \to \infty} u(x^n) = u(x)$; *(ii) locally nonsatiated* if for any $x \in X$ and for any $\epsilon > 0$, there exists $y \in B(x, \epsilon)$ where $B(x, \epsilon) = \{z \in X | \quad ||z - x|| \leq \epsilon\}$ such that $u(y) > u(x)$; *(iii) strictly increasing* if for $x, y \in X$, $x \geq y$ implies $u(x) > u(y)$; and *(iv) concave* if for any $x, y \in X$, we have $u(x) - u(y) \geq \xi\,(y - x)$, for $\xi \in \partial u(y)$, where $\partial u(y)$ is the subdifferential of u.

Afriat's (1967) fundamental theorem is well known:

Theorem B. (Afriat's theorem, Varian 1982) Consider a finite data set $O^T = \{p^t, x^t\}_{t\in\mathbb{T}}$. The following statements are equivalent:

(i) The data O^T can be weakly rationalized by a locally nonsatiated utility function.

(ii) The data O^T satisfies GARP.

(iii) There exist numbers U^t and $\lambda^t > 0$ for all $t \in \mathbb{T}$ such that the Afriat inequalities:

$$U^t - U^s \geq \lambda^t p^t(x^t - x^s),$$

hold for all $s, t \in \mathbb{T}$.

(iv) There exist numbers V^t for all $t \in \mathbb{T}$ such that the Varian inequalities:

$$\begin{aligned} \text{if } p^t(x^t - x^s) &\geq 0 \text{ then, } V^t - V^s \geq 0, \\ \text{if } p^t(x^t - x^s) &> 0 \text{ then, } V^t - V^s > 0, \end{aligned}$$

hold for all $s, t \in \mathbb{T}$.

(v) The data O^T can be weakly rationalized by a continuous, strictly increasing, and concave utility function.

There are several interesting features of Afriat's theorem.[8] Statements (ii), (iii), and (iv) give testable conditions that are easy to implement in practice. Perhaps the most interesting theoretical implication of Afriat's theorem is that statements (i) and (v) are equivalent, which means that continuity, monotonicity, and concavity are nontestable properties. In other words, separate violations of any of these properties cannot be detected in finite data sets.

Varian (1982) shows that the numbers U^t and λ^t in statement (iii) can be interpreted as measures of the utility level and marginal utility level of income at observation $t \in \mathbb{T}$. Analogously, Demuynck and Hjertstrand (2019) shows that the numbers V^t in statement (iv) can be interpreted as measures of the utility levels at the observed demands.

Matzkin and Richter (1991) provides an analogous result for strict rationalization, by showing that SARP is a necessary and sufficient condition for a data set O^T to be strictly rationalized by a continuous, strictly increasing, and strictly concave utility function:

Theorem C. (Matzkin and Richter 1991) Consider a finite data set $O^T = \{p^t, x^t\}_{t\in\mathbb{T}}$. The following statements are equivalent:

[8] Statements (i), (ii), (iii), and (v) comprise Varian's (1982) original formulation of Afriat's theorem. Statement (iv) is rather new to the revealed preference literature.

(i) The data O^T can be strictly rationalized by a locally nonsatiated utility function.

(ii) The data O^T satisfies SARP.

(iii) There exist numbers U^t and $\lambda^t > 0$ for all $t \in \mathbb{T}$ such that the inequalities:

$$\begin{aligned} \text{if } x^t &\neq x^s \text{ then, } U^t - U^s > \lambda^t p^t \left(x^t - x^s\right), \\ \text{if } x^t &= x^s \text{ then, } U^t - U^s = 0, \end{aligned}$$

hold for all $s, t \in \mathbb{T}$.

(iv) There exist numbers V^t for all $t \in \mathbb{T}$ such that the inequalities:

$$\begin{aligned} \text{if } x^t &\neq x^s \text{ and } p^t(x^t - x^s) \geq 0 \text{ then, } V^t - V^s > 0, \\ \text{if } x^t &= x^s \text{ then, } V^t - V^s = 0, \end{aligned}$$

hold for all $s, t \in \mathbb{T}$.

(v) The data O^T can be strictly rationalized by a continuous, strictly increasing and strictly concave utility function.

Matzkin and Richter's (1991) theorem mirrors Afriat's, in the sense that it shows that continuity, monotonicity, and strict concavity are nontestable properties.[9] Although much is known about the types of consumer behavior that characterize finite data sets satisfying SARP and GARP, there are no analogous characterizations for WARP and WGARP. The current paper, starting with the next section, fills this gap.

3. Characterizations of WGARP and WARP

In this section, we provide revealed-preference characterizations analogous to the ones in Afriat's and Matzkin and Richter's theorems, for WGARP and WARP. We begin by introducing the maximin preference model, which, to the best of our knowledge, is a new model of consumer behavior.

3.1. The Maximin Preference Model

We start with some preliminaries.

[9] Matzkin and Richter's (1991) original formulation consists of statements (i), (ii), (iii), and (v). Talla Nobibon et al. (2016) proves the equivalence of statements (ii) and (iv).

Definition 4. *(Preference function) A preference function is a mapping $r : X \times X \to \mathbb{R}$, that maps ordered pairs of commodity bundles to real numbers.*

A preference function is a numerical representation of a consumer's preferences. If $r(x, y) \geq 0$ then the consumer prefers bundle x to y. Similarly, if $r(x, y) > 0$, x is strictly preferred to y.

In what follows, we focus on a particular representation of r, namely the maximin model. Let U be an arbitrary finite set. The elements in this set can be thought as reference points. Let a *reference pair* be defined as any pair of points $i, j \in U$. For every reference pair $i, j \in U$ let $u_{ij}(x) : X \to \mathbb{R}$ denote a *reference dependent utility function*. Since every reference dependent utility function depends on a pair of elements in U, there are $|U| \times |U|$ such utility functions, where $|\cdot|$ denotes cardinality. We assume that u_{ij} is independent of permutations of the indices in the subscript, so that $u_{ij} = u_{ji}$ for all $i, j \in U$.

It is natural to think of the reference points as representing attributes, or state of moods of the consumer. As such, the utility u_{ij} represents the preference of the consumer over consumption bundles when considering attributes $i, j \in U$. To give a different example from the one in the introduction, suppose a consumer has preferences consisting of $|U| = 2$ distinct attributes, *healthy* (h) and *tasty* (t), over bundles of hamburger and salad. In this case, she evaluates the bundle differently from a pure healthy perspective than she would from a pure tasty perspective, which also differ if she combines the two attributes and evaluates the bundle from a "healthy and tasty" perspective. The reference-dependent utilities u_{hh}, u_{tt}, and $u_{ht} = u_{th}$ represent the utilities from the three different states. How should she aggregate these moods in order to formulate her global preference?

Let $\Delta(U)$ be the probability simplex defined on U. Our maximin preference model is defined as:

Definition 5. *(Maximin (strict) preference model) We say that the preference function $r(x, y)$ is a maximin (strict) preference function if, for any $x, y \in X$, it can be written as:*

$$r(x, y) = \max_{\mu \in \Delta(U)} \min_{\lambda \in \Delta(U)} \sum_{i \in U} \sum_{j \in U} \lambda_i \mu_j (u_{ij}(x) - u_{ij}(y)),$$

where, for any reference point indexed by $i, j \in U$, the local utility function, u_{ij}, is continuous, strictly increasing, and (strictly) concave.

The maximin preference function assigns a numerical value to the comparison of any pair of commodity bundles $x, y \in X$, by additively aggregating over local preferences that are defined for any reference pair. More specifically, the aggregation is a maximin function, which in the first dimension takes the maximal difference between the utility gains of x over y, and in the second dimension, takes the minimal value of that difference, over the different local utility functions. As will be argued, the maximin aggregation effectively extends the incomplete direct revealed preference relation when it is asymmetric (i.e., when WGARP holds) to the commodity space X. As such, this representation provides us with an interpretation of the direct revealed preference relation when rationality –GARP– fails, but WGARP is satisfied.

For any reference pair $i, j \in U$, the local utility function, u_{ij}, is continuous, strictly increasing, and (strictly) concave. Moreover, the maximin is attained at a particular average utility (with endogenous weights) , i.e.,

$$r(x,y) = \sum_{i \in U} \sum_{j \in U} \lambda_i^* \mu_j^* (u_{ij}(x) - u_{ij}(y)),$$

for some $\lambda^*, \mu^* \in \Delta(U)$. Thus, for any pairwise comparison, the model is "locally" (i.e., for a fixed pair of bundles) equivalent to one where the consumer behaves as if she is rational. However, note from the definition of the maximin model that, for any distinct pairwise comparison preferences may change, in which case, preferences are not stable across all observations. In the maximin (strict) preference model, a consumer still behaves locally "as if she were rational" in that behavior according to this model rules out binary inconsistencies. More precisely, we show that a data set O^T satisfies (WARP) WGARP if and only if the maximin (strict) preference model (strictly) rationalizes the data.

Next, we present some properties of the preference function. We begin with a property that turns out to be key in our characterizations of WARP and WGARP:

Definition 6. *(Skew-symmetry) We say that a preference function $r : X \times X \mapsto \mathbb{R}$ is* skew-symmetric *if $r(x,y) = -r(y,x)$ for all $x, y \in X$.*

Skew-symmetry means that the preference function r induces a preference order on X that is complete and asymmetric. Note that, when a data set O^T satisfies WGARP, the direct preference relation is an (incomplete) asymmetric relation on X, the preference function r extends it if it is a maximin function. We have the following result:

Lemma 1. *If r is a maximin (strict) preference function, then for any $x, y \in X$, we have:*

$$\begin{aligned} r(x,y) &= \max_{\mu \in \Delta(U)} \min_{\lambda \in \Delta(U)} \sum_{i \in U} \sum_{j \in U} \lambda_i \mu_j (u_{ij}(x) - u_{ij}(y)) \\ &= \min_{\lambda \in \Delta(U)} \max_{\mu \in \Delta(U)} \sum_{i \in U} \sum_{j \in U} \lambda_i \mu_j (u_{ij}(x) - u_{ij}(y)), \end{aligned}$$

and moreover, r is skew-symmetric.

The proof of this lemma is omitted, as it follows directly from the classical von Neumann's minimax theorem (because Δ is convex and compact, and the sum is linear in λ and μ). It is easy to see that the maximin (strict) preference function is skew-symmetric, making this model a generalization of the general nontransitive consumer model, considered in Shafer (1974).

The following examples illustrate the maximin model:

Example 1. *Consider $L = 2$ goods, hamburger (x_1) and salad (x_2). As above, suppose that the preferences of the consumer consists of two attributes, healthy (h) and tasty (t). Let the consumer be endowed with the following three reference-dependent utilities: $u_{tt}(x) = \overline{\alpha}x_1 + x_2$, $u_{hh} = \underline{\alpha}x_1 + x_2$, and $u_{th}(x) = u_{ht}(x) = x_1 + x_2$, where $\overline{\alpha} > 1 > \underline{\alpha}$. The maximin preference function has a closed form solution:*

$$r(x, y) = u_{ht}(x) - u_{ht}(y) = x_1 + x_2 - y_1 - y_2.$$

This implies that $r(x, y) = -r(y, x)$.

In this example, behavior is not only consistent with WGARP but also rational (consistent with GARP by Theorem A) because the consumer endogenously decides to focus on both attributes to make her decision independently of the quantities consumed of both goods.

The next example illustrates a more complicated solution with reference dependence, and describes how reference dependence arises endogenously from the maximin aggregation of preferences.

Example 2. *Consider $L = 3$ goods, such that x_1, x_2, x_3 represent consumption of vegetables, chocolate, and meat. Suppose there are two reference points, $U = \{h, s\}$, representing two different moods, where h stands for hedonistic, and s for stoic. The utility of the hedonistic mood is $u_{hh}(x) = \frac{1}{2}log(x_1) + log(x_2) + log(x_3)$, and the utility of the stoic mood is $u_{ss}(x) = 2log(x_1) + log(x_2) + log(x_3)$. The combination of the two moods has utility $u_{hs}(x) = u_{sh}(x) = log(x_1) + log(x_2) + log(x_3)$. When the consumer is on the stoic mood she has a boost for eating vegetables, when she is on the hedonistic mood she experiences a lessening of the utility of eating vegetables. The maximin preference function has a closed form solution:*

$$r(x, y) = \begin{cases} u_{hh}(x) - u_{hh}(y) & x_1 \leq y_1 \\ u_{ss}(x) - u_{ss}(y) & x_1 > y_1. \end{cases}$$

Note the preference function r above has an endogenous reference point at $x_1 = y_1$. When comparing bundle x to y, the maximin consumer valuation depends on whether the former bundle has weakly more quantity of vegetables than the latter. If a bundle has more vegetables than another bundle, then she evaluates the bundles according to the the stoic mood. Otherwise the evaluation will correspond to the hedonistic mood. The higher presence of vegetables activates the stoic mood in the consumer. Following Shafer (1974) we can maximize the preference function r subject to the budget constraint $p_1x_1 + p_2x_2 + p_3x_3 = w$, and obtain the demand system:

$$x_1(p, w) = \frac{p_2 w}{p_1(3p_2 + 2p_3)},$$

$$x_2(p,w) = \frac{2w}{3p_2 + 2p_3},$$
$$x_3(p,w) = \frac{2p_2 w}{p_3(3p_2 + 2p_3)}.$$

The demand system does not depend directly on the reference pair. Nonetheless, we can easily verify that the entries of the associated Slutsky matrix[10] *of this system of equations is not symmetric:*

$$s_{13}(p,w) = \frac{2p_2(p_2 - p_3)w}{p_1 p_3(3p_2 + 2p_3)^2},$$
$$s_{31}(p,w) = \frac{2p_2^2 w}{p_1 p_3(3p_2 + 2p_3)^2}.$$

In this example, reference dependence breaks transitivity. This means that this demand system is not rational or violates GARP/SARP. However, it satisfies WGARP.

Our last example connects our work with the recent maximin model of rationality in Frick et al. (2019).

Example 3. *Consider a case where there is one physical good (money) and* 3 *states of the world (good, business-as-usual, bad). The Arrow-Debreu state-contingent securities are* x_l *for* $l = 1, 2, 3$*. Suppose that the consumer knows that state* 1 *happens with probability* $1/2$*, and state* 2 *or* 3 *happens with probability* $1/2$ *but there is uncertainty about which of these two will occur. The consumer has two reference points* $U = \{r, m\}$*, where* r *stands for realistic and* m *for pessimistic. The utility of an Arrow-Debreu security* x *associated with the attitudes* i, j *is:*

$$u_{ij} = 1/2v(x_1) + (1/2 - \pi_{ij})v(x_2) + \pi_{ij}v(x_3) = E_{\pi_{i,j}}[v(x)],$$

where $0 \leq \pi_{i,j} \leq 1/2$ *and* v *is a Bernoulli utility defined over money. The subjective probability associated with* $r, r \in U$ *is* $\pi_{r,r} = 1/5$*, for* $m, m \in U$ *it is* $\pi_{m,m} = 1/3$*, and for* $r, m \in U$ *it is* $\pi_{r,m} = \pi_{m,r} = 1/4$*. We note that this model is similar to the ambiguity framework posed in Frick et al. (2019), where the "act" corresponds to choosing bundle* x *over bundle* y*, such that the utility of such "act" is given by:*

$$r(x,y) = \max_{i \in \{r,m\}} \min_{j \in \{r,m\}} (E_{\pi_{i,j}}(v(x) - v(y))).$$

Here we assume that the maximin is achieved at pure strategies. This is done for illustrative purposes and without loss of generality as v *is left unspecified.*

The next definition lists some additional important properties of preference functions, which will connect with the maximin preference model in the sequel:

[10] The Slutsky matrix is an $L \times L$ matrix with entry $s_{lk}(p,w) = \frac{\partial x_l(p,w)}{\partial p_k} + \frac{\partial x_l(p,w)}{\partial w} x_k(p,w)$.

Definition 7. *Consider a preference function $r : X \times X \to \mathbb{R}$. We say that:*

(i) r *is* continuous *if for all $y \in X$ and any sequence $\{x^n\}$ of elements in X that converges to $x \in X$ it must be that $\lim_{n\to\infty} r(x^n, y) = r(x, y)$.*[11]

(ii) r *is* locally nonsatiated *if for any $x, y \in X$ such that $r(x, y) = 0$ and for any $\epsilon > 0$, there exists a $y' \in B(y, \epsilon)$ such that $r(x, y') < 0$.*

(iii) r *is* strictly increasing *if for all $x, y, z \in X$, $x \geq z$ implies $r(x, y) > r(z, y)$.*

(iv) r *is* quasiconcave *if for all $x, y, z \in X$ and any $0 \leq \lambda \leq 1$ we have $r(\lambda x + (1 - \lambda) z, y) \geq \min\{r(x, y), r(z, y)\}$ and* strictly quasiconcave *if, for any $0 < \lambda < 1$, the inequality is strict whenever $x \neq z$.*

(v) r *is* concave *if for all $x, y, z \in X$ and any $0 \leq \lambda \leq 1$ we have $r(\lambda x + (1 - \lambda) z, y) \geq \lambda r(x, y) + (1 - \lambda) r(z, y)$, and* strictly concave *if, for any $0 < \lambda < 1$, the inequality is strict whenever $x \neq z$.*

(vi) r *is* piecewise concave *if there is a sequence of concave functions in the first argument $f_t(x, y)$ for $t \in \mathbb{K}$, where $\mathbb{K}$ is a compact set, such that $r(x, y) = \max_{t\in\mathbb{K}} \{f_t(x, y)\}$, and* strictly piecewise concave *if there is a similar sequence of strictly concave functions.*

Continuity is a technical condition that is convenient to ensure existence of a maximum in the constrained maximization of the preference function (Sonnenschein 1971). Local nonsatiation rules out thick indifference curves: if we take an arbitrarily small neighborhood of a bundle that is indifferent to a given bundle x, the neighborhood contains bundles that are dominated by x. Strict monotonicity simply means that "more is better". Quasiconcavity says that for any fixed point $y \in X$, a mixture of two bundles $x, z \in X$ is at least as good as the worst of the two bundles, according to the preference function. Concavity is a cardinal version of quasiconcavity. Quasiconcavity and concavity are important properties because they ensure well-behaved optimization problems. More precisely, quasiconcavity guarantees that a function defined on a compact set has a convex set of maxima points, while a strictly concave function defined on a compact set always has a unique global maximum. In the general nontransitive consumer model considered by Shafer (1974), the preference function is assumed to be concave.

Piecewise concavity and its strict version are new properties, which turn out to be especially important for our characterizations of WGARP/WARP.[12] The property says that for a fixed $y \in X$, a mixture of two bundles $x, z \in X$ is at least as good as the worst one of the two bundles, but only if x, z are close enough. In other words, this is a local version of concavity, implying local quasiconcavity.

[11] We state the weaker versions of these properties, as all we need is to work with movements in one of the arguments.

[12] See Zangwill (1967) for a detailed discussion of piecewise concave functions.

3.2. Preference Function Rationalization

We now introduce the notion of (strict) rationalization by a preference function, which is analogous to utility rationalization in Definition 3:

Definition 8. *(Preference function rationalization) Consider a finite data set* $O^T = \{p^t, x^t\}_{t\in\mathbb{T}}$ *and a preference function* $r : X \times X \mapsto \mathbb{R}$*. For all* $y \in X$ *and all* $t \in \mathbb{T}$ *such that* $p^t y \leq p^t x^t$*,*

- *the data* O^T *is weakly rationalized by* r *if* $r(x^t, y) \geq 0$*.*
- *the data* O^T *is strictly rationalized by* r *if* $r(x^t, y) > 0$ *whenever* $y \neq x^t$*.*

3.3. WGARP

The next theorem provides a revealed-preference characterization of WGARP for finite data sets. This result mirrors Afriat's theorem in terms of preference-function rationalization (as opposed to utility rationalization):

Theorem 1. *Consider a finite data set* $O^T = \{p^t, x^t\}_{t\in\mathbb{T}}$*. The following statements are equivalent:*

(i) The data O^T *can be weakly rationalized by a locally nonsatiated and skew-symmetric preference function.*

(ii) The data O^T *satisfies WGARP.*

(iii) There exist numbers $R^{t,s}$ *and* $\lambda^t_{ts} > 0$ *for all* $s, t \in \mathbb{T}$ *with* $R^{t,s} = -R^{s,t}$ *and* $\lambda^t_{ts} = \lambda^t_{st}$ *such that inequalities:*

$$R^{t,s} \geq \lambda^t_{ts} p^t (x^t - x^s),$$

hold for all $s, t \in \mathbb{T}$*.*

(iv) There exist numbers $W^{t,s}$ *for all* $s, t \in \mathbb{T}$ *with* $W^{t,s} = -W^{s,t}$ *such that inequalities:*

$$\text{if } p^t(x^t - x^s) \geq 0 \text{ then, } W^{t,s} \geq 0,$$
$$\text{if } p^t(x^t - x^s) > 0 \text{ then, } W^{t,s} > 0,$$

hold for all $s, t \in \mathbb{T}$*.*

(v) The data O^T *can be weakly rationalized by a maximin preference function.*

(vi) The data O^T *can be weakly rationalized by a continuous, strictly increasing, piecewise concave, and skew-symmetric preference function.*

The equivalence of statements (i) and (vi) shows that, if the data can be weakly rationalized by any nontrivial preference function at all, it can, in fact, be weakly rationalized by a preference function that satisfies continuity, monotonicity, and piecewise concavity. Put differently, separate violations of these three properties cannot be detected in finite data sets.

The numbers $R^{t,s}$ and λ^t_{st} in statement (iii) have a similar interpretation as in Afriat's theorem for each reference point; that is, if we consider $t, s \in \mathbb{T}$, then $R^{t,s}$ is a measure of the utility difference $u_{ts}(x^t) - u_{ts}(x^s)$ for that particular pairwise data set, while λ^t_{st} is a measure of the marginal utility level of income at observation $t \in \mathbb{T}$ in that pairwise data set.

3.4. WARP

The next theorem provides a revealed-preference characterization of WARP for finite data sets, and mirrors Matzkin and Richter's (1991) theorem:

Theorem 2. *Consider a finite data set $O^T = \{p^t, x^t\}_{t\in\mathbb{T}}$. The following statements are equivalent:*

(i) The data O^T can be strictly rationalized by a locally nonsatiated and skew-symmetric preference function.

(ii) The data O^T satisfies WARP.

(iii) There exist numbers $R^{t,s}$ and $\lambda^t_{ts} > 0$ for all $s, t \in \mathbb{T}$ with $R^{t,s} = -R^{s,t}$ and $\lambda^t_{ts} = \lambda^t_{st}$ such that inequalities:

$$\begin{aligned} \text{if } x^t &\neq x^s \text{ then, } R^{t,s} > \lambda^t_{ts} p^t \left(x^t - x^s\right), \\ \text{if } x^t &= x^s \text{ then, } R^{t,s} = 0, \end{aligned}$$

hold for all $s, t \in \mathbb{T}$.

(iv) There exist numbers $W^{t,s}$ for all $s, t \in \mathbb{T}$ with $W^{t,s} = -W^{s,t}$ such that inequalities:

$$\begin{aligned} \text{if } x^t &\neq x^s \text{ and } p^t(x^t - x^s) \geq 0 \text{ then, } W^{t,s} > 0, \\ \text{if } x^t &= x^s \text{ then, } W^{t,s} = 0, \end{aligned}$$

hold for all $s, t \in \mathbb{T}$.

(v) The data O^T can be strictly rationalized by a maximin strict preference function.

(vi) The data O^T can be strictly rationalized by a continuous, strictly increasing, piecewise strictly concave, and skew-symmetric preference function.

Analogous to Theorem 1, this result shows that separate violations of continuity, monotonicity, and strict piecewise concavity cannot be detected in finite data sets.

4. Demand Counterfactuals

In this section, we investigate the important open question of whether WGARP is suitable for counterfactual demand analysis. That is, for a new (possibly unobserved) price vector p^{T+1}, we study whether WGARP is able to predict demand generically. As such, we need to formalize what it means for WGARP to make out-of-sample predictions. We are interested in the following object:

Definition 9. *(W-Demand Set) We define the W-demand set, or the set of all bundles compatible with WGARP, by*

$$D^{\downarrow}(p^{T+1}, w^{T+1}) = \{x \in X : O^T \cup \{p^{T+1}, x\} \text{ satisfies WGARP and } p^{T+1}x = w^{T+1}\}.$$

We obtain a negative result that seems to be new in the literature. In particular, the following lemma shows that, for some prices, the W-demand set may be empty, which implies that, for these prices, it is not possible to predict demand using WGARP.

Lemma 2. *(Impossibility) There are data sets O^T such that for an open set of out-of-sample prices $p^{T+1} \in P$, the W-demand set is empty, i.e., $D^{\downarrow}(p^{T+1}, w^{T+1}) \equiv \emptyset$.*

We notice that the W-demand set for a new price-vector can be formulated as a linear program by making use of the following result:

Corollary 1. *The bundle $x^{T+1} \in X$ is in $D^{\downarrow}(p^{T+1}, w^{T+1})$ if and only if it satisfies:*

(i) $p^{T+1}x^{T+1} = w^{T+1}$,

(ii) $p^t x^{T+1} \geq p^t x^t$, *for all* $t \in \mathbb{T}$, *for which* $p^{T+1}x^t \leq p^{T+1}x^{T+1}$,

(iii) $p^t x^{T+1} > p^t x^t$, *for all* $t \in \mathbb{T}$, *for which* $p^{T+1}x^t < p^{T+1}x^{T+1}$.

This program is a simplification of the procedure for the same purpose under GARP, proposed by Varian (1982). The first condition imposes Walras' law for the target income level. The second condition imposes the restriction that, if the observed bundles are cheaper than the new bundle at the new prices, then the new bundle cannot be affordable at the observed prices. The third condition strengthens the second, for the case of a strict inequality.

We now illustrate the foregoing lemma by means of a counterexample for a single price. The fact that there exists an open set of prices for which the W-demand set is empty shows that our counterexample does not constitute a degenerate case, but that it is also robust to perturbations of the out-of-sample price.

Example 4. *(Empty-demand counterfactuals) (Keiding and Tvede 2013, Example 1, p.467) Consider the data set O^3 with prices $p^1 = (4,1,5)'$, $p^2 = (5,4,1)'$, $p^3 = (1,5,4)'$, and bundles $x^1 = (4,1,1)'$, $x^2 = (1,4,1)'$, $x^3 = (1,1,4)'$. Note that the income level in all observations is the same, i.e., $p^t x^t = 22$ for all $t \in \{1,2,3\}$. This data set satisfies WGARP. Suppose the out-of-sample budget is: $p^{T+1} = \frac{1}{3}(p^1 + p^2 + p^3) = \frac{10}{3}(1,1,1)'$ and $w^{T+1} = 22$. Now, assume towards a contradiction that there exists a bundle x^{T+1} in the set $D^{\downarrow}(p^{T+1}, w^{T+1})$. Note that x^{T+1} is directly revealed preferred to x^t, because $22 = p^{T+1}x^{T+1} > p^{T+1}x^t = 20$ for all $t \in \{1,2,3\}$. By definition, it must be that $p^t x^t < p^t x^{T+1}$ for all $t \in \{1,2,3\}$, such that WGARP (and WARP) holds. However, averaging these inequalities, we get $22 = \frac{1}{3}(p^1x^1 + p^2x^2 + p^3x^3) < p^{T+1}x^{T+1} = 22$, where the right-hand side of the inequality follows from the definition of p^{T+1}. Hence, we obtain a contradiction, and can conclude that $D^{\downarrow}(p^{T+1}, w^{T+1}) = \emptyset$.*

4.1. A New Approach to Counterfactual Analysis Using WGARP

Maximin rationalization allows to understand the source of the failure of WGARP in producing nonempty demand counterfactuals. Simply put, the multiple-selves involved in the maximin may not be able to produce aggregated preferences that provide a solution for the consumer's maximization problem. As already argued, one key interpretation of maximin rationalization is to see it as a model of preference formation, and emptiness of the demand correspondence can be viewed as a failure of this preference formation process. As a solution for this failure, in this section we propose to allow the maximin rational consumer to drop reference points when such a solution does not exist. Recall that the set U is the set of all reference points.

Definition 10. *(Decisive maximin (strict) preference model) We say that the preference function $r(x,y)$ is a decisive maximin (strict) preference function if, for any $x, y \in X$, it can be written as:*

$$r(x,y) = \max_{\mu \in \Delta(V_{p,w})} \min_{\lambda \in \Delta(V_{p,w})} \sum_{i \in V_{p,w}} \sum_{j \in V_{p,w}} \lambda_i \mu_j (u_{ij}(x) - u_{ij}(y)),$$

where, $V_{p,w}$ is a subset of U, such that for any reference point indexed by $i, j \in U$, the local utility function, u_{ij}, is continuous, strictly increasing, and (strictly) concave; and such that maximizing r for any arbitrary linear budget set defined by prices $p \in P$ and wealth $w \in W$ has a solution.

The decisive maximin preference model is well defined because there is always a subset $V_{p,w} \subseteq U$ that satisfies the definition, namely, we can set $V_{p,w}$ to be a singleton. In that case, continuity of the only utility function associated with $V_{p,w}$ guarantees the existence of a solution to the preference maximization problem for any budget set.

The following result is a trivial consequence of Theorem 1.

Corollary 2. *If a data set O^T can be rationalized by a maximin preference function, it can be rationalized by a decisive maximin preference function.*

Definition 11. *($W_{\mathbb{T}\setminus S}$ decisive demand set) We define the $W_{\mathbb{T}\setminus S}$ decisive demand set, or the set of all commodity bundles compatible with WGARP, when we leave all observations indexed by $S \subseteq \mathbb{T}$ out from O^T, by:*

$$D_{\mathbb{T}\setminus S}(p^{T+1}, w^{T+1}) = \{x \in X : O^T \setminus \{p^s, x^s\}_{s\in S} \cup \{p^{T+1}, x\} \text{ satisfies WGARP and } p^{T+1}x = w^{T+1}\}.$$

In the proof of Theorem 1 we show that we can identify the index set $\mathbb{T}$ with the set of reference points U. This means that we can drop reference points by removing observations from the data set O^T. The following result holds:

Lemma 3. *For any data set O^T with $T \geq 2$ satisfying WGARP, there is a set $S \subseteq \mathbb{T}$, with at least 2 elements such that the decisive demand set is nonempty, i.e., $D_{\mathbb{T}\setminus S}(p, w) \neq \emptyset$ for all $p, w \in P \times W$.*

It is obvious that if S is a singleton then the decisive demand set will be nonempty. If O^T and $T \geq 2$, then we can always drop all observations but any 2. In this new data set $\{p^s, x^s\}_{s\in S}$, WGARP holds and is equivalent to GARP, in which case the set $D_{\mathbb{T}\setminus S}(p, w)$ is nonempty (Varian 1982). Now we can define the following counterfactual demand set as:

Definition 12. *(Upper decisive demand set) We define the upper decisive demand set for any price and wealth pair $(p, w) \in P \times W$ by*

$$D^{\uparrow}(p, w) = \cup_{S\subset\mathbb{T}} D_{\mathbb{T}\setminus S}(p, w).$$

We are interested in the upper decisive demand set because it allow us to make well-defined (i.e., nonempty) counterfactual predictions, using the largest amount of information contained in a data set O^T that satisfies WGARP. The following result follows from the previous arguments:

Theorem 3. *If a data set O^T is rationalized by any decisive maximin preference function r in the set R, then for a given price-wealth pair $(p, w) \in P \times W$ the demand correspondence $\mathbf{x}_r(p, w)$ –associated with maximizing r– is such that:*

$$D^{\downarrow}(p, w) \subseteq \cup_{r\in R} \mathbf{x}_r(p, w) \subseteq D^{\uparrow}(p, w).$$

There is an obvious caveat to the previous exercise. The upper bound $D^{\uparrow}(p, w)$ might be uninformative. Consider the case where we drop all but one observation, in which case

WGARP is trivially satisfied.[13] We can refine the decisive maximin preference model in order to get more informative bounds. For instance, we can require that $V_{p,w}$ is the largest possible set (in the set inclusion sense) that delivers a nonempty demand correspondence. This requirement implies that we can focus our attention on the largest possible set S that has nonempty decisive demand sets.

Definition 13. *(Minimal upper decisive demand set) We define the minimal upper decisive demand set for any price and wealth pair* $(p, w) \in P \times W$ *by*

$$D^{\uparrow}(p, w) = \cup_{S \subseteq \mathcal{S}} D_{\mathbb{T} \setminus S}(p, w),$$

where the collection of sets $\mathcal{S} \subseteq 2^{\mathbb{T}}$ *is such that any* $S' \subseteq \mathbb{T}$ *that contains some* $S \in \mathcal{S}$*, has an empty* $D_{\mathbb{T} \setminus S'}(p, w)$*.*

Next, by means of an example, we illustrate the fact that the minimal upper decisive demand set can be informative.

Example 5. *(Illustration of the minimal upper decisive demand set) Consider a data set* O^3 *as in Example 4. Set the out-of-sample budget to* $p^{T+1} = \frac{1}{3}(p^1 + p^2 + p^3) = \frac{10}{3}(1, 1, 1)'$*, and* $w^{T+1} = 22$*.*

- *The set* $D_{\{2,3\}}(p^{T+1}, w^{T+1})$ *is characterized by all consumption bundles* x*, with corresponding shares* $s_l = p_l^{T+1} x_l / w^{T+1}$*, for* $l = \{1, 2, 3\}$*, such that: (i)* $s_1 \leq \frac{1}{3}$ *and* $s_2 > \frac{1}{9}(7 - 12s_1)$*, or (ii)* $s_1 > \frac{1}{3}$ *and* $s_2 > \frac{1}{3}(-2 + 9s_1)$*.*
- *The set* $D_{\{1,3\}}(p^{T+1}, w^{T+1})$ *is characterized by all consumption bundles* x*, with corresponding shares* s_l*, for* $l = \{1, 2, 3\}$*, such that: (i)* $s_1 < \frac{1}{3}$ *and* $\frac{1}{3}(-2 + 9s_1) < s_2 < \frac{1}{12}(5 - 3s_1)$*.*
- *The set* $D_{\{1,2\}}(p^{T+1}, w^{T+1})$ *is characterized by all consumption bundles* x*, with corresponding shares* s_l*, for* $l = \{1, 2, 3\}$*, such that: (i)* $s_1 > \frac{1}{3}$ *and* $\frac{1}{9}(7 - 12s_1) < s_2 < \frac{1}{12}(5 - 3s_1)$*.*

The set $D^{\uparrow}(p^{T+1}, w^{T+1})$ *is the union of the previous three sets and is depicted in Figure 1.*

Remarkably, we can observe that the intersection of the three sets corresponding to dropping one observation from the data set, is empty. This empty set corresponds to the WGARP counterfactual demand set. In contrast, the set $D^{\uparrow}(p^{T+1}, w^{T+1})$ (the union of the aforementioned sets) is nonempty and informative. In fact, a large region in the budget share-simplex is excluded as a possible prediction. As we just mentioned, this set is the most informative counterfactual set that is compatible with the decisive maximin preference model, because it uses all of the available information imposing WGARP.

[13] Note that even in this case counterfactuals may be informative albeit containing very little information.

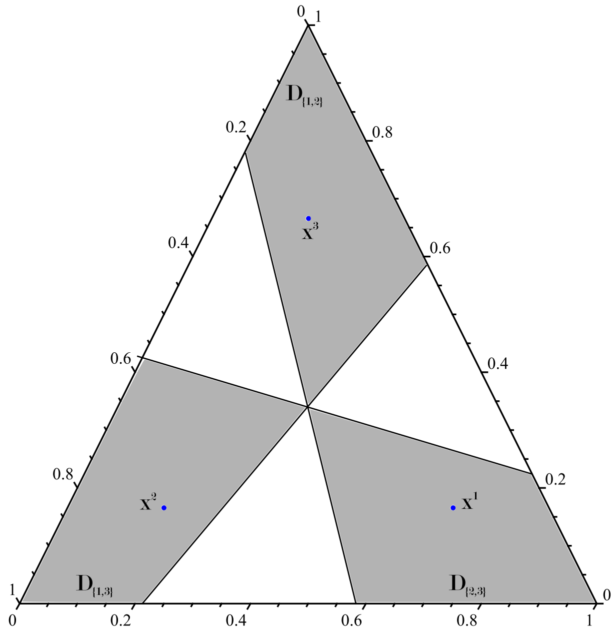

Figure 1 – Minimal Upper Decisive Demand Set: Shaded areas represent the Minimal Upper Decisive Demand Set.

4.2. Computing the Minimal Upper Decisive Demand Set

In this subsection, we outline a simple and practical procedure to compute the minimal upper decisive demand set just introduced. This procedure is based on the following equivalent reformulation of the (minimal) upper decisive demand set:

Proposition 1. *The following statements hold:*

(i) The bundle $x^{T+1} \in X$ is in the upper decisive demand set if and only if there exist binary variables $e^t \in \{0,1\}$ for all $t \in \mathbb{T}$ such that the following inequalities hold:

$$p^{T+1}x^{T+1} = w^{T+1}, \tag{1}$$

$$p^t x^{T+1} \geq e^t p^t x^t, \quad \textit{for all} \quad t \in \mathbb{T}, \quad \textit{for which} \quad p^{T+1}x^t \leq p^{T+1}x^{T+1}, \tag{2}$$

$$p^t x^{T+1} > e^t p^t x^t, \quad \textit{for all} \quad t \in \mathbb{T}, \quad \textit{for which} \quad p^{T+1}x^t < p^{T+1}x^{T+1}. \tag{3}$$

(ii) The minimal upper decisive demand set is given by the observations corresponding to the maximal number of binary variables $\{e^t\}_{t\in\mathbb{T}}$ equal to unity such that the expressions in Eqs. (1)-(3) hold.

The minimal upper decisive demand set in (ii) can be computed in the L1-norm by solving the following problem:

$$\arg \min_{\{e^1,\ldots,e^T,x^{T+1}\}} \sum_{t\in\mathbb{T}} (1-e^t) \quad \text{s.t.} \quad \text{Eqs. (1)} - \text{(3)} \quad \text{hold.} \tag{4}$$

However, this problem is difficult to implement because of the logical relations appearing in (2) and (3). An efficient practical implementation is possible by formulating the problem as a mixed integer linear programming (MILP) problem. Specifically, the inequalities in Eqs. (2) and (3) are equivalent to the existence of binary variables $\{y_1^t\}_{t\in\mathbb{T}}$ and $\{y_2^t\}_{t\in\mathbb{T}}$, with $y_1^t \leq y_2^t$, such that the following inequalities hold:

$$A^t(y_1^t - 1) \leq p^{T+1}(x^{T+1} - x^t), \tag{5}$$

$$B_1^t(y_1^t - 1) \leq p^t(x^{T+1} - e^t x^t), \tag{6}$$

$$A^t(y_2^t - 1) + \epsilon \leq p^{T+1}(x^{T+1} - x^t), \tag{7}$$

$$B_2^t(y_1^t - 1) \leq p^t(x^{T+1} - e^t x^t), \tag{8}$$

where $A^t \geq p^{T+1}x^t$, $B_1^t \geq p^t x^t$, $B_2^t \geq p^t x^t + \epsilon$ and $0 < \epsilon < w^{T+1}$. We have that $y_1^t = 1$ if and only if $p^{T+1}(x^{T+1} - x^t) \geq 0$ and $p^t(x^{T+1} - e^t x^t) \geq 0$, which correspond to (2). Similarly, $y_2^t = 1$ if and only if $p^{T+1}(x^{T+1} - x^t) > 0$ and $p^t(x^{T+1} - e^t x^t) > 0$, corresponding to (3). Hence, we suggest to replace (2) and (3) in the problem (4) with the linear inequalities (5)-(8), and compute the minimal decisive demand set by solving the following MILP problem:

$$\arg \min_{\{e^1,\ldots,e^T, y_1^1,\ldots,y_1^T, y_2^1,\ldots,y_2^T, x^{T+1}\}} \sum_{t\in\mathbb{T}} (1 - e^t) \quad \text{s.t.} \quad \text{Eqs. (1) and (5)} - \text{(8)} \quad \text{hold.} \tag{9}$$

This problem gives an exact and global solution, and there exist efficient algorithms for solving such MILP problems in practice (e.g., branch and bound and cutting plane).

5. Shape Constraints: Concave Rationalization and the Law of Demand

We have shown, in sections 3.3 and 3.4, that WGARP (WARP) is a necessary and sufficient condition for a data set to be rationalized by a continuous, strictly increasing, skew-symmetric, and piecewise (strictly) concave preference function. In this section, we consider conditions that are necessary and sufficient for a data set to be rationalized under stronger shape restrictions, while also addressing the issue of nonemptiness of the solution in the consumer's maximization problem. The results offered here highlight new applications of the maximin preference model, and add, with respect to Theorem 1, quasihomothetic and quasilinear restrictions of the preference function.

5.1. Concave Rationalization

John (2001) provides conditions under which a data set can be weakly rationalized by a continuous, strictly increasing, skew-symmetric, and *concave* preference function (Shafer preference function). Given that (quasi)concavity is a testable condition (See Example 6 in the next section) and the generic failure of WGARP to produce convex preferences, John's conditions are stronger than WGARP. Nevertheless, these conditions are weaker than GARP because they still relax transitivity. In this section, we extend the results in John (2001) by providing restrictions on the local utilities associated with the maximin preference function in order to guarantee a rationalization with convex preferences. Specifically, we show that when the local utilities are restricted to be *quasihomothetic*, i.e., when the indirect local utility functions are consistent with the Gorman polar form, then the data can be rationalized with a Shafer concave preference function. More precisely, we require that the marginal utility of income associated with each local utility function is the same, or in other words, the marginal utility is *reference-independent*. This result might be of interest in its own right because it links reference-dependence in the presence of multiple attributes/selves with aggregation theory. Empirically, these restrictions are helpful because they guarantee that nonparametric counterfactual demand analysis can always be performed without imposing transitivity. Moreover, our new notion of maximin rationalization with the additional restriction that the local utilities are quasihomothetic opens the door to provide tractable models of nontransitive behavior using well-known Gorman polar functional forms in parametric demand modelling. In passing we highlight, that consistent aggregation of reference-dependent models of behavior may be important to guarantee well-behaved counterfactual analysis.

In order to introduce our results, we need some preliminaries. The indirect utility function is defined as:

$$v(p, w) = \max_{x} \{u(x) | px \leq w\}.$$

Next, we define the concept of rationalization by quasihomothetic (Gorman polar) preferences.

Definition 14. *(Quasihomothetic/Gorman polar utility rationalization) A data set O^T is rationalized by quasihomothetic (Gorman polar) preferences if there exists a continuous, strictly increasing and concave utility function, u, that rationalizes the data, and such that the associated indirect utility function can be written $v(p, w) = \frac{w-a(p)}{b(p)}$, where the functions $a(p) : P \to \mathbb{R}$ and $b(p) : P \to \mathbb{R}$ are homogeneous of degree one.*

The function b can be interpreted as the inverse of the marginal utility of income, and it will play a crucial role in obtaining the additional restrictions on the local utilities in the maximin model to obtain a rationalization with a concave preference function. To state these results, we define in the obvious manner the maximin quasihomothetic preference model and the notion of maximin quasihomothetic preference rationalization, where for any reference point indexed $i, j \in U$, the local utility function u_{ij} is continuous, strictly increasing, concave, and quasihomothetic. In particular, the indirect utility functions, v_{ij}, associated with each local utility function are Gorman polar forms with reference-independent marginal utilities, i.e., $v_{ij}(p, w) = \frac{w}{b(p)} - \frac{a_{ij}(p)}{b(p)}$, such that $a_{ij} = a_{ji}$ for all $i, j \in U$. The next theorem states our main result in this section.

Theorem 4. *Consider a finite data set $O^T = \{p^t, x^t\}_{t \in \mathbb{T}}$. The following statements are equivalent:*

(i) The data O^T can be weakly rationalized by a locally nonsatiated, concave, and skew-symmetric preference function.

(ii) There exist numbers $R^{t,s}$ and $\lambda^t > 0$ for all $s, t \in \mathbb{T}$ with $R^{t,s} = -R^{s,t}$ such that the inequalities:

$$R^{t,s} \geq \lambda^t p^t (x^t - x^s),$$

hold for all $s, t \in \mathbb{T}$.

(iii) The data set O^T can be weakly rationalized by a maximin quasihomothetic preference function.

(iv) The data O^T can be weakly rationalized by a continuous, strictly increasing, concave, and skew-symmetric preference function.

John (2001) proves the equivalence of (*i*), (*ii*), and (*iv*). We prove the equivalence of (*ii*) and (*iii*). In particular, this is a new result that connects the general nontransitive consumer of Shafer (1974) with the maximin preference model introduced in this paper. Strictly speaking, the Shafer preference function and the maximin preference function are different models of behavior. However, Theorem 4 establishes their empirical equivalence. As representations, both formulations have their own pros and cons but we can use them interchangeably when modelling consumer demand.

In comparison with Theorem 1, it is easy to see that the testable condition (ii) in Theorem 4 imply any of the conditions in our theorem, but not vice versa. In particular, note that, in contrast to statement (iii) in Theorem 1, the indices λ^t are constant across all pairs $s, t \in \mathbb{T}$. This ensures that the maximin quasihomothetic preference function is indeed concave.

5.2. The Law of Demand and Quasilinear Preference Functions

This subsection derives necessary and sufficient conditions for a finite data set to be rationalized by a continuous, strictly increasing, skew-symmetric, concave, and quasilinear preference function. Interestingly, we show that one such condition is the law of demand,

and consequently, this is equivalent to rationalization by a maximin quasilinear preference function. Before presenting these results, we briefly recall the revealed-preference characterization for quasilinear-utility maximization.

5.2.a. *The Quasilinear Utility Maximization Model.–* First, we consider the definition of quasilinear utility maximization:

Definition 15. *(Quasilinear utility maximization) Consider a locally nonsatiated utility function $u(x)$. We say that a consumer facing prices $p \in P$ and income $w \in W$ is a quasilinear utility maximizer if she solves*

$$\max_{x \in X} u(x) + w - px \iff \max_{x \in X, y \in \mathbb{R}} u(x) + y \;\; s.t.\; px + y \leq w.$$

As in standard applications of quasilinear utility maximization, we allow the numeraire y to be negative in order to avoid technicalities related to corner solutions.[14]

Brown and Calsamiglia (2007) shows that the axiom of cyclical monotonicity is a necessary and sufficient condition for a data set to be rationalized by a continuous, strictly increasing, concave, and quasilinear utility function.

Axiom 5. *(Cyclical monotonicity) Cyclical monotonicity holds if, for all distinct choices of indices* $(1, 2, 3, \ldots, n) \in \mathbb{T}$*:*

$$p^1(x^1 - x^2) + p^2(x^2 - x^3) + \cdots + p^n(x^n - x^1) \leq 0.$$

The next theorem recalls the revealed-preference characterization of quasilinear utility maximization from Brown and Calsamiglia (2007) and Allen and Rehbeck (2018):

Theorem E. (Brown and Calsamiglia 2007; Allen and Rehbeck 2018) Consider a finite data set $O^T = \{p^t, x^t\}_{t \in \mathbb{T}}$. The following statements are equivalent:

(i) The data O^T can be rationalized by a locally nonsatiated and quasilinear utility function.

(ii) The data O^T satisfies cyclical monotonicity.

(iii) There exist numbers U^t for all $t \in \mathbb{T}$ such that the inequalities:

$$U^t - U^s \geq p^t(x^t - x^s),$$

hold for all $s, t \in \mathbb{T}$.

(iv) The data O^T can be rationalized by a continuous, strictly increasing, concave, and quasilinear utility function.

[14] Allen and Rehbeck (2018) shows the equivalence of the unconstrained quasilinear maximization and the constrained version with a numeraire in Definition 15.

5.2.b. *The Law of Demand and the Quasilinear Preference Function Maximization Model.–* In this subsection, we provide analogous results for the quasilinear preference function model, and show that the law of demand is a necessary and sufficient condition for a data set to be rationalized by a continuous, strictly increasing, concave, skew-symmetric, and quasilinear preference function. The axiom of the law of demand is formally defined as:

Axiom 6. *(Law of demand) The law of demand holds if, for all observations $s, t \in \mathbb{T}$:*

$$(p^t - p^s)(x^t - x^s) \leq 0.$$

For any sequence consisting of only two (distinct) observations $s, t \in \mathbb{T}$, it is easy to see that cyclical monotonicity and the law of demand are equivalent. To state our revealed-preference characterization of the law of demand, we introduce the maximin quasilinear preference function, which is a maximin preference function where, for any reference point indexed by $i, j \in U$, the local utility function $u_{i,j}$ is continuous, strictly increasing, concave, and quasilinear. This model is a special case of the maximin quasihomothetic preference model, in which the functions $a_{ij}(p)$ and $b(p)$ in the indirect local utility function $v_{ij}(p, w) = \frac{w}{b(p)} - \frac{a_{ij}(p)}{b(p)}$ takes the more restrictive forms: $a_{ij}(p) = p_l \phi(p)$ and $b(p) = p_l$, where p_l is the price of a numeraire good and ϕ is a function that is homogeneous of degree one. The next theorem provides a revealed-preference characterization of the law of demand:

Theorem 5. *Consider a finite data set $O^T = \{p^t, x^t\}_{t \in \mathbb{T}}$. The following statements are equivalent:*

(i) The data O^T can be rationalized by a locally nonsatiated, skew-symmetric, and quasilinear preference function.

(ii) The data O^T satisfies the law of demand.

(iii) There exist numbers $R^{t,s}$, for all $s, t \in \mathbb{T}$, with $R^{t,s} = -R^{s,t}$, such that inequalities:

$$R^{t,s} \geq p^t(x^t - x^s),$$

hold for all $s, t \in \mathbb{T}$.

(iv) The data O^T can be rationalized by a maximin quasilinear preference function.

(v) The data O^T can be rationalized by a continuous, strictly increasing, concave, skew-symmetric, and quasilinear preference function.

6. Recoverability of Preferences

Next, we tackle the question of when one can use the direct revealed-preference relation elicited from the observed consumer behavior in order to make inferences about her true preferences. We begin by showing that recovering preferences using WGARP does not follow as a trivial corollary of the original approach proposed by Varian (1982). Subsequently, we propose an alternative method to recover bounds on preferences using WGARP.

6.1. Varian's Approach to Recover Bounds on Preferences Using WGARP

At this point, it is useful to briefly recall the classical approach from Varian (1982), which finds upper and lower bounds to the true preferences of a consumer. These are captured by the strict upper contour set of a commodity bundle x according to the true preference function r:

Definition 16. *(Set of strictly better alternatives) We define the set of strictly better alternatives than a (possibly unobserved) commodity bundle $x \in X$ as:*

$$U_r(x) = \{y \in X : r(y, x) > 0\},$$

for the true preference function r.

Varian (1982) defines the supporting set of prices for any new commodity bundle $x \in X$, so that the *extended data set*, $O^T \cup \{p, x\}$, satisfies GARP as:

$$S(x) = \{p \in P : O^T \cup \{p, x\} \text{ satisfies GARP}\}.$$

Varian then uses the set $S(x)$ to create upper and lower bounds for the set of interest $U_r(x)$. We need to define two new sets. The *revealed worse set* is:

$$RW(x) = \{y \in X : \forall p \in S(x), x \succ^{R,D}_{O^T \cup \{(p,x)\}} y\}$$

for $\succ^{R,D}$, defined on the extended data set $O^T \cup \{(p, x)\}$. The *nonrevealed worse set* $NRW(x)$ is the complement of $RW(x)$. The *revealed preferred set* is:

$$RP(x) = \{y \in X : \forall p \in S(y), y \succ^{R,D}_{O^T \cup \{(p,y)\}} x\}.$$

Varian (1982) shows that, in the case of utility maximization (i.e., $r(x, y) = u(x) - u(y)$), for some $u : X \to \mathbb{R}$ and all $x, y \in X$, we have:

$$RP(x) \subseteq U_r(x) \subseteq NRW(x).$$

One would be tempted to use the same construction for WGARP by replacing the definition of the supporting set $S(x)$ with one where the extended data set satisfies WGARP. Of course, when the data consists of two goods (i.e., $L = 2$), this does not cause any problems since, in such a case, WGARP and GARP are equivalent (See Theorem A). However, if $L > 2$, as we show, performing such an exercise is generally not advisable. In particular, we illustrate this by means of an example that, in some cases, yields an uninformative upper bound set $NRW(x)$.

Example 6. *Consider again the data set O^3 with prices $p^1 = (4, 1, 5)'$, $p^2 = (5, 4, 1)'$, $p^3 = (1, 5, 4)'$, and bundles $x^1 = (4, 1, 1)'$, $x^2 = (1, 4, 1)'$, $x^3 = (1, 1, 4)'$. It is easy to verify that this data set satisfies WGARP. Given the observed behavior, suppose the goal is to recover the preferences of this consumer for a new commodity bundle, x^{T+1}. Suppose that the unobserved commodity bundle is:*

$$x^{T+1} = \frac{1}{3}(x^1 + x^2 + x^3) = (2, 2, 2)'.$$

If one were to use the methods in Varian (1982), it is necessary to recover all prices p^{T+1} such that the extended data set $O^3 \cup (p^{T+1}, x^{T+1})$ satisfies WGARP. In the extended data set, we have $p^t(x^t - x^{T+1}) = 2 > 0$, for all $t = 1, 2, 3$. However, there is no $p \in P$ for which $p(x^{T+1} - x^t) < 0$, for all $t = 1, 2, 3$. This implies that the extended data set violates WGARP. This presents a problem if the goal is to recover preferences using Varian's (1982) approach, because this method implicitly assumes that there always exists at least one such vector of prices satisfying WGARP.

In this example, Varian's supporting set is empty, i.e., $S(x^{T+1}) = \emptyset$. Moreover, it directly follows that the set $U_r(x)$ may contain any monotonically dominated bundle such as $x^- = (1, 1, 1)$. Consequently, the upper bound of $U_r(x^{T+1})$ is uninformative, i.e., $NRW(x^{T+1}) = X \setminus x^{T+1}$. Thus, any analysis based on this approach is problematic, since the observed behavior can be rationalized by a preference function that is strictly increasing (in the first argument). In other words, Varian's method to bound preferences does not provide any valuable information in Example 6.

We can also clarify the source of this failure, traced to a violation of convexity of preferences. Consider once again Example 6, and note that the observed data satisfies WGARP, which implies that there is a preference function r rationalizing the data. Moreover, we have $r(x_1, x_2) > 0$, $r(x_2, x_3) > 0$, and $r(x_3, x_1) > 0$. In addition, we know that the new bundle, x^{T+1}, is a convex combination of the observed bundles. In our case, convexity of preferences is implied by the assumption that the preference function is quasiconcave in its first argument, in which case, we must have $r(x^{T+1}, x^{T+1}) = 0 \geq \min_{t=1,2,3}\{r(x^t, x^{T+1})\}$. This implies that x^{T+1} must be revealed to be weakly better than at least one of the three observed bundles x^1, x^2, or x^3. However, note that, for all $t = 1, 2, 3$, we have $p^t(x^t - x^{T+1}) = 2 > 0$. Thus, if the consumer were maximizing a quasiconcave preference function, then all observed commodity bundles must be strictly preferred to the new bundle, i.e., $r(x^t, x^{T+1}) > 0$ for all $t = 1, 2, 3$. Hence, the extended data set $O^3 \cup \{p, x^{T+1}\}$ cannot be weakly rationalized by a quasiconcave and skew-symmetric preference function.

Interestingly, this example shows that quasiconcavity of the preference function is, in fact, a testable property in finite data sets. As such, this is also a counterexample to Samuelson's *eternal darkness conjecture*, saying that any finite data set always can be rationalized by a convex preference relation. Summarizing these results, the lack of convexity of preferences, which can be inferred from behavior consistent with WGARP, limits the applicability of the tools developed in the classical treatment by Varian (1982).

6.2. A New Approach to Recover Bounds on Preferences Using WGARP

In this subsection, we use the new notion of maximin preference rationalization as a way to provide new informative bounds on the true preferences. We show that these new bounds escape the problems associated with Varian's approach.

The proof of Theorem 1 shows that, without loss of generality, we can identify the set of reference points in U with the set of observations $\mathbb{T}$, such that the true global preferences for any $x', x \in X$ are given by:

$$r(x', x) = \max_{\mu \in \Delta(\mathbb{T})} \min_{\lambda \in \Delta(\mathbb{T})} \sum_{t \in \mathbb{T}} \sum_{s \in \mathbb{T}} \lambda_s \mu_t (u_{st}(x') - u_{st}(x)).$$

The proof also shows that any data set O^T satisfying WGARP can be broken into T^2 pairwise data sets $O^2_{st} = \{(p^t, x^t), (p^s, x^s)\}$, and we argue that each one of these data sets satisfies GARP. For any pair of observations $s, t \in \mathbb{T}$, we define the local support set $S_{st}(x)$ for any $x \in X$ as in Varian (1982). Hence, for a data set of T observations, we have a collection of T^2 such local support sets. Note, by definition, that everyone of these sets is never empty. Thus, consider the following two definitions:

Definition 17. *(WGARP-robust revealed preferred set) For each $s, t \in \mathbb{T}$ let*

$$RP_{st}(x) = \{y \in X : \forall p \in S_{st}(y), py > px\}$$

be the pairwise revealed preferred set. We define the (WGARP-)robust revealed preferred set as:

$$RP^W(x) = \cup_{s \in \mathbb{T}} \cap_{t \in \mathbb{T}} RP_{st}(x).$$

Next, we argue that the robust revealed preferred set is a lower bound of $U_r(x)$ for all $x \in X$. If $x' \in RP^W(x)$, this implies $x' \in RP_{st}(x)$ for all $t \in \mathbb{T}$ and for some $s^* \in \mathbb{T}$. Thus, it must be the case that, for s^* and for all $t \in \mathbb{T}$, $u_{s^*t}(x') > u_{s^*t}(x)$, which means that $r(x', x) \geq \min_\mu \sum_t \mu_t(u_{s^*t}(x') - u_{s^*t}(x)) > 0$. Hence, if $x' \in RP^W(x)$, then we have $r(x', x) > 0$, which can be equivalently stated as: $RP^W(x) \subseteq U_r(x)$.

Definition 18. *(WGARP-robust (non)revealed worse set) For each $s, t \in \mathbb{T}$, let*

$$RW_{st}(x) = \{y \in X : \forall p \in S_{st}(x), px > py\}$$

be the pairwise revealed worse set. Let $NRW_{st}(x)$ be the complement of $RW_{st}(x)$. Define the (WGARP-)robust nonrevealed worse set as

$$NRW^W(x) = \cap_{s\in\mathbb{T}} \cup_{t\in\mathbb{T}} NRW_{st}(x).$$

From this definition, it directly follows that, if $r(x', x) > 0$, then $x' \in NRW^W(x)$. In particular, note that, if $r(x', x) > 0$, then there must be some $t^* \in \mathbb{T}$ such that $u_{st^*}(x') > u_{st^*}(x)$ for all $s \in \mathbb{T}$. By a direct application of the results in Varian (1982), we have $x' \in NRW_{st^*}(x)$ for all $s \in \mathbb{T}$. Then, by Definition 18, it follows that $x' \in NRW^W(x)$. Hence, this proves that $U_r(x) \subseteq NRW^W(x)$. The following theorem summarizes these results, confirming that the bounds recovered using Varian's approach in this context are not sharp:

Theorem 6. *The upper contour set of the true preferences at any given $x \in X$ is:*

$$RP^W(x) \subseteq U_r(x) \subseteq NRW^W(x).$$

Moreover, (i) the upper bound, $NRW(x)$, recovered using Varian's approach is not sharp, i.e., $NRW^W(x) \subseteq NRW(x)$ for all $x \in X$ (with strict containment for some $x \in X$); and (ii) the lower bound, $RP(x)$, recovered using Varian's approach is not sharp, i.e., $RP(x) \subseteq RP^W(x)$ for all $x \in X$ (with strict containment for some $x \in X$).

We note that, in the context of Example 6, $NRW^W(x^{T+1})$ does not contain the dominated bundle $x^- = (1\,1\,1)'$. In fact, $NRW^W(x^{T+1})$ excludes all commodity bundles that are monotonically dominated by x^{T+1}, which is a desirable property, lacking in Varian's analogous set $NRW(x^{T+1}) = X \setminus x^{T+1}$. Similar statements can be made about the $RP^W(x^-)$ set.

Thus, the first part of Theorem 6 shows that the new method of using subsets of data sets to calculate bounds on preferences yields informative bounds. The second part highlights that a naive application of the methodology in Varian (1982), when the assumption of convex preferences does not hold, is problematic.

7. Infinite Data Sets: Characterizations of WGARP, WARP, and the Law of Demand

Thus far, our results have been derived under the assumption that the researcher only observes a finite number of choices. In the original formulation of revealed-preference theory, Samuelson (1938) and Houthakker (1950) implicitly assume that the entire demand function, or a demand correspondence, is observed. In this section, we show that our main results from the previous sections can be transported to the case of infinite data sets, namely, when we observe a demand correspondence $\mathbf{x} : P \times W \to 2^X \setminus \emptyset$, where $w \in W \equiv \mathbb{R}_{++}$ denotes wealth. We focus on compact sets of prices and consumption bundles, where, by abusing notation slightly, we denote $P \subset \mathbb{R}^L_{++}$ and $X \subset \mathbb{R}^L_+ \setminus 0$, as the sets of prices and consumption bundles, respectively. We continue to assume Walras' law, so that $x \in \mathbf{x}(p, w)$, $px = w$. We define the image of $\mathbf{x}$ as $\mathbf{X} = \cup_{p \in P, w \in W} \mathbf{x}(p, w)$. A central assumption throughout this section is that we can write the demand correspondence as a data set consisting of an infinite number of demand observations, which we denote by $O^\infty = \{p, x\}_{p \in P, x \in \mathbf{X} : x \in \mathbf{x}(p, px)}$.

7.1. WGARP and WARP

We begin by providing revealed-preference characterizations for WGARP and WARP. In doing so, we define the direct-preference relations for infinite data sets as:

Definition 19. *(Direct Revealed Preferences) We say that $x \in \mathbf{X}$ is directly revealed preferred to $y \in \mathbf{X}$, written $x \succeq^{R,D} y$, when $px \geq py$ such that $x \in \mathbf{x}(p, px)$. Also, $x \in \mathbf{X}$ is directly and strictly revealed preferred to $y \in \mathbf{X}$, written $x \succ^{R,D} y$ when $px > py$ and $x \in \mathbf{x}(p, px)$.*

Under this definition, the data O^∞ satisfies WGARP if there is no pair $x, y \in \mathbf{X}$ such that $x \succeq^{R,D} y$ and $y \succ^{R,D} x$. Analogously, the data O^∞ satisfies WARP if there is no pair $x, y \in \mathbf{X}$ such that $x \succeq^{R,D} y$ and $y \succeq^{R,D} x$ with $x \neq y$.

We begin by generalizing the maximin rational preference function to the case of infinite data sets. For this, we need some preliminaries. For any reference point in the data set O^∞, we rearrange the observations into a vector $o = (p' \; a' \; q' \; b')' \in \mathbf{O}^2$, with $\mathbf{O} \subseteq P \times \mathbf{X}$ and $x \in \mathbf{x}(p, px)$, such that each reference point can be viewed as a column vector. In addition, we define $o_1 = (p' \; a')'$ and $o_2 = (q' \; b')'$, such that $o = (o_1' \; o_2')'$.

We assume that the set of reference points $\mathbf{O}$ is compact, i.e., closed and bounded. There are several examples satisfying this condition; for instance, when there are a finite number of reference points, or when the demand correspondence that generates the data set is compact-valued. In the latter case, compactness of $\mathbf{O}$ follows from assuming that

$\mathbf{x}(p, px)$ is a compact set, which ensures that the entire set $P \times \mathbf{X}$ is compact.[15]

We consider a reference-dependent utility function, $u_\bullet : \mathbf{O}^2 \times X \to \mathbb{R}$, which rationalizes the data. That is, for every pair $o = (o_1' \, o_2')' \in \mathbf{O}^2$ and for all $y \in X$, if $px \geq (>) py$, then it must be the case that $u_o(x) \geq (>) u_o(y)$. We further assume that $u_\bullet$ is continuous, or more precisely, that it is continuous at the reference point for every commodity bundle. Moreover, we assume that the reference-dependent utility functions are independent of permutations, so that $u_{o_1 o_2} = u_{o_2 o_1}$ for all $o_1, o_2 \in \mathbf{O}$.

Let Σ denote a Borel σ-algebra defined on $\mathbf{O}$, and let $\Delta(\mathbf{O}, \Sigma)$ denote the simplex of Borel probability measures defined on $\mathbf{O}$ (We write $\Delta(\mathbf{O}) = \Delta(\mathbf{O}, \Sigma)$).[16] The next definition introduces the generalized maximin preference function:

Definition 20. *(Generalized Maximin (strict) preference model) We say that the preference function $r(x, y)$ is a generalized maximin (strict) preference function if, for any $x, y \in X$, it can be written as:*

$$r(x, y) = \max_{\mu \in \Delta(\mathbf{O})} \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} (u_{o_1 o_2}(x) - u_{o_1 o_2}(y)) d\lambda(o_1) d\mu(o_2),$$

where, for any reference point $o \in \mathbf{O}^2$, the local utility function, $u_{o_1 o_2}(\cdot)$, is continuous, strictly increasing, and (strictly) concave.

First, we have:

Lemma 4. *If r is a generalized maximin preference function, then for any $x, y \in X$, we have:*

$$\begin{aligned} r(x, y) &= \min_{\lambda \in \Delta(\mathbf{O})} \max_{\mu \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} (u_{o_1 o_2}(x) - u_{o_1 o_2}(y)) d\lambda(o_1) d\mu(o_2) \\ &= \max_{\mu \in \Delta(\mathbf{O})} \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} (u_{o_1 o_2}(x) - u_{o_1 o_2}(y)) d\lambda(o_1) d\mu(o_2), \end{aligned}$$

and moreover, r is skew-symmetric.

The first part of Lemma 4 follows from the continuous version of Von-Neumann's minimax theorem in Glicksberg (1950), and by the definition of $u_\bullet$ that guarantees that it is a continuous mapping.[17] In this framework, we can obviously define the concept of rationalization as before.

[15] Note that compact-valuedness of the demand correspondence graph is a very general assumption and holds for the case of continuous preferences maximized over linear budget sets when the sets of prices and wealth are compact.

[16] The set $\Delta(\mathbf{O}, \Sigma)$ is endowed with the usual weak* topology. Specifically, since O is a metrizable space, the topology is endowed with the Prokhorov metric. Also note that $\Delta(\mathbf{O}, \Sigma)$ is a compact metric space, because $\mathbf{O}$ is assumed to be a compact metric space. This follows from Alaoglu's theorem (See e.g., Dunford and Schwartz 1958, p.424, Theorem V.4.2).

[17] Note that we can always construct a continuous $u_\bullet$, if we build the utilities associated with each reference point following Varian (1982). We elaborate this technical point in the proof of the main theorem of this section.

The next theorem shows that WGARP (WARP) is a necessary and sufficient condition for an infinite data set to be rationalized by a generalized maximin (strict) preference function:

Theorem 7. *Consider an infinite data set O^∞. The following statements are equivalent:*

(i) The data O^∞ can be (strictly) weakly rationalized by a locally nonsatiated and skew-symmetric preference function.

(ii) The data O^∞ satisfies (WARP) WGARP.

(iii) The data O^∞ can be (strictly) weakly rationalized by a generalized maximin preference function.

(iv) The data O^∞ can be (strictly) weakly rationalized by a continuous, strictly increasing, (strictly) piecewise concave, and skew-symmetric preference function.

Some remarks are pertinent:

First, if the data O^∞ satisfies WARP, then it must hold for any observation $(p, x) \in O^\infty$ that $x = \mathbf{x}(p, px)$, in which case the demand correspondence is actually a demand function. Hence, in the weak sense, Theorem 7 rationalizes demand correspondences satisfying WGARP, and in the strict sense, it rationalizes demand functions satisfying WARP.

Second, Theorem 7 generalizes the results in Kim and Richter (1986) and Quah (2006). Specifically, the key assumption in these papers is that the demand correspondence satisfies an invertibility condition, i.e., that for every commodity bundle $x \in X$, there exists a price $p \in P$ at which x is demanded (with wealth px). In contrast, the results in Theorem 7 are not based on any such invertibility condition. Note that this condition is violated in our motivating example (Example 6).

Third, as discussed above, our main assumption is that the graph of the demand correspondence is compact. Note that, for the case of demand functions, this is trivially true. For demand correspondences, maximizing a continuous preference function on compact sets of prices and wealth, implies, by Berge's maximum theorem, that the correspondence is compact-valued. Consequently, this assumption is indeed a very weak condition.[18]

Finally, we have not explicitly assumed homogeneity of degree zero. In fact, homogeneity can be imposed, since it is implied by maximin rationalization, and as such, we can normalize wealth to 1 without loss of generality.

[18]We note that the compactness condition can be relaxed to some degree if we substitute the min and max operators by supremum and infimum. However, this will result in some additional technicalities that are not of any practical interest.

7.2. The Law of Demand

This subsection is devoted to a characterization of the law of demand for infinite data sets, defined as:

Definition 21. *(Law of demand) The law of demand holds if, for all $x, y \in \mathbf{X}$:*

$$(p - q)(x - y) \leq 0,$$

such that $x \in \mathbf{x}(p, px)$ and $y \in \mathbf{x}(q, qy)$.

Under this definition, with the notation and assumptions from the previous subsection, we define the generalized maximin quasilinear preference function as follows:

Definition 22. *We say that the preference function $r(x, y)$ is a generalized maximin quasilinear preference function if, for any $x, y \in X$, it can be written as:*

$$r(x, y) = \max_{\mu \in \Delta(\mathbf{O})} \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} (u_{o_1 o_2}(x) - u_{o_1 o_2}(y)) d\lambda(o_1) d\mu(o_2),$$

where, for any reference point $o \in \mathbf{O}^2$, the local utility function, $u_{o_1, o_2}(\cdot)$, is continuous, strictly increasing, concave, and quasilinear.

Of course, we can then adapt the definitions of rationalization and state the following theorem:

Theorem 8. *Consider an infinite data set O^∞. The following statements are equivalent:*

(i) The data O^∞ can be rationalized by a locally nonsatiated, skew-symmetric, and quasilinear preference function.

(ii) The data O^∞ satisfies the law of demand.

(iii) The data O^∞ can be rationalized by a generalized maximin quasilinear preference function.

(iv) The data O^∞ can be rationalized by a continuous, strictly increasing, concave, skew-symmetric, and quasilinear preference function.

8. Related Literature

In this section, we extend our discussion of the relationship with previous works. Afriat (1967) and Varian (1982) show that the classical notion of rationality is equivalent to the Generalized Axiom of Revealed Preference (GARP). In the current study, we show that a data set that satisfies WGARP, but perhaps not GARP, is consistent with a local (reference-dependent) notion of rationalization. Of course, classical utility maximization is a special case, when there is a (global) utility function u that is capable of rationalizing the data set O^T. In that case, $r(x, y) = u(x) - u(y)$ for all $x, y \in X$; in other words, $u = u_{ij}$ for all $i, j \in U$.

The closest works to our paper are Kim and Richter (1986) and Quah (2006). Both works provide rationalizations of demand correspondences, or functions, consistent with WGARP or WARP, using additional conditions on the invertibility of demand, with preferences that are convex (in a certain sense).[19] Our paper generalizes these contributions by (i) providing a rationalization of WGARP/WARP for finite data sets, and (ii) relaxing the invertibility requirement that for every commodity bundle in X, there is a price in its supporting set (i.e., the set of prices at which the commodity bundle is chosen is nonempty) for the case of infinite data sets. As we have seen in Example 6, there are commodities with an empty supporting set.[20] Instead, our work imposes the weak technical condition of compactness of the graph of the demand correspondence.

It is worth briefly stressing this point made in the introduction. Preference functions with the skew-symmetry property were introduced by Shafer (1974). We have shown that rationalization by skew-symmetric preferences is essentially equivalent to WGARP. Moreover, we have also shown that WGARP is equivalent to rationalization by a new kind of preference function, the maximin preference function, and our results answer in the negative the conjecture posed in Kihlstrom et al. (1976), concerning the equivalence between Shafer's skew-symmetric preference functions and WGARP.

Krauss (1985) provides a representation of 2-monotone operators (effectively equivalent to the law of demand), by means of a skew-symmetric preference function. To our knowledge, our results regarding WGARP are new in the mathematical literature on monotone operators as well, extending the contribution of Krauss (1985) to 2-cyclical consistent operators (effectively equivalent to WGARP). We also provide an extension for the original representation of the law of demand, connecting it with maximin quasilinear rationalization, as in Brown and Calsamiglia (2007), as well as covering the case of limited data sets.

Some papers have extended Varian's (1982) method to recover preferences to different types of consumer demand models. One notable example is Blundell et al. (2003, 2008) which show that it is possible to substantially enhance recovery and prediction results by combining revealed-preference theory with the nonparametric estimation of Engel curves.

[19] The notions of convexity of preferences in Kim-Richter and Quah are strictly weaker than standard convexity of preferences.

[20] Mariotti provides a study of WGARP in abstract environments with complete data sets (all choice sets are observed), and shows that WARP is equivalent to the maximization (in a new sense that he calls justified, of a binary preference relation that is asymmetric). Mariotti does not provide a representation theorem for WGARP nor does he deal with limited data sets.

However, their analysis is based on WARP, which we have shown may be problematic in a setup with more than two goods (Blundell et al. (2003) considers 22 goods and Blundell et al. (2008) uses three goods in their empirical applications). Blundell et al. (2015) shows how the methods in Blundell et al. (2003, 2008) can be modified to derive sharp bounds on welfare measures under SARP (i.e., global rationality).

Finally, Halevy et al. (2017) shows that Varian's method to recover preferences under GARP does not apply to nonconvex preferences, and suggests an alternative method based on monotonicity. However, when GARP holds, concavity is not a testable restriction. Our analysis provides a different solution based on local concavity (i.e., piecewise concavity), which, being more informative than monotonicity, is also robust to the possible lack of convexity of preferences, when the data satisfies WGARP but violates GARP. Note that, in our setup, convexity of preferences is, in fact, a testable condition.

9. Conclusion

This paper has provided a new notion of rationalization, the maximin preference function, which is equivalent to Samuelson's WGARP. It has built a comprehensive theory of revealed preference on the basis of this notion. Our findings should be helpful for practitioners of revealed preferences since, from an empirical perspective, WGARP is significantly easier to work with than Varian's GARP. In applications, it is common for practitioners to use WGARP as a synonymous of GARP. However, as shown in Cherchye et al. (2018), this is only true if price variation is limited.[21] For example, it may happen that a finite data set of prices and observed consumption choices is consistent with WGARP, but cannot be rationalized by a utility function. If this occurs, the interpretation of the direct revealed-preference relation is unclear, yet we have shown that meaningful welfare and counterfactual analysis is possible.

References

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

Allen, R. and Rehbeck, J. (2018). Assessing Misspecification and Aggregation for Struc-

[21] For a generalized treatment of when WARP implies rationality, see Caradonna (2018).

tured Preferences. *Unpublished working paper.*

Banerjee, S. and Murphy, J. (2006). A simplified test for preference rationality of two-commodity choice. *Experimental Economics*, 9(1):67–75.

Blundell, R., Browning, M., Cherchye, L., Crawford, I., De Rock, B., and Vermuelen, F. (2015). Sharp for SARP: Nonparametric bounds on counterfactual demands. *American Economic Journal: Microeconomics*, 7(1):43–60.

Blundell, R., Browning, M., and Crawford, I. (2003). Nonparametric engel curves and revealed preference. *Econometrica*, 71(1):1227–1262.

Blundell, R., Browning, M., and Crawford, I. (2008). Best Nonparametric Bounds on Demand Responses. *Econometrica*, 76(6):1227–1262.

Brown, D. and Calsamiglia, C. (2007). The nonparametric approach to applied welfare analysis. *Economic Theory*, 31(1):183–188.

Caradonna, P. (2018). How strong is the weak axiom? Technical report, Georgetown University, Department of Economics.

Cherchye, L., Crawford, I., De Rock, B., and Vermeulen, F. (2016). Gorman revisited: nonparametric conditions for exact linear aggregation. *SERIEs*, 7(2):203–220.

Cherchye, L., Demuynck, T., and De Rock, B. (2018). Transitivity of preferences: when does it matter? *Theoretical Economics*, 13(3):1043–1076.

Cherchye, L., Demuynck, T., and De Rock, B. (2019). Bounding counterfactual demand with unobserved heterogeneity and endogenous expenditures. *Journal of Econometrics*, forthcoming.

Cosaert, S. and Demuynck, T. (2018). Nonparametric welfare and demand analysis with unobserved individual heterogeneity. *Review of Economics and Statistics*, 100(2):349–361.

Demuynck, T. and Hjertstrand, P. (2019). Samuelson's approach to revealed preference theory: Some recent advances. In Anderson, R., Barnett, W., and Cord, R., editors, *Paul Samuelson: Master of Modern Economics*. Palgrave MacMillan, London.

Dunford, N. and Schwartz, J. T. (1958). *Linear operators part I: General theory*, volume 7. Interscience publishers New York.

Frick, M., Iijima, R., and Le Yaouanq, Y. (2019). Boolean expected utility. Working paper.

Glicksberg, I. L. (1950). *Minimax theorem for upper and lower semi-continuous payoffs*. Rand Corporation.

Halevy, Y., Persitz, D., and Zrill, L. (2017). Non-parametric bounds for non-convex preferences. *Journal of Economic Behavior and Organization*, 137:105–112.

Hoderlein, S. and Stoye, J. (2014). Revealed preferences in a heterogeneous population. *Review of Economics and Statistics*, 96(2):197–213.

Houthakker, H. (1950). Revealed preference and the utility function. *Economica*, 17(66):159–174.

John, R. (2001). The concave nontransitive consumer. *Journal of Global Optimization*, 20(3-4):297–308.

Keiding, H. and Tvede, M. (2013). Revealed smooth nontransitive preferences. *Economic Theory*, 54(3):463–484.

Kihlstrom, R., Mas-Colell, A., and Sonnenschein, H. (1976). The demand theory of the weak axiom of revealed preference. *Econometrica*, 44(5):971–978.

Kim, T. and Richter, M. K. (1986). Nontransitive-nontotal consumer theory. *Journal of Economic Theory*, 38(2):324–363.

Kitamura, Y. and Stoye, J. (2018). Nonparametric analysis of random utility models. *Econometrica*, 86(6):1883–1909.

Kőszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.

Krauss, E. (1985). A representation of arbitrary maximal monotone operators via subgradients of skew-symmetric saddle functions. *Nonlinear Analysis: Theory, Methods & Applications*, 9(12):1381–1399.

Mariotti, M. What kind of preference maximization does the weak axiom of revealed preference characterize?

Matzkin, R. L. and Richter, M. K. (1991). Testing strictly concave rationality. *Journal of Economic Theory*, 53(2):287–303.

Moore, J. (1999). *Mathematical Methods for Economic Theory 2*. Springer.

Quah, J. K.-H. (2006). Weak axiomatic demand theory. *Economic Theory*, 29(3):677–699.

Quah, J. K.-H. (2008). The existence of equilibrium when excess demand obeys the weak axiom. *Journal of Mathematical Economics*, 44(3-4):337–343.

Richter, M. and Rubinstein, A. (2019). Convex preferences: A new definition. *Theoretical Economics*, 14(4):1169–1183.

Rose, H. (1958). Consistency of preference: The two-commodity case. *Review of Economic Studies*, 25(2):124–125.

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17):61–71.

Samuelson, P. A. (1948). Consumption theory in terms of revealed preference. *Economica*, 15(60):243–253.

Shafer, W. J. (1974). The nontransitive consumer. *Econometrica*, 42(5):913–919.

Sonnenschein, H. (1971). Demand theory without transitive preferences, with applications to the theory of competitive equilibrium. In Chipman, J., Hurwich, L., and Richter, M., S.-H., editors, *Preferences, Utility and Demand*. Harcourt Brace Jovanovich, New York.

Talla Nobibon, F., Cherchye, L., Crama, Y., Demuynck, T., De Rock, B., and Spieksma, F. (2016). Revealed preference tests of collectively rational consumption behavior: Formulations and algorithms. *Operations Research*, 64(6):1197–1216.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1):31–48.

Varian, H. R. (1982). The nonparametric approach to demand analysis. *Econometrica*, 50(4):945–973.

Zangwill, W. (1967). The piecewise concave function. *Management Science*, 13(11):900–912.

Appendix

Proofs of Section 3: Characterizations of WGARP and WARP

Proof of Theorem 1

(i) $\implies$ (ii).– Let $r(x,y)$ be skew-symmetric utility function that weakly rationalizes the data. Suppose there is a violation of WGARP, so that $p^t x^t \geq p^t x^s$ and $p^s x^s > p^s x^t$ for some pair of observations $s,t \in \mathbb{T}$. Then by weak rationalization in Definition 8 we have $r(x^t, x^s) \geq 0$. Suppose first that $r(x^s, x^t) > 0$. But this results in a contradiction, since by skew-symmetry $-r(x^s, x^t) = r(x^t, x^s)$, which implies $r(x^t, x^s) \geq 0 > -r(x^s, x^t) = r(x^t, x^s)$. Suppose next that $r(x^s, x^t) = 0$. But then by local nonsatiation there exists $y \in B(x^t, \epsilon)$ for some small $\epsilon > 0$ such that $p^s x^s > p^s y$ with $r(x^s, y) < 0$, which contradicts that r weakly rationalizes the data. Thus, there cannot exist a locally nonsatiated function $r(x^s, x^t) = 0$ such that $p^s x^s > p^s x^t$.

(ii) $\implies$ (v).– Suppose that WGARP in condition (ii) holds. For every pair of observations in the data set O^T, we let O^2_{st} denote the data set consisting of the two observations $s,t \in \mathbb{T}$. Overall, we have T^2 such data sets, which exhausts all possible pairwise comparisons in O^T. For the two observations in every data set O^2_{st}, we define the Afriat function $u_{st} : X \to \mathbb{R}$ as in Afriat's theorem (See e.g., Varian 1982). From Afriat's theorem we know that u_{st} is continuous, concave and strictly increasing. Next, for all $x, y \in X$, we define the mapping: $r_{st} : X \times X \to \mathbb{R}$ as:

$$r_{st}(x,y) = \begin{cases} u_{st}(x) - u_{st}(y) & \text{if} \quad s \neq t, \\ p^t (x - y) & \text{if} \quad s = t. \end{cases}$$

Clearly, r_{st} is continuous in x and y, concave in x and convex in y (since u_{st} is continuous and concave). Moreover, it is skew-symmetric since $r_{st}(y,x) = u_{st}(y) - u_{st}(x) = -r_{st}(x,y)$. Notice that since the function r_{st} is constructed for every $(s,t)-$ pair of observations in O^T we have a collection of T^2 functions r_{st}.

Let the $T-1$ dimensional simplex be denoted as $\Delta = \{\lambda \in \mathbb{R}^T_+ | \sum_{t=1}^T \lambda_t = 1\}$. Define the preference function $r(x,y)$ for any $x, y \in X$ as:

$$\begin{aligned} r(x,y) &= \min_{\lambda \in \Delta} \max_{\mu \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x,y) \\ &= \max_{\mu \in \Delta} \min_{\lambda \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x,y). \end{aligned}$$

We prove that the function r weakly rationalizes the data set O^T. Consider $y \in X$ and some fixed $t \in \mathbb{T}$ such that $p^t x^t \geq p^t y$. Let $\mu^t \in \Delta$ be the $T-1$ simplex such that $\mu^t_j = 0$

if $j \neq t$ and $\mu_j^t = 1$ if $j = t$. Then we have:

$$\begin{aligned} r(x^t, y) &= \max_{\mu\in\Delta} \min_{\lambda\in\Delta} \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t r_{st}(x^t, y) \\ &\geq \min_{\lambda\in\Delta} \sum_{i\in\mathbb{T}} \sum_{j\in\mathbb{T}} \lambda_i \mu_j^t r_{ij}(x^t, y) \\ &= \min_{\lambda\in\Delta} \sum_{i\in\mathbb{T}} \lambda_i r_{it}(x^t, y). \end{aligned}$$

It suffices to show that $r_{it}(x^t, y) \geq 0$ whenever $p^t x^t \geq p^t y$ for each data set O_{it}^2. But this follows directly from the definition of r_{it} and Afriat's theorem. Hence, $r(x^t, y) \geq 0$.

(v) $\implies$ (vi) .– Here, we verify that the preference function r constructed in condition (v) is skew-symmetric, continuous, strictly increasing and piecewise concave (in x and also piecewise convex in y). First, we show skew-symmetry. We have:

$$\begin{aligned} -r(x, y) &= -\min_{\lambda\in\Delta} \max_{\mu\in\Delta} \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t r_{st}(x, y) \\ &= \max_{\lambda\in\Delta} \min_{\mu\in\Delta} \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t (-r_{st}(x, y)), \end{aligned}$$

Since r_{st} is skew-symmetric (i.e., $-r_{st}(x, y) = r_{st}(y, x)$), we have (using Lemma 1):

$$\begin{aligned} -r(x, y) &= \max_{\lambda\in\Delta} \min_{\mu\in\Delta} \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t (-r_{st}(x, y)) \\ &= \max_{\lambda\in\Delta} \min_{\mu\in\Delta} \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t r_{st}(y, x) \\ &= \min_{\mu\in\Delta} \max_{\lambda\in\Delta} \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t r_{st}(y, x) \\ &= r(y, x), \end{aligned}$$

which proves that r is skew-symmetric.

Second, we show that r is continuous. The simplex Δ consists of a finite number of elements and is therefore compact. Moreover, from above, we know that r_{st} is continuous. Hence, for any $\lambda, \mu \in \Delta$, the function

$$f(x, y; \lambda, \mu) = \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t r_{st}(x, y),$$

is continuous. By a direct application of Berge's maximum theorem (e.g., Moore 1999, p.280) it follows that $r(x, y) = \min_{\lambda\in\Delta} \max_{\mu\in\Delta} f(x, y; \lambda, \mu)$ is a continuous function of $x, y \in X$.

Third, we show that r is strictly increasing. Consider any $x, y, z \in X$ such that $x > y$.

Then:

$$\begin{aligned} r_{st}(x,z) &= u_{st}(x) - u_{st}(z) \\ &> u_{st}(y) - u_{st}(z) \\ &= r_{st}(y,z), \end{aligned}$$

where $u_{st}(x) > u_{st}(y)$ follows by Afriat's theorem. This implies:

$$\max_{\mu\in\Delta} \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t r_{st}(x,z) > \max_{\mu\in\Delta} \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t r_{st}(y,z),$$

for all $\lambda \in \Delta$. Thus, $r(x,z) > r(y,z)$.

Fourth, we show that r is piecewise concave in its first argument (and piecewise convex in its second argument). Consider any $x \in X$ and a fixed $z \in X$. The function $r_{st}(x,z) = u_{st}(x) - u_{st}(z)$ is concave in x since $u_{st}(x)$ is concave by Afriat's theorem and the difference between a concave function and a constant is concave. Moreover, the function $f_z(x;\lambda,\mu) = \sum_{s\in\mathbb{T}} \sum_{t\in\mathbb{T}} \lambda_s \mu_t r_{st}(x,z)$ is concave for any $\lambda,\mu \in \Delta$, since the linear combination of concave functions is concave. Since concavity is preserved under the pointwise minimum operator, it follows that the function $g_z(x;\lambda) = \min_{\mu\in\Delta} f_z(x;\lambda,\mu)$ is concave in the first argument for all $\lambda \in \Delta$. Thus, by Definition 7 the function $h_z(x) = \max_{\lambda\in\Delta} g_z(x;\lambda)$ is piecewise concave, proving that r is piecewise concave. By skew-symmetry the mapping r is piecewise convex in the second argument.

(vi) $\Longrightarrow$ (i).– Trivial.

(ii) $\Longrightarrow$ (iii).– Suppose that WGARP holds. Consider once again the data set O^2_{ts}, and recall that we have T^2 such data sets, which exhausts all possible pairwise comparisons in the data set O^T. Obviously, for the two observations in each data set O^2_{ts}, WGARP is equivalent to GARP. By a direct application of Afriat's theorem, the following conditions are equivalent: *(i)* the data set O^2_{st} satisfies WGARP, *(ii)* there exist numbers U^k_{ts} and $\lambda^k_{ts} > 0$ for all $k \in \{t,s\}$ such that the Afriat inequalities: $U^k_{ts} - U^l_{ts} \geq \lambda^k_{ts} p^k(x^k - x^l)$ hold for all $k,l \in \{t,s\}$. Now, notice that the two data sets O^2_{ts} and O^2_{st} contain the same two bundles and that permuting the data is insignificant for Afriat's theorem. Thus, without loss of generality, we can set $U^k_{ts} = U^k_{st}$ and $\lambda^k_{ts} = \lambda^k_{st}$ for all $k \in \{t,s\}$. By defining $R^{t,s} = U^t_{ts} - U^s_{ts}$ and $R^{s,t} = U^s_{ts} - U^t_{ts}$, we get the inequalities in condition (iii).

(iii) $\Longrightarrow$ (iv).– Suppose that condition (iii) holds. Since $\lambda^t_{ts} > 0$, if $p^t(x^t - x^s) \geq 0$ then $R^{t,s} \geq 0$, and if $p^t(x^t - x^s) > 0$ then $R^{t,s} > 0$. Define $W^{t,s} = R^{t,s}$ for all $s,t \in \mathbb{T}$ and the proof follows.

(iv) $\Longrightarrow$ (ii).– Suppose that the inequalities in condition (iv) holds, but that WGARP is violated, i.e., $p^t(x^t - x^s) \geq 0$ and $p^s(x^s - x^t) > 0$ for some $s,t \in \mathbb{T}$. Then $W^{t,s} \geq 0$ and $W^{s,t} > 0$. Thus, $W^{t,s} + W^{s,t} > 0$, which violates the inequalities in condition (iv).

Remarks.– The numbers $W^{t,s}$ in condition (iv) can be constructed directly from WGARP by considering the following simple proof that condition (ii) implies (iv). Suppose that WGARP holds. For all $s, t \in \mathbb{T}$ set $W^{t,s} = p^t(x^t - x^s) - p^s(x^s - x^t)$. We verify that this construction works. First, notice that $W^{s,t} = p^s(x^s - x^t) - p^t(x^t - x^s)$. Thus, $W^{t,s} + W^{s,t} = 0$ since:

$$
\begin{aligned}
W^{t,s} + W^{s,t} &= (p^t(x^t - x^s) - p^s(x^s - x^t)) + (p^s(x^s - x^t) - p^t(x^t - x^s)) \\
&= p^t(x^t - x^s) - p^t(x^t - x^s) - p^s(x^s - x^t) + p^s(x^s - x^t) \\
&= 0.
\end{aligned}
$$

Second, notice that if $p^t(x^t - x^s) \geq 0$ then $p^s(x^s - x^t) \leq 0$, otherwise we would have a violation of WGARP. Thus, $W^{t,s} = p^t(x^t - x^s) - p^s(x^s - x^t) \geq 0$. Also, if $p^t(x^t - x^s) > 0$ then $W^{t,s} = p^t(x^t - x^s) - p^s(x^s - x^t) > 0$.

Proof of Theorem 2

(i) $\Longrightarrow$ (ii).– Let $r(x, y)$ be skew-symmetric utility function that strictly rationalizes the data. Suppose there is a violation of WARP, so that $p^t x^t \geq p^t x^s$ and $p^s x^s \geq p^s x^t$ with $x^s \neq x^t$ for some pair of observations $s, t \in \mathbb{T}$. Then, by strict rationalization in Definition 8, we have $r(x^t, x^s) > 0$ and $r(x^s, x^t) > 0$. But this violates skew-symmetry.

(ii) $\Longrightarrow$ (v).– Since this proof is very similar to the proof of Theorem 1, we only give the main parts (and the parts that differ).

Suppose that WARP in condition (ii) holds. For all $s, t \in \mathbb{T}$, we let the data set O^2_{st} consist of the two observations $s, t \in \mathbb{T}$. Overall, this gives T^2 such data sets. For the two observations in each data set O^2_{st}, we define the function $u_{st} : X \to \mathbb{R}$ as in Matzkin and Richter's (1991) theorem. From this, we know that each function u_{st} is continuous, strictly concave and strictly increasing. Next, for all $x, y \in X$, we define the mapping: $r_{st} : X \times X \to \mathbb{R}$ as:

$$
r_{st}(x, y) = \begin{cases} u_{st}(x) - u_{st}(y) & \text{if} \quad s \neq t, \\ p^t(x - y) - \varepsilon(g(x - x^t) - g(y - x^t)), & \text{if} \quad s = t. \end{cases}
$$

for some small $\varepsilon > 0$ and where the function g is defined in Matzkin and Richter (1991). Clearly, each function $r_{s,t}$ is continuous, strictly concave and skew-symmetric.

Define the preference function $r(x, y)$ for any $x, y \in X$ as:

$$
\begin{aligned}
r(x, y) &= \min_{\lambda \in \Delta} \max_{\mu \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x, y) \\
&= \max_{\mu \in \Delta} \min_{\lambda \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x, y).
\end{aligned}
$$

We prove that the function r strictly rationalizes the data set O^T. Consider $y \in X$ and some fixed $t \in \mathbb{T}$ such that $x^t \neq y$ and $p^t x^t \geq p^t y$. Let $\mu^t \in \Delta$ be the $T-1$ simplex such that $\mu^t_j = 0$ if $j \neq t$ and $\mu^t_j = 1$ if $j = t$. By the same argument as in the proof of Theorem 1, we have

$$\begin{aligned} r(x^t, y) &\geq \min_{\lambda \in \Delta} \sum_{i \in \mathbb{T}} \sum_{j \in \mathbb{T}} \lambda_i \mu^t_j r_{ij}(x^t, y) \\ &= \min_{\lambda \in \Delta} \sum_{i \in \mathbb{T}} \lambda_i r_{it}(x^t, y). \end{aligned}$$

It suffices to show that $r_{it}(x^t, y) > 0$ whenever $x^t \neq y$ and $p^t x^t \geq p^t y$ for each data set O^2_{it}. But this follows directly from the definition of r_{it} and Matzkin and Richter's (1991) theorem. Hence, $r(x^t, y) > 0$.

(v) $\Longrightarrow$ (vi).– We verify that the preference function r constructed in condition (v) is skew-symmetric, continuous, strictly increasing and piecewise strictly concave in x (and piecewise strictly convex in y).

By the exact same arguments as in the proof of Theorem 1, it can be shown that the function $r(x, y)$ is skew-symmetric, continuous and strictly increasing. Thus, it suffices to show that it is piecewise strictly concave in x (and piecewise strictly convex in y). Consider any $x \in X$ and a fixed $z \in X$. The function $r_{st}(x, z) = u_{st}(x) - u_{st}(z)$ is strictly concave in x since $u_{st}(x)$ is strictly concave and $u_{st}(z)$ can be treated as a constant. Since the linear combination of strictly concave functions is strictly concave, it follows that the function $f_z(x; \lambda, \mu) = \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x, z)$ is strictly concave for any $\lambda, \mu \in \Delta$. It then follows that the function $g_z(x; \lambda) = \min_{\mu \in \Delta} f_z(x; \lambda, \mu)$ is strictly concave in the first argument for all $\lambda \in \Delta$. Hence, the function $h_z(x) = \max_{\lambda \in \Delta} g_z(x; \lambda)$ is piecewise strictly concave, proving that r is piecewise strictly concave By skew-symmetry the mapping r is piecewise strictly convex in y.

(vi) $\Longrightarrow$ (i).– Trivial.

(ii) $\Longrightarrow$ (iii).– Suppose that WARP holds. Consider the T^2 data sets O^2_{ts} for every pair of observations $s, t \in \mathbb{T}$. By a direct application of Matzkin and Richter's (1991) theorem, the following conditions are equivalent: *(i)* the data set O^2_{st} satisfies WARP, *(ii)* there exist numbers U^k_{ts} and $\lambda^k_{ts} > 0$ for all $k \in \{t, s\}$ such that the inequalities: *if* $x^k \neq x^l$ *then,* $U^k_{ts} - U^l_{ts} > \lambda^k_{ts} p^k \left(x^k - x^l\right)$, *and if* $x^k = x^l$ *then,* $U^k_{ts} - U^l_{ts} = 0$ hold for all $k, l \in \{t, s\}$. Since permuting the data is insignificant for Matzkin and Richter's (1991) theorem, we can without loss of generality set $U^k_{ts} = U^k_{st}$ and $\lambda^k_{ts} = \lambda^k_{st}$ for all $k \in \{t, s\}$. We obtain the inequalities in condition (iii) by defining $R^{t,s} = U^t_{ts} - U^s_{ts}$ and $R^{s,t} = U^s_{ts} - U^t_{ts}$.

(iii) $\Longrightarrow$ (iv).– Suppose that condition (iii) holds. If $x \neq x^t$ and $p^t(x^t - x^s) \geq 0$ then $R^{t,s} > 0$. We obtain condition (iv) by defining $W^{t,s} = R^{t,s}$ for all $s, t \in \mathbb{T}$.

(iv) $\implies$ *(ii).*– Suppose that the inequalities in condition (iv) holds, but that WARP is violated, i.e., $p^t(x^t - x^s) \geq 0$ and $p^s(x^s - x^t) \geq 0$ with $x^t \neq x^s$ for some $s, t \in \mathbb{T}$. Then $W^{t,s} > 0$ and $W^{s,t} > 0$. Thus, $W^{t,s} + W^{s,t} > 0$, which violates the inequalities in condition (iv).

Proofs of Section 4: Demand Counterfactuals.

Proof of Lemma 2

Consider the data set O^3 with prices $p^1 = (4\,1\,5)'$, $p^2 = (5\,4\,1)'$, and $p^3 = (1\,5\,4)'$, and bundles $x^1 = (4\,1\,1)'$, $x^2 = (1\,4\,1)'$, $x^3 = (1\,1\,4)'$. This data set satisfies WGARP. Notice that $p^t x^t = 22$ for all $t = 1, 2, 3$. Define the out-of-sample price: $p^{T+1} = \frac{22}{k}(p^1 + p^2 + p^3)$ for some $k \geq 60$, and the income level $w^{T+1} = p^{T+1}x^{T+1} = 22$. Then we have:

$$p^{T+1} = \frac{220}{k}(1 \quad 1 \quad 1)',$$

$$(x_1^{T+1} + x_2^{T+1} + x_3^{T+1}) = \frac{k}{10}.$$

More important, we observe that:

$$22 = p^{T+1}x^{T+1} \geq p^{T+1}x^t = \frac{22 \cdot 60}{k}.$$

Assume towards contradiction that x^{T+1} is in $D(p^{T+1}, w^{T+1})$, then it must be that, $p^t x^t < p^t x^{T+1}$ for $t = 1, 2, 3$. Adding up inequalities we obtain, $66 = (p^1x^1 + p^2x^2 + p^3x^3) < 10(x_1^{T+1} + x_2^{T+1} + x_3^{T+1}) = k$. This produces a contradiction whenever $60 \leq k < 66$ for WGARP, and WARP. There is a continuum of examples.

Proof of Theorem 3

If $x \in D^{\downarrow}(p, w)$, then there is an r that is a maximin preference function, in fact, r is also in the set of decisive maximin preference functions that rationalizes the data R. Then this means that $x \in \cup_{r\in R}\mathbf{x}_r(p, w)$. If $x \in \cup_{r\in R}\mathbf{x}_r(p, w)$, this means that there is some r for which $x \in \mathbf{x}_r(p, w)$, for this r and for the budget p, w, there is a set $V_{p,w} \subseteq \mathbb{T}$ that characterizes the decisive maximin preference model. This is true by analogous arguments to the proof of Theorem 1, because we can identify U with $\mathbb{T}$ and therefore we can identify $V_{p,w}$ with *some* subset of $\mathbb{T}$. Hence, $x \in D_{\mathbb{T}\setminus S}(p, w)$ for some $S \subseteq T$. This implies that $x \in D^{\uparrow}(p, w)$.

Proofs of Section 5: Shape Constraints: Concave Rationalization and the Law of Demand

Proof of Theorem 4

John (2001) proves the equivalence of conditions (i), (ii), and (iv). We establish the equivalence of conditions (ii) and (iii). Our proof makes use of a result in Cherchye et al. (2016), which is stated in Theorem F.

Theorem F. (Cherchye et al. 2016) Consider a finite data set $O^T = \{p^t, x^t\}_{t\in\mathbb{T}}$. The following statements are equivalent:

(1) The data O^T can be rationalized by a locally nonsatiated, continuous, strictly increasing, concave and quasihomothetic utility function.

(2) There exist numbers $V^t > 0$, a^t and $b^t > 0$ for all $t \in \mathbb{T}$ such that the inequalities:

$$V^t - V^s \geq \frac{1}{b^t} p_t(x^t - x^s),$$
$$V^t = \frac{p^t x^t - a^t}{b^t},$$

hold for all $s, t \in \mathbb{T}$, with $a^t = \delta a^s$ and $b^t = \delta b^s$ if $p^t = \delta p^s$, where $\delta > 0$.

We can now prove the following lemma:

Lemma 5. *Consider any pairwise data set O^2_{st} in $O^T = \{p^t, x^t\}_{t\in\mathbb{T}}$. The following statements are equivalent:*

(a) There exist numbers $V^t > 0$, a^t and $b^t > 0$ for all observations t in O^2_{st} such that the inequalities:

$$V^t - V^s \geq \frac{1}{b^t} p^t(x^t - x^s),$$
$$V^t = \frac{p^t x^t - a^t}{b^t},$$

hold for both observations s, t in O^2_{st}, with $a^t = \delta a^s$ and $b^t = \delta b^s$ if $p^t = \delta p^s$, where $\delta > 0$.

(b) The data O^2_{st} satisfies WGARP.

Proof. We first prove (b) $\Longrightarrow$ (a). For notational convenience, we rename the observations in O^2_{st} as 1 and 2. By Afriat's theorem, if the data set O^2_{st} satisfies WGARP, then there exist numbers $U^1, U^2, \lambda^1 > 0$ and $\lambda^2 > 0$ satisfying Afriat's inequalities.

There are two cases of interest: First, we consider $p^1 \neq \delta p^2$ for some $\delta > 0$. In that case, Cherchye et al. (2016) prove that Afriat's inequalities and the inequalities in condition (2) of Theorem F are equivalent.

Consider the second case when $p^2 = \delta p^1$ with $\delta > 0$. Using Algorithm 3 in Varian (1982), we can compute the Afriat numbers $U^1, U^2, \lambda^1, \lambda^2$ as follows: Without loss of generality we assume that:

$$p^1(x^1 - x^2) \geq 0.$$

Hence, Afriat's inequalities can be written as:

$$\begin{aligned} U^1 - U^2 &\geq \lambda^1 p^1(x^1 - x^2), \\ U^2 - U^1 &\geq \lambda^2 \delta p^1(x^2 - x^1). \end{aligned}$$

Applying Algorithm 3 in Varian (1982), we can calculate the numbers U^1 and U^2 as follows:

$$\begin{aligned} U^1 &= 1 + \min(\delta p^1(x^1 - x^2), 0) = 1, \\ U^2 &= 1 + \min(p^1(x^2 - x^1), 0) = 1 + p^1[x^2 - x^1]. \end{aligned}$$

Now, note that we must have $p^1 x^2 = \frac{p^2 x^2}{\delta}$, in which case U^2 can be computed as: $U_2 = \frac{p^2 x^2}{\delta} - p^1 x^1 + 1$. Also, notice that $U^1 - U^2 = p^1(x^1 - x^2)$ and $U^2 - U^1 = p^1(x^2 - x^1)$.

Next, applying Algorithm 3 in Varian (1982) once again to compute $\lambda^1 > 0$, we have:

$$\lambda^1 = \max(1, \frac{U^2 - U^1}{p^1(x^2 - x^1)}) = 1.$$

To compute $\lambda^2 > 0$, there are two cases. First, if $p^1(x^1 - x^2) > 0$, then:

$$\lambda^2 = \max(1, \frac{U^1 - U^2}{\delta p^1(x^1 - x^2)}) = \max(1, \frac{1}{\delta}).$$

Second, if $p^1(x^1 - x^2) = 0$, then $\lambda^2 = 1$.

Now, we define the numbers $a^1 = p^1 x^1 - 1$, $a^2 = \delta a^1 = \delta(p^1 x^1 - 1)$, $b^1 = 1$ and $b^2 = \delta b^1 = \delta$. As such, it is easy to verify that $V^1 = U^1$, $V^2 = U^2$ and $\lambda^1 = \frac{1}{b^1} = 1$, which implies that the first inequality in condition (a) of Lemma 5 holds, i.e.,

$$V^1 - V^2 \geq \frac{1}{b^1} p^1(x^1 - x^2).$$

To verify that the second inequality holds, simply note that $V^2 - V^1 = U^2 - U^1 = p^1(x^2 - x^1) = \frac{1}{b^2} \delta p^1(x^2 - x^1)$, which implies:

$$V^2 - V^1 \geq \frac{1}{b^2} p^2(x^2 - x^1),$$

since $p^2 = \delta p^1$. Hence, condition (b) imply (a).

Next, we prove (a) $\Longrightarrow$ (b). Define $\lambda_{12}^1 = \frac{1}{b^1}, \lambda_{12}^2 = \frac{1}{b^2}, R^{21} = V^2 - V^1$ and $R^{12} = V^1 - V^2$, in which case we have $R^{12} = -R^{21}$, and moreover,

$$R^{12} \geq \lambda_{12}^1 p^1 (x^1 - x^2),$$
$$R^{21} \geq \lambda_{12}^2 p^2 (x^2 - x^1),$$

which implies by Theorem 1 that WGARP holds. ■

Now we are ready to prove Theorem 4.

(iii) $\Longrightarrow$ *(ii).–* Consider the number:

$$r(x^t, x^s) = \max_{\lambda \in \Delta(U)} \min_{\mu \in \Delta(U)} \sum_{i \in U} \sum_{j \in U} \lambda_i \mu_j (u_{ij}(x^t) - u_{ij}(x^s)),$$

where u_{ij} is a continuous, strictly increasing and concave utility function that admits a Gorman indirect utility function, $v_{ij}(p, w)$, with reference-independent marginal utility of wealth, i.e., $v_{ij}(p, w) = \frac{w}{b(p)} - \frac{a_{ij}(p)}{b(p)}$, such that $a_{ij} = a_{ji}$ for all $i, j \in U$. We have $u_{ij}(x^t) = v_{ij}(p^t, p^t x^t) = \frac{p^t x^t}{b(p^t)} - \frac{a_{ij}(p^t)}{b(p^t)}$. Then, any λ^* and μ^* that solves the maximin optimization problem produces an aggregate of Gorman utility functions that is a Gorman utility function as well, such that:

$$r(x^t, x^s) = \sum_{i \in U} \sum_{j \in U} \lambda_i^* \mu_j^* (u_{i^* j^*}(x^t) - u_{i^* j^*}(x^s)) = \overline{u}(x^t) - \overline{u}(x^s) \geq \frac{1}{b(p^t)} p^t (x^t - x^s),$$

for all $t, s \in \mathbb{T}$, where $\overline{u}$ is the Gorman utility function, which follows by Theorem 2 in Cherchye et al. (2016). Let $R^{st} = r(x^t, x^s)$ and $\lambda^t = \frac{1}{b(p^t)}$, in which case condition (ii) holds.

(ii) $\Longrightarrow$ *(iii).–* First, we break O^T into T^2 pairwise sub data sets O_{st}^2. Without loss of generality, let $U = \mathbb{T}$. For any O_{st}^2, condition (ii) implies that WGARP holds. By directly applying Lemma 5 and Theorem 2 in Cherchye et al. (2016), we can construct local utility functions u_{ts} that weakly rationalizes O_{st}^2 and are quasihomothetic (i.e., consistent with the Gorman polar form) given by:

$$v_{ts}(p, w) = \frac{w}{b_{ts}(p)} - \frac{a_{ts}(p)}{b_{ts}(p)},$$

where $a_{ts}(p) = a_{st}(p)$, and a and b are both homogeneous of degree 1.

Condition (ii) implies by Theorem 3 in Cherchye et al. (2016) that we can always set $b_{ts}(p) = b(p)$ for all $t, s \in \mathbb{T}$. Specifically, this holds because $\frac{1}{b_{ts}(p^t)} = \lambda^t$ for all $t \in \mathbb{T}$. Then, we can construct a quasihomothetic preference function, r, with a reference-independent marginal utility of wealth, i.e., $b_{ts}(p) = b(p)$, as:

$$r(x, y) = \max_{\lambda \in \Delta(\mathbb{T})} \min_{\mu \in \Delta(\mathbb{T})} \sum_{t \in \mathbb{T}} \sum_{s \in \mathbb{T}} \lambda_t \mu_s (u_{ts}(x) - u_{ts}(y)),$$

where each local utility function u_{ts} is continuous, strictly increasing, concave and quasi-homothetic such that it admits a Gorman polar form with reference-independent marginal utility of wealth, i.e., $v_{ts}(p, w) = \frac{w}{b(p)} - \frac{a_{ts}(p)}{b(p)}$, where $a_{ts} = a_{st}$ for all $t, s \in \mathbb{T}$.

Finally, since each local utility function, u_{st}, rationalizes O^2_{st}, we have by the same arguments as in Theorem 1 that r weakly rationalizes the data set O^T.

Proof of Theorem E

(i) $\implies$ *(ii).–* By the definition of quasilinear rationalization, we have for any observation $s \in \mathbb{T}$ with $x = x^s$,

$$u(x^t) - p^t x^t \geq u(x^s) - p^t x^s.$$

Thus, after rearranging terms, for any sequence of distinct choices of indices $(1, 2, 3, ..., n) \in \mathbb{T}$, we have:

$$\begin{aligned} p^1x^2 - p^1x^1 &\geq u(x^2) - u(x^1), \\ p^2x^3 - p^2x^2 &\geq u(x^3) - u(x^2), \\ &\vdots \\ p^nx^1 - p^nx^n &\geq u(x^1) - u(x^n). \end{aligned}$$

Adding up both sides, we get:

$$\begin{aligned} & (p^1x^2 - p^1x^1) + (p^2x^3 - p^2x^2) + \cdots + (p^nx^1 - p^nx^n) \\ \geq\ & (u(x^2) - u(x^1)) + (u(x^3) - u(x^2)) + \cdots + (u(x^1) - u(x^n)) \\ =\ & 0. \end{aligned}$$

Thus,

$$p^1(x^1 - x^2) + p^2(x^2 - x^3) + \cdots + p^n(x^n - x^1) \leq 0,$$

which is cyclical monotonicity.

(ii) $\Longrightarrow$ *(iii).–* Suppose that condition (ii) holds and define:

$$U^t = \min_{\{1,2,3,...,n,t\}\in\mathbb{T}} \{p^1(x^2 - x^1) + p^2(x^3 - x^2) + \cdots + p^n(x^t - x^n)\},$$

for all $t \in \mathbb{T}$. That is, U^t is a minimum of the given expression over all sequences starting anywhere and terminating at t. Note that there are only finitely many sequences because their elements are distinct. Hence, the minimum always exists. To show that the numbers U^t do satisfy the inequalities in statement (iii), suppose that:

$$\begin{aligned} U^t &= p^1(x^2 - x^1) + p^2(x^3 - x^2) + \cdots + p^n(x^t - x^n), \\ U^s &= p^a(x^b - x^a) + p^b(x^c - x^b) + \cdots + p^m(x^s - x^m), \end{aligned}$$

for some distinct sequences $\{1, 2, 3, ...n, t\} \in \mathbb{T}$ and $\{a, b, c, ..., m, s\} \in \mathbb{T}$. Then:

$$\begin{aligned} U^t &= p^1(x^2 - x^1) + p^2(x^3 - x^2) + \cdots + p^n(x^t - x^n) \\ &\leq p^a(x^b - x^a) + p^b(x^c - x^b) + \cdots + p^m(x^s - x^m) + p^s(x^t - x^s) \\ &= U^s + p^s(x^t - x^s), \end{aligned}$$

since the value on the left-hand side of the inequality is a minimum over all paths to t. Hence,

$$U^t \leq U^s + p^s(x^t - x^s),$$

for all $s, t \in \mathbb{T}$, which are the inequalities in statement (iii).

(iii) $\Longrightarrow$ *(iv).–* Suppose that condition (iii) holds. For all $x \in X$, define the function:

$$u(x) = \min_{s\in\mathbb{T}} \{U^s + p^s(x - x^s)\}$$

Since u is defined as the lower envelope of a set of linear functions, it is continuous, strictly increasing and concave. Moreover, it is easy to show that $u(x^t) = U^t$ for all $t \in \mathbb{T}$. Finally, for all $x \in X$ and all $t \in \mathbb{T}$:

$$\begin{aligned} u(x) - p^t x &= \min_{s\in\mathbb{T}} \{U^s + p^s(x - x^s)\} - p^t x \\ &\leq U^t + p^t(x - x^t) - p^t x \\ &= U^t - p^t x^t \\ &= u(x^t) - p^t x^t. \end{aligned}$$

Thus, u rationalizes the data set O^T.

(iv) $\Longrightarrow$ *(i).–* Trivial.

Proof of Theorem 5

(i) $\Longrightarrow$ *(ii).–* If the data set O^T can be rationalized by a skew-symmetric and quasilinear preference function, then for all $t \in \mathbb{T}$ and all $y \in X$,

$$r(x^t, y) \geq p^t(x^t - y).$$

In particular, it must be that for $y = x^s$, $r(x^t, x^s) \geq p^t(x^t - x^s)$. Analogously, we have $r(x^s, x^t) \geq p^s(x^s - x^t)$ for all $s, t \in \mathbb{T}$. Adding these inequalities, and by skew-symmetry, we have:

$$0 = r(x^t, x^s) + r(x^s, x^t) \geq p^t(x^t - x^s) + p^s(x^s - x^t).$$

Rearranging terms, we get:

$$(p^t - p^s)(x^t - x^s) \leq 0,$$

for all $s, t \in \mathbb{T}$, which is the law of demand.

(ii) $\Longrightarrow$ *(iii).–* Assume that condition (ii) holds and define:

$$R^{s,t} = \frac{1}{2}(p^s(x^s - x^t) - p^t(x^t - x^s)).$$

Clearly, $R^{s,t} = -R^{t,s}$ for all $s, t \in \mathbb{T}$. Moreover,

$$\begin{aligned} R^{s,t} &= \frac{1}{2}(p^s(x^s - x^t) - p^t(x^t - x^s)) \\ &= \frac{1}{2}(p^s(x^s - x^t) + p^t(x^s - x^t)) \\ &= \frac{1}{2}(p^s(x^s - x^t) - p^t(x^s - x^t) + 2p^t(x^s - x^t)). \end{aligned}$$

By condition (ii), we have

$$p^s(x^s - x^t) - p^t(x^s - x^t) = (p^s - p^t)(x^s - x^t) \leq 0.$$

Hence,

$$\begin{aligned} R^{s,t} &= \frac{1}{2}(p^s(x^s - x^t) - p^t(x^s - x^t) + 2p^t(x^s - x^t) \\ &= \frac{1}{2}(p^s(x^s - x^t) - p^t(x^s - x^t)) + p^t(x^s - x^t) \Longleftrightarrow \\ R^{s,t} - p^t(x^s - x^t) &= \frac{1}{2}(p^s(x^s - x^t) - p^t(x^s - x^t)) \\ &\leq 0, \end{aligned}$$

implying:

$$
\begin{aligned}
-R^{s,t} &\geq -p^t\left(x^s - x^t\right) \Longleftrightarrow \\
R^{t,s} &\geq p^t\left(x^t - x^s\right),
\end{aligned}
$$

which are the inequalities in statement (iii).

(iii) $\Longrightarrow$ *(iv).–* As in the proofs of Theorems 1 and 2, we break O^T into T^2 pairwise data sets $O^2_{st} = \{(p^t, x^t), (p^s, x^s)\}$ for all $s, t \in \mathbb{T}$. For the two observations in every data set O^2_{st}, we define the function $u_{st}(x) : X \to \mathbb{R}$ as in the proof of Theorem E. From Theorem E, we know that u_{st} is continuous, strictly increasing, concave, quasilinear and rationalizes the data O^2_{st}.

For all $x, y \in X$, we define the mapping: $r_{st} : X \times X \to \mathbb{R}$ as:

$$
r_{st}(x, y) = \begin{cases} u_{st}(x) - u_{st}(y) & \text{if} \quad s \neq t, \\ p^t(x - y) & \text{if} \quad s = t. \end{cases}
$$

Next, we define the maximin quasilinear preference function, r, for any $x, y \in X$, as:

$$
\begin{aligned}
r(x, y) &= \min_{\lambda \in \Delta} \max_{\mu \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x, y) \\
&= \max_{\mu \in \Delta} \min_{\lambda \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x, y).
\end{aligned}
$$

We show that the maximin quasilinear preference function rationalizes the data set O^T. Consider $y \in X$ and some fixed $t \in \mathbb{T}$. Let $\mu^t \in \Delta$ be the $T - 1$ simplex such that $\mu^t_j = 0$ if $j \neq t$ and $\mu^t_j = 1$ if $j = t$. We have:

$$
\begin{aligned}
r(x^t, y) &= \max_{\mu \in \Delta} \min_{\lambda \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x^t, y) \\
&\geq \min_{\lambda \in \Delta} \sum_{i \in \mathbb{T}} \sum_{j \in \mathbb{T}} \lambda_i \mu^t_j r_{ij}(x^t, y) \\
&= \min_{\lambda \in \Delta} \sum_{i \in \mathbb{T}} \lambda_i r_{it}(x^t, y).
\end{aligned}
$$

It suffices to show that $r_{it}(x^t, y) \geq p^t x^t - p^t y$, for each data set O^2_{it}. But this follows directly from the definition of r_{it} and weak rationalization. Hence, $r(x^t, y) \geq p^t x^t - p^t y$ for all $y \in X$ and all $t \in \mathbb{T}$.

(iv) $\Longrightarrow$ *(i).–* Using the same arguments as in the proof of Theorem 1 it follows that the maximin quasilinear preference function r constructed above is continuous, skew-symmetric, strictly increasing (in x), piecewise concave in x, piecewise convex in y and quasilinear.

(iii) $\Longrightarrow$ *(v).–* Suppose that condition (iii) holds and define for all $x, y \in X$ the functions:

$$r_{st}(x,y) = R^{s,t} + p^s(x - x^s) - p^t\left(y - x^t\right).$$

Clearly, the function r_{st} is continuous, strictly increasing and concave in x and convex in y. Since $R^{s,t} = -R^{t,s}$, we have:

$$\begin{aligned} -r_{st}(x,y) &= -\left(R^{s,t} + p^s(x - x^s) - p^t\left(y - x^t\right)\right) \\ &= R^{t,s} + p^t\left(y - x^t\right) - p^s(x - x^s) \\ &= r_{ts}(y,x). \end{aligned}$$

Let the $T-1$ dimensional simplex be denoted $\Delta = \{\lambda \in \mathbb{R}^T_+ | \sum_{t=1}^T \lambda_t = 1\}$. Define the preference function $r(x,y)$ for any $x, y \in X$ as:

$$\begin{aligned} r(x,y) &= \min_{\lambda \in \Delta} \max_{\mu \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x,y) \\ &= \max_{\mu \in \Delta} \min_{\lambda \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x,y). \end{aligned}$$

We show that the function r is skew-symmetric, continuous, strictly increasing and concave. First, we show skew-symmetry:

$$\begin{aligned} -r(x,y) &= -\min_{\lambda \in \Delta} \max_{\mu \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x,y) \\ &= \max_{\lambda \in \Delta} \min_{\mu \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} -\lambda_s \mu_t r_{st}(x,y) \\ &= \max_{\lambda \in \Delta} \min_{\mu \in \Delta} \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} -\lambda_s \mu_t r_{ts}(y,x) \\ &= r(y,x), \end{aligned}$$

since $-r_{st}(x,y) = r_{ts}(y,x)$.

Second, we show that r is continuous. The simplex Δ consists of a finite number of elements and is therefore compact. Moreover, from above, we know that r_{st} is continuous. Hence, for any $\lambda, \mu \in \Delta$, the function

$$f(x,y;\lambda,\mu) = \sum_{s \in \mathbb{T}} \sum_{t \in \mathbb{T}} \lambda_s \mu_t r_{st}(x,y),$$

is continuous. By a direct application of Berge's maximum theorem it follows that $r(x,y) = \min_{\lambda \in \Delta} \max_{\mu \in \Delta} f(x,y;\lambda,\mu)$ is a continuous function of $x, y \in X$.

Third, we show that r is strictly increasing. Consider $x, y, z \in X$ such that $x > y$.

Since each function r_{st} is strictly increasing we have:

$$\max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\lambda_t\mu_s r_{st}(x,z) > \max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\lambda_t\mu_s r_{st}(y,z),$$

for all Δ. Hence,

$$\min_{\lambda\in\Delta}\max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\lambda_t\mu_s r_{st}(x,z) > \min_{\lambda\in\Delta}\max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\lambda_t\mu_s r_{st}(y,z),$$

which shows that r is strictly increasing in the first argument x.

Fourth, we will show that $r(x,y)$ is concave in x. Fix y and $\lambda\in\Delta$, and consider the function:

$$r_\lambda(x) = \max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\lambda_s\mu_t r_{st}(x,y).$$

We have:

$$\begin{aligned}
r_\lambda(x) &= \max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\lambda_s\mu_t r_{st}(x,y)\\
&= \max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\lambda_s\mu_t\left(R^{s,t}+p^s(x-x^s)-p^t\left(y-x^t\right)\right)\\
&= \max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\lambda_s\left(\sum_{t\in\mathbb{T}}\left(\mu_t R^{s,t}+\mu_t p^s(x-x^s)-\mu_t p^t\left(y-x^t\right)\right)\right)\\
&= \max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\lambda_s\left(p^s(x-x^s)+\sum_{t\in\mathbb{T}}\left(\mu_t R^{s,t}-\mu_t p^t\left(y-x^t\right)\right)\right)\\
&= \sum_{s\in\mathbb{T}}\lambda_s p^s(x-x^s)+\max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\left(\mu_t R^{s,t}-\mu_t p^t\left(y-x^t\right)\right).
\end{aligned}$$

Clearly, $r_\lambda(x)$ is linear in x and, as such, concave. Hence, $r(x,y)=\min_{\lambda\in\Delta} r_\lambda(x)$ is the minimum over a set of linear function and is therefore also concave.

Finally, we show that r is a quasilinear preference function that rationalizes the data. For all $y\in X$ and all $t\in\mathbb{T}$:

$$\begin{aligned}
r\left(x^t,y\right) &= \min_{\lambda\in\Delta}\max_{\mu\in\Delta}\sum_{s\in\mathbb{T}}\sum_{t\in\mathbb{T}}\lambda_s\mu_t r_{st}\left(x^t,y\right)\\
&\geq \min_{\lambda\in\Delta}\sum_{s\in\mathbb{T}}\sum_{v\in\mathbb{T}}\lambda_s\mu_v^t r_{sv}\left(x^t,y\right)\\
&= \min_{\lambda\in\Delta}\sum_{t\in\mathbb{T}}\lambda_s r_{st}\left(x^t,y\right),
\end{aligned}$$

where $\mu_v^t=1$ when $v=t$ and zero otherwise. Note that the term $p^t(y-x^t)$ does not depend on s, which implies:

$$
\begin{aligned}
r(x^t, y) &\geq \min_{\lambda\in\Delta}\sum_{s\in\mathbb{T}}\lambda_s r_{st}\left(x^t, y\right) \\
&= \min_{\lambda\in\Delta}\sum_{s\in\mathbb{T}}\lambda_s\left(R^{s,t} + p^s\left(x^t - x^s\right) - p^t\left(y - x^t\right)\right) \\
&= -\sum_{s\in\mathbb{T}}\lambda_s p^t\left(y - x^t\right) + \min_{\lambda\in\Delta}\sum_{s\in\mathbb{T}}\lambda_s\left(R^{s,t} + p^s\left(x^t - x^s\right)\right) \\
&= -p^t\left(y - x^t\right) + \min_{\lambda\in\Delta}\sum_{s\in\mathbb{T}}\lambda_s\left(R^{s,t} + p^s\left(x^t - x^s\right)\right).
\end{aligned}
$$

Thus, r is a quasilinear preference function that rationalizes the data O^T since:

$$
\begin{aligned}
r\left(x^t, y\right) - p^t\left(x^t - y\right) &= \min_{\lambda\in\Delta}\sum_{s\in\mathbb{T}}\lambda_s\left(R^{s,t} + p^s\left(x^t - x^s\right)\right) \\
&\geq 0,
\end{aligned}
$$

because $R^{s,t} + p^s\left(x^t - x^s\right) \geq 0$ by condition (iii) and $\lambda_s \geq 0$ for all $s, t \in \mathbb{T}$.

(v) $\Longrightarrow$ *(i)*.– Trivial.

Proofs of Section 6: Recoverability of Preferences

Proof of Theorem 6

The proof of the first part of the theorem follows from the discussion in Section 6.2. We prove the second part. Since $RW_{st}(x) \supseteq RW(x)$, we have that $NRW_{st}(x) \subseteq NRW(x)$. Thus, by construction $NRW^W(x) \subseteq NRW(x)$. We are going to show that $NRW^w(x) \subset NRW(x)$ for some $x \in X$. We will do this in the context if Example 6. Clearly, the bundle $x^- = (1\,1\,1)'$ is monotonically dominated by x^{T+1}. First, note that the upper bound using Varian's method contains this dominated option, i.e., $x^- \in NRW(x) = X \setminus x^{T+1}$. Second, note that for all $s, t \in \mathbb{T}$ and by strict monotonicity, we must have $u_{st}(x^{T+1}) > u_{st}(x^-)$ (this follows by Afriat's theorem applied to the data O^2_{st}). This implies that $x^- \notin NRW_{st}(x^{T+1})$ for all $t, s \in \mathbb{T}$. It also follows that $RP(w) \subseteq RP^W(x)$, since $RP(x) \subseteq RP_{st}(x)$ holds for all $s, t \in \mathbb{T}$.

Consider again, in the context of Example 6, the bundle $x^- = (1\,1\,1)'$. We are going to show that $x^{T+1} = (2\,2\,2)'$ is not in $RP(x^-)$, but that it is in $RP^W(x^-)$. From Example 6, we know that $S(x^{T+1}) = \emptyset$, which implies $x^{T+1} \notin RP(x^-)$. However, we also have that $p(x^{T+1} - x^-) > 0$ for all $p \in P$, which means that for the rationalizing local utility function, we have $u_{st}(x^{T+1}) > u_{st}(x^-)$ for all $s, t \in \mathbb{T}$. This means that $x^{T+1} \in RP_{st}(x^-)$ for all $s, t \in \mathbb{T}$. Hence, $x^{T+1} \in RP^W(x^-)$.

Proofs of Section 7: Infinite Data Sets: Characterizations of WGARP and WARP.

In order to prove Theorem 7, we need an auxiliary lemma, which is a modification of Algorithm 3 in Varian (1982).

Lemma 6. *Consider a finite data set O^T, and suppose that O^T satisfies SARP. Then:*

(i) There exist numbers U^t and $\lambda^t > 0$ for all $t \in \mathbb{T}$, such that the inequalities:

$$U^t - U^s \geq \lambda^t p^t(x^t - x^s),$$

hold for all $s, t \in \mathbb{T}$, with a strict inequality when $x^t \neq x^s$.

(ii) There exists a continuous function that maps the data set O^T to the numbers U^t, λ^t for all $t \in \mathbb{T}$.

Proof. We begin with (i). Let $g : \mathbb{R}^L \to \mathbb{R}_{++}$ be any continuous function such that $g(x) = 0$ if and only if $x = 0$, and let $\epsilon > 0$ be a scalar. For any subset I of $\mathbb{T}$, let $\max I$ denote the index of a maximum element of O^T relative to the revealed preference order. Consider the following algorithm:

Input: A set of price-quantity observations O^T satisfying SARP.
Output: Numbers U^t and $\lambda^t > 0$ for all $t \in \mathbb{T}$ satisfying the inequalities in statement (i).

1. Let $I = \{1, ..., n\}$ and $B \neq \emptyset$.

2. Let $m = \max(I)$.

3. Set $E = \{i \in I : x^i \succeq^R x^m\}$. If $B = \emptyset$, then set $U^m = \lambda^m = 1$ and go to step 6; else go to step 4.

4. Set $U^m = \min_{i \in E} \min_{j \in B} \min\{U^j + \lambda^j p^j(x^i - x^j) - \epsilon g(x^i - x^j), U^j\}$.

5. Set $\lambda^m = \max_{i \in E} \max_{j \in B} \max\{(U^j - U^m + \epsilon g(x^j - x^m))/p^m(x^i - x^m), 1\}$.

6. Set $I = I \setminus E$ and $B = B \cup E$. If $I \neq \emptyset$, then stop; otherwise, go to step 2.

We now prove that this algorithm generates numbers U^t and $\lambda^t > 0$ for all $t \in \mathbb{T}$ that satisfies the inequalities in statement (i). From step 4, we have:

$$U^m \leq U^j + \lambda^j p^j(x^m - x^j) - \epsilon g(x^m - x^j),$$

and

$$\lambda^m \geq (U^j - U^m + \epsilon g(x^m - x^j))/p^m(x^j - x^m),$$

for all $m, j \in \mathbb{T}$. This implies:

$$U^j - U^m \leq \lambda^m p^m (x^j - x^m) - \epsilon g(x^m - x^j),$$

for all $m, j \in \mathbb{T}$. Hence, this shows that the algorithm guarantees that there exist numbers U^t and $\lambda^t > 0$ for all $t \in \mathbb{T}$ satisfying the inequalities in statement (i).

Moreover, it is clear that this algorithm provides a continuous function that maps the data set O^T to the numbers U^t, λ^t for all $t \in \mathbb{T}$, which proves statement (ii). ■

Proof of Theorem 7

(i) $\Longrightarrow$ *(ii).–* Trivial.

(ii) $\Longrightarrow$ *(iii).–* Suppose that WGARP in statement (ii) holds. For every pair of observations in the data set O^∞, we let $O^2 = \{(p, a), (q, b)\}$ denote the data set consisting of any pair $(p, a), (q, b) \in O^\infty$. Overall, we have a continuum of such data sets, which exhausts all possible pairwise comparisons in O^∞. We rearrange O^2 into a vector $o = (p' \, a' \, q' \, b')' \in \mathbf{O}^2$, where the set of reference points is defined to be $\mathbf{O} = P \times \mathbf{X}$ with $x \in \mathbf{x}(p, px)$, such that each data set O^2 can be thought of as a column vector (Recall that under our assumptions $\mathbf{O}$ is a compact and metric space). We define $o_1 = (p' \, a')'$ and $o_2 = (q' \, b')'$, such that $o = (o_1' \, o_2')'$.

Since every data set o satisfies WGARP, we can directly apply Algorithm 3 in Varian (1982), which specifies a continuous function that maps finite data sets to the numbers U_{o_i} and $\lambda_{o_i} > 0$ for $i \in \{1, 2\}$ satisfying the Afriat inequalities for every data set o. We can then use these numbers to define the utility function $u_{o_1 o_2}(x) = \min_{i \in \{1,2\}} \{U_{o_i} + \lambda_{o_i} p^i (x - x^i)\}$ as in Afriat's theorem that is continuous for all $x \in X$, strictly increasing and concave on o (In the case of strict rationalizability, we simply apply Lemma 6, and then define the function $u_{o_1 o_2}$ as in Matzkin and Richter's (1991) theorem. As such, we know that $u_{o_1 o_2}$ is continuous, strictly concave and strictly increasing).[22]

Next, we define the mapping: $r_{o_1 o_2} : X \times X \to \mathbb{R}$ as:

$$r_{o_1 o_2}(x, y) = u_{o_1 o_2}(x) - u_{o_1 o_2}(y),$$

for all $x, y \in X$. Clearly, $r_{o_1 o_2}$ is continuous (since $u_{o_1 o_2}$ is continuous) and (strictly) concave in the first argument (since $u_{o_1 o_2}$ is (strictly) concave). Moreover, it is skew-symmetric since $r_{o_1 o_2}(y, x) = u_{o_1 o_2}(y) - u_{o_1 o_2}(x) = -r_{o_1 o_2}(x, y)$. Notice that since $r_{o_1 o_2}$ is constructed for every $o-$vector of observations in O^∞ we have an infinite collection of

[22] If $o_1 = o_2$ then we set $u_{o_1 o_2} = p(x - y)$ in the weak case and $u_{o_1 o_2} = p(x - y) - \epsilon(g(x - a) - g(y - b))$ in the strict case, for some small scalar $\epsilon > 0$ and a function g defined in Matzkin and Richter's (1991) theorem.

functions $r_{o_1o_2}$.

Define the preference function $r(x,y)$ for any $x,y \in X$ as:

$$\begin{aligned} r(x,y) &= \min_{\lambda\in\Delta(\mathbf{O})}\max_{\mu\in\Delta(\mathbf{O})}\int_{o_1\in\mathbf{O}}\int_{o_2\in\mathbf{O}} r_{o_1o_2}(x,y)d\lambda(o_1)d\mu(o_2) \\ &= \max_{\mu\in\Delta(\mathbf{O})}\min_{\lambda\in\Delta(\mathbf{O})}\int_{o_1\in\mathbf{O}}\int_{o_2\in\mathbf{O}} r_{o_1o_2}(x,y)d\lambda(o_1)d\mu(o_2), \end{aligned}$$

where the second equality follows from Lemma 4.[23]

We prove that the function r weakly (strictly) rationalizes the data set O^∞. Consider $y \in X$ and some fixed $\bar{o}_1 = (p'\ x')' \in \mathbf{O}$ such that $px \geq py$. Let $\mu^{\bar{o}_1} \in \Delta(\mathbf{O})$ be the probability measure such that $\mu^{\bar{o}_1}(q,b) = 0$ if $(q'\,b)' \neq \bar{o}_1$ and $\mu^{\bar{o}_1}(q,b) = 1$ when $(q'\,b)' = \bar{o}_1$. We have:

$$\begin{aligned} r(x,y) &= \max_{\mu\in\Delta(\mathbf{O})}\min_{\lambda\in\Delta(\mathbf{O})}\int_{o_1\in\mathbf{O}}\int_{o_2\in\mathbf{O}} r_{o_1o_2}(x,y)d\lambda(o_1)d\mu(o_2) \\ &\geq \min_{\lambda\in\Delta(\mathbf{O})}\int_{o_1\in\mathbf{O}}\int_{o_2\in\mathbf{O}} r_{o_1o_2}(x,y)d\lambda(o_1)d\mu^{\bar{o}_1}(o_2) \\ &= \min_{\lambda\in\Delta(\mathbf{O})}\int_{o_1\in\mathbf{O}} r_{o_1\bar{o}_1}(x,y)d\lambda(o_1). \end{aligned}$$

It suffices to show that $r_{o_1\overline{o_1}}(x,y)(>) \geq 0$ whenever $px \geq py$, for each data set $o = (o_1'\,\bar{o}_1)'$. But this follows directly from the definition of $r_{o_1\bar{o}_1}$ and Afriat's theorem (In the case of strict rationalizability, it follows from Matzkin and Richter's (1991) theorem). Specifically, notice that if $px \geq py$ then $x \succeq^{R,D} y$ for $\bar{o}_1 = (p'\ x)'$ since, in such case, $u_{o_1\bar{o}_1}(x)(>) \geq u_{o_1\bar{o}_1}(y)$. Hence, $r(x,y)(>) \geq 0$.

(iii) $\implies$ *(iv)*.– We verify that the preference function r is skew-symmetric, continuous, strictly increasing and (strictly) piecewise concave in x (and (strictly) piecewise convex in y). First, we show skew-symmetry. We have:

$$\begin{aligned} -r(x,y) &= -\min_{\lambda\in\Delta(\mathbf{O})}\max_{\mu\in\Delta(\mathbf{O})}\int_{o_1\in\mathbf{O}}\int_{o_2\in\mathbf{O}} r_{o_1o_2}(x,y)d\lambda(o_1)d\mu(o_2) \\ &= \max_{\lambda\in\Delta(\mathbf{O})}\min_{\mu\in\Delta(\mathbf{O})}\int_{o_1\in\mathbf{O}}\int_{o_2\in\mathbf{O}} (-r_{o_1o_2}(x,y))d\lambda(o_1)d\mu(o_2). \end{aligned}$$

Since $r_{o_1o_2}$ is skew-symmetric (i.e., $-r_{o_1o_2}(x,y) = r_{o_1o_2}(y,x)$), we have $-r(x,y) = r(y,x)$, which proves that r is skew-symmetric.

Second, we show that r is continuous. The simplex $\Delta(\mathbf{O})$ is a compact set and since $r_{o_1o_2}$ is defined as the difference between two continuous functions (by Afriat's theorem),

[23] The minimax theorem in Glicksberg (1950) requires that $\Delta(\mathbf{O})$ is a compact metric space, and that $r_\bullet$ is a continuous function. We have already shown that this holds by construction, so we can directly apply this version of the minimax theorem.

it is continuous itself. Thus, for any $\lambda, \mu \in \Delta(\mathbf{O})$, the function

$$f(x, y; \lambda, \mu) = \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, y) d\lambda(o_1) d\mu(o_2),$$

is continuous. By a direct application of Berge's maximum theorem it follows that $r(x, y) = \min_{\lambda \in \Delta(\mathbf{O})} \max_{\mu \in \Delta(\mathbf{O})} f(x, y; \lambda, \mu)$ is a continuous function of $x, y \in X$.

Third, we show that r is strictly increasing. Consider any $x, y, z \in X$ such that $x > y$. Then:

$$\begin{aligned} r_{o_1 o_2}(x, z) &= u_{o_1 o_2}(x) - u_{o_1 o_2}(z) \\ &> u_{o_1 o_2}(y) - u_{o_1 o_2}(z) \\ &= r_{o_1 o_2}(y, z), \end{aligned}$$

where $u_{o_1 o_2}(x) > u_{o_1 o_2}(y)$ follows by Afriat's theorem. Hence, $r(x, z) > r(y, z)$.

In the case of strict rationalizability, it follows that r satisfies skew-symmetry, continuity and strict monotonicity simply by replacing Afriat's theorem with Matzkin and Richter's (1991) theorem in the proofs above.

Finally, we show that r is (strictly) piecewise concave in x (and (strictly) piecewise convex in y). Consider any $x \in X$ and a fixed $z \in X$. The function $r_{o_1 o_2}(x, z) = u_{o_1 o_2}(x) - u_{o_1 o_2}(z)$ is (strictly) concave in x since $u_{o_1 o_2}(x)$ is (strictly) concave by Afriat's theorem (Matzkin and Richter's (1991) theorem) and the difference between a (strictly) concave function and a constant is (strictly) concave. Moreover, the function $f_z(x; \lambda, \mu) = \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, z) d\lambda(o_1) d\mu(o_2)$ is (strictly) concave for any $\lambda, \mu \in \Delta$, since the linear combination of (strictly) concave functions is (strictly) concave. It then follows that the function $g_z(x; \lambda) = \min_{\mu \in \Delta(\mathbf{O})} f_z(x; \lambda, \mu)$ is (strictly) concave in the first argument for all $\lambda \in \Delta(\mathbf{O})$. Thus, by definition the function $h_z(x) = \max_{\lambda \in \Delta(\mathbf{O})} g_z(x; \lambda)$ is (strictly) piecewise concave, proving that r is piecewise (strictly) concave. By skew-symmetry the mapping r is (strictly) piecewise convex in the second argument.

(iv) $\Longrightarrow$ *(i)*.– Trivial.

Proof of Theorem 8

(i) $\Longrightarrow$ *(ii)*.– Suppose that the data O^∞ satisfies statement (i), in which case, for all $x, y \in X$:

$$r(x, y) \geq p(x - y),$$

and

$$r(y, x) \geq q(y - x),$$

such that $x \in \mathbf{x}(p, px)$ and $y \in \mathbf{x}(q, qy)$. Adding these inequalities, and by skew-symmetry, we have:

$$0 = r(x, y) + r(y, x) \geq p(x - y) + q(y - x).$$

Rearranging terms, we get:

$$(p - q)(x - y) \leq 0,$$

which is the law of demand.

(ii) $\implies$ (iii).– Suppose that statement (ii) holds. Define $o_1 = (p' \; a')'$ and $o_2 = (q' \; b')'$, such that $o = (o_1' \; o_2')'$ as in the proof of Theorem 7. Since the law of demand holds by definition, we can use the algorithm in the proof of Theorem E to find numbers that satisfies the inequalities in statement (iii) in Theorem E. As shown in the proof of Theorem E, these numbers can then be used to construct a continuous, strictly increasing, concave and quasilinear utility function $u_{o_1 o_2}$ for all $x \in X$ and every o.

We define the mapping: $r_{o_1 o_2} : X \times X \to \mathbb{R}$ as:

$$r_{o_1 o_2}(x, y) = u_{o_1 o_2}(x) - u_{o_1 o_2}(y),$$

for all $x, y \in X$. Clearly, $r_{o_1 o_2}$ is continuous, concave (in the first argument and convex in the second argument), and skew-symmetric.

Next, define the preference function $r(x, y)$ for any $x, y \in X$ as:

$$\begin{aligned} r(x, y) &= \min_{\lambda \in \Delta(\mathbf{O})} \max_{\mu \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, y) d\lambda(o_1) d\mu(o_2) \\ &= \max_{\mu \in \Delta(\mathbf{O})} \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, y) d\lambda(o_1) d\mu(o_2). \end{aligned}$$

We prove that the function r rationalizes the data set O^∞. Consider $y \in X$ and some fixed $\bar{o}_1 = (p' \; x')' \in \mathbf{O}$. Let $\mu^{\bar{o}_1} \in \Delta(\mathbf{O})$ be the probability measure such that $\mu^{\bar{o}_1}(q, b) = 0$ if $(q' \; b)' \neq \bar{o}_1$ and $\mu^{\bar{o}_1}(q, b) = 1$ when $(q' \; b)' = \bar{o}_1$. We have:

$$\begin{aligned} r(x, y) &= \max_{\mu \in \Delta(\mathbf{O})} \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, y) d\lambda(o_1) d\mu(o_2) \\ &\geq \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, y) d\lambda(o_1) d\mu^{\bar{o}_1}(o_2) \\ &= \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} r_{o_1 \bar{o}_1}(x, y) d\lambda(o_1). \end{aligned}$$

It suffices to show that $r_{o_1 \overline{o_1}}(x, y) \geq p(x - y)$ for each data set $o = (o_1' \; \bar{o}_1)'$. But this follows directly from the definition of $r_{o_1 \bar{o}_1}$ and Theorem E. Hence, $r(x, y) \geq p(x - y)$.

(iii) $\implies$ (i).– It follows directly from the proof of (ii) $\implies$ (iii) that the constructed generalized maximin quasilinear preference function is also locally nonsatiated and skew-symmetric.

(ii) $\implies$ *(iv)*.– Suppose that statement (ii) holds. Define $o_1 = (p' \, a')'$ and $o_2 = (q' \, b')'$, such that $o = (o_1' \, o_2')'$ as in the proof of Theorem 7. Since the law of demand holds by definition, we can use the algorithm in the proof of Theorem E to find numbers that satisfy the inequalities in statement (iii) in Theorem E. We can define for every data set o_i, $i = 1, 2$:

$$\begin{aligned} u_{o_1}(x) &= u_{o_1} + p(x - a), \\ u_{o_2}(y) &= u_{o_2} + q(y - b). \end{aligned}$$

Next, we define the mapping: $r_{o_1 o_2} : X \times X \to \mathbb{R}$ as:

$$\begin{aligned} r_{o_1 o_2}(x, y) &= u_{o_1}(x) - u_{o_2}(y) \\ &= (u_{o_1} + p(x - a)) - (u_{o_2} + q(y - b)) \\ &= R_{o_1, o_2} + p(x - a) - q(y - b), \end{aligned}$$

for all $x, y \in X$, where $R_{o_1, o_2} = u_{o_1} - u_{o_2}$. Clearly, $r_{o_1 o_2}$ is continuous, linear in both arguments and satisfies $-r_{o_1 o_2} = r_{o_2 o_1}$.

Next, define the preference function $r(x, y)$ for any $x, y \in X$ as:

$$\begin{aligned} r(x, y) &= \min_{\lambda \in \Delta(\mathbf{O})} \max_{\mu \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, y) d\lambda(o_1) d\mu(o_2) \\ &= \max_{\mu \in \Delta(\mathbf{O})} \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, y) d\lambda(o_1) d\mu(o_2). \end{aligned}$$

Exactly as in the proof of $(ii) \implies (iii)$, we conclude that r is strictly increasing and continuous. We show that r is concave: Fix y and $\lambda \in \Delta(\mathbf{O})$. Consider:

$$f_{\lambda, y}(x) = \max_{\mu \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} r_{o_1 o_2}(x, y) d\lambda(o_1) d\mu(o_2).$$

We have:

$$\begin{aligned} f_{\lambda, y}(x) &= \max_{\mu \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} \int_{o_2 \in \mathbf{O}} [R_{o_1, o_2} + p(x - a) - q(y - b)] d\lambda(o_1) d\mu(o_2) \\ &= \int_{o_1 \in \mathbf{O}} [p(x - a) \max_{\mu \in \Delta(\mathbf{O})} \int_{o_2 \in \mathbf{O}} [R_{o_1, o_2} - q(y - b)] d\mu(o_2)] d\lambda(o_1). \end{aligned}$$

Note that $f_{\lambda, y}(x)$ is linear in x and therefore concave. Hence,

$$r(x, y) = \min_{\lambda \in \Delta(\mathbf{O})} \int_{o_1 \in \mathbf{O}} f_{\lambda, y}(x) d\lambda(o_1),$$

is the minimum over a set of linear functions and therefore concave.

We omit the proof that r rationalizes the data O^∞ because it is completely analogous to the proof of $(iii) \implies (v)$ in Theorem 5.

(iii) $\implies$ *(i).–* It follows directly from the proof of (ii) $\implies$ (iii) that the constructed generalized maximin quasilinear preference function is also locally nonsatiated and skew-symmetric.

(iv) $\implies$ *(i).–* Trivial.