Cardinal Revealed Preference, Price-Dependent Utility, and Consistent Binary Choice*

Victor H. Aguiar† and Roberto Serrano‡

This version: January 2018

Abstract We present a new notion of cardinal revealed preference that exploits the expenditure information in classical consumer theory environments with finite data. We propose a new behavioral axiom, Acyclic Enticement (AE), that requires the acyclicity of the cardinal revealed-preference relation. AE is logically independent from the Weak Axiom of Revealed Preference (WARP). We show that the Generalized Axiom of Revealed Preference (GARP), which characterizes the standard rational consumer, is logically equivalent to AE and WARP. We propose a new notion of rationalization by means of a price-dependent utility function that characterizes AE, which in particular is suitable for welfare analysis. We also propose a consistency condition for preference functions that is equivalent to WARP. We use our axiomatic decomposition to show, in experimental and scanner consumer-panel data sets, that AE explains the majority of the predictive success of GARP. Moreover, AE taken alone is superior in predictive success to both WARP and GARP.

JEL classification numbers: C60, D10.

Keywords: consumer theory; rationality; revealed preference approach; bounded rationality.

*We thank Roy Allen, Francis Bloch, Efe Ok, and John Quah for useful comments and encouragement. Judith Levi did a superb editing job.

†Department of Economics, The University of Western Ontario, vaguiar@uwo.ca

‡Department of Economics, Brown University, roberto_serrano@brown.edu

1. Introduction

Since the ground-breaking works of Afriat (1967) and Varian (1983), we have known that the rationalization of a list of price-demand observations by means of the maximization of a utility function subject to a linear budget constraint is equivalent to the Generalized Axiom of Revealed Preference (GARP). GARP requires that the revealed-preference relation be acyclic. We say that x^t is revealed preferred to x^s whenever x^t is chosen when x^s is affordable at prices p^t. This definition of revealed preference, due to Samuelson (1948), uses only ordinal information (i.e., it relies only on the fact that one commodity bundle is selected over another).

This result has allowed practitioners of revealed-preference analysis to test the null hypothesis of utility-maximizing behavior. Indeed, the exercise has been performed in experimental budget allocation data sets,[1] household consumption survey data,[2] and scanner consumption panels.[3] The empirical success of GARP in these different environments is usually limited, as quantified by predictive power measures such as those proposed by Beatty and Crawford (2011). In particular, the pass-rate of GARP is usually very small for high-powered environments (for detecting model inconsistencies) with substantial price variation (e.g., experimental and scanner data sets).[4] In the face of these facts, a natural question to ask is:

Question. *What consistency conditions are more primitive than GARP, but are themselves more empirically successful?*

We propose two primitive and logically-independent consistency conditions that, taken together, are equivalent to GARP. The first primitive condition is new; we call it *Acyclic Enticement* (AE), and it underlies our main contribution. Moreover, AE is suitable for meaningful welfare analysis in the absence of GARP, and is empirically more successful than the latter condition.

We define the *Acyclic Enticement* (AE) condition as the requirement that a cardinal revealed-preference relation be acyclic. We use the (cardinal) expenditure information available in a traditional consumer environment. In particular, we define the expenditure premium of x^t over x^s under prices p^t as the additional amount of dollars that the consumer spends on x^t at prices p^t when x^s was affordable. We say that a bundle x^t is cardinally revealed preferred to (more enticing than) x^s whenever the expenditure premium of x^t over x^s under prices p^t is not smaller than the expenditure premium of x^s over x^t under prices p^s.

The second primitive condition is the well-known *Weak Axiom of Revealed Preference* (WARP). WARP turns out to have an empirical performance similar to GARP. In fact, recent work by Cherchye et al. (2017) provides the necessary and sufficient conditions on price variation under which WARP is equivalent to GARP. Empirically, survey data price variation usually satisfies

[1] See Andreoni and Miller (2002), Choi et al. (2007), and Ahn et al. (2014).

[2] See Beatty and Crawford (2011) and Blundell et al. (2008).

[3] See Echenique et al. (2011) and Dean and Martin (2016).

[4] The pass-rate is the fraction of subjects in a sample that are consistent with a given condition such as GARP. When GARP has a high pass-rate (although this occurs only exceptionally), such as in survey data sets, this high pass-rate can be attributed to limited price variation, which implies that its predictive power is usually close to zero.

these conditions (Cherchye et al., 2017), which means that in this type of data set, WARP and GARP are indistinguishable. In our application, we find in experimental and scanner data sets that WARP has low empirical success, and its performance is relatively equivalent to that of GARP.[5] This empirical and theoretical finding supports the need for different primitive conditions from WARP, such as AE.

A second contribution of our work is to provide separate rationalization results for Acyclic Enticement and for WARP, in the form of a price-dependent utility maximization and a pairwise-consistent preference function, respectively. We establish that standard utility maximization is the intersection of these two classes of models (price-dependent preferences and pairwise consistency).

Price-dependent utilities were studied early by Samuelson (1968) to model the demand for money (Basmann et al., 1987). However, it quickly became apparent that any data set can be rationalized by a price-dependent utility without further restrictions (Pollak, 1977, Shafer, 1974). We require that price-dependent utility functions satisfy a form of monotonicity in the expenditure premium. This restriction allows us to connect cardinal revealed preferences with information about the indirect price-dependent utilities. Due to this feature, we show that welfare analysis is possible even if WARP –or the law of demand– fails.[6] Reference-dependent utility functions are at the center of many behavioral models, including prospect theory (Kahneman and Tversky, 1979). Our new model of reference dependence contributes to this literature from a revealed-preference point of view.[7] The work of Deb et al. (2017) also provides a new revealed-preference theory, in terms of prices, that corresponds to an expenditure-dependent utility function. We show that their new revealed-preference condition is logically independent of AE.

Consistency in binary choices leads to a new representation of WARP. We show that WARP is equivalent to the rationalization of a data set by means of maximizing a preference function that satisfies a consistency property, which we call *sign asymmetry*. The result can be seen as a variant of the work of Quah (2006) in demand functions, for finite data sets.

The only previous work we are aware of that has broken down GARP into more primitive conditions has done it exclusively in infinite data sets (i.e., for demand functions). Hurwicz and Richter (1979) shows that for demand functions GARP is equivalent to both WARP and an axiom called the Ville Axiom of Revealed Preference.[8] The Ville Axiom is not testable with finite data sets; in addition, due to its differential nature, it is not easily compared with our AE condition. However, we demonstrate that AE is different from the Ville Axiom.[9] More importantly, we show

[5]This is so even when the price variation in these data sets may not satisfy the conditions in Cherchye et al. (2017).

[6]Some other works that allow one to perform welfare analysis in such situations from a revealed-preference perspective include: (i) Ok et al. (2015), introducing an endogenous reference point; (ii) limited consideration models initiated by Masatlioglu and Ok (2005), which separate attention from preference rationalizing violations of WARP; and (iii) DellaVigna et al. (2017) proposing a structural job search model with reference dependence, where the behavior of agents is affected by their level of consumption.

[7]See the behavioral consumer model of Kőszegi and Rabin (2006), for a model with endogenous reference dependence that allows for violations of WARP.

[8]The Ville Axiom of Revealed Preference rules out the existence of a differential version of revealed-demand cycles. Hurwicz and Richter (1979) shows the Ville Axiom to be equivalent to the symmetry of the Slutsky substitution matrix function.

[9]We show that demand functions that are consistent with the Ville Axiom can generate finite consumption data

that AE and WARP are together logically equivalent to GARP, which, to the best of our knowledge, provides the first axiomatic decomposition of GARP in finite data.

Section 2 establishes the axiomatic decomposition of GARP into AE and WARP. Section 3 presents the notion of price-dependent utility maximization, states its connection with AE, and elaborates on how welfare analysis can be conducted under AE alone (in the absence of WARP). Section 4 presents the notion of rationalization by maximizing a preference function, and establishes its connection with WARP. Section 5 presents an empirical application of our new axiom AE, along with WARP and GARP to both experimental and scanner consumer-panel data sets. We find that AE has higher predictive success (as proposed by Beatty and Crawford (2011)) than WARP and GARP. Section 6 proposes an additive decomposition of the predictive success of GARP into the marginal contributions of AE and WARP; in the same experimental and scanner data sets, we find that AE explains the majority of the empirical success of GARP. Section 7 expands on the formal relation of the new AE condition with numerous models presented in others' previous work. Finally, Section 8 concludes. All the proofs are collected in an appendix.

2. Setup and Axiomatic Decomposition of GARP

2.1. Setup: Finite Data Sets and Afriat's Theorem

We consider a finite data set consisting of a list of prices and consumption bundles $O^K = \{p^t, x^t\}_{t=1}^K$ for a finite $K \geq 2$, where $p^t \in P \equiv \mathbb{R}^L_{++}$ and $x^t \in X \equiv \mathbb{R}^L_{+} \setminus \{0\}$ for all $t = 1, \cdots, K$. In what follows, we denote the inner product of two vectors $v, w \in \mathbb{R}^L$ with $vw = \sum_{l=1}^L v_l w_l$.

First, we define the traditional revealed-preference framework, which takes into account only ordinal information. Indeed, the objective of this section is to formally establish the notion of rationalization by means of maximizing a utility function subject to a budget constraint, and its equivalence with GARP. Then we will proceed to decompose GARP axiomatically into two more primitive conditions, which will be the basis of our empirical and theoretical analysis. We begin with some preliminaries.

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

If x^t is directly revealed preferred to x^s, this means that the consumer chose x^t and not x^s, when she could have chosen the latter. If x^t is directly and strictly revealed preferred to x^s, then, in addition, the consumer could have saved some money by choosing x^s.

Next, we use this binary relation to define a set of behavioral conditions that characterize a rational consumer-demand behavior. We define the revealed-preference relation $\succeq^R$ as the transitive

sets that fail AE.

closure of the direct revealed-preference relation $\succeq^{R,D}$. We let $\succ^{R}$ be the transitive closure of the direct strict preference relation $\succ^{R,D}$, and, we let $\sim^{R}$ be the revealed indifference relation. (The transitive closure of $\sim^{R,D}$. The relation $\sim^{R,D}$ is defined as follows: $x^t \succeq^{R,D} x^s$ and not $x^t \succ^{R,D} x^s$, namely, $p^t[x^t - x^s] = 0$.)

Definition 2. (Revealed Preference) We say that x^t is revealed preferred to x^s ($x^t \succeq^{R} x^s$) whenever there is a chain $x^1, x^2, \ldots, x^n$ with $x^s = x^1$ and $x^t = x^n$ such that $x^1 \succeq^{R,D} x^2 \succeq^{R,D} \ldots \succeq^{R,D} x^n$. We say that x^t is strictly revealed preferred to x^s ($x^t \succ^{R} x^s$) whenever there is a chain like that above with strict and direct revealed preferences.

We state here a classic axiom in revealed preference theory:

Axiom 1. *(Generalized Axiom of Revealed Preference, GARP) There is no pair $s, t \in \{1, \cdots, K\}$ such that $x^t \succeq^{R} x^s$, and $x^s \succ^{R,D} x^t$.*[10]

It follows from Afriat's theorem, that GARP is necessary and sufficient for (weak) rationalization, as defined below.

Definition 3. (Data Rationalization) We say that a data set O^K is (weakly) rationalized by a utility function $u : X \mapsto \mathbb{R}$ whenever for each $k = 1, \ldots, K$, it follows that $u(x^k) \geq u(x)$ for all $x \in X$ such that $p^k x \leq p^k x^k$.

We also say that a utility function $u : X \mapsto \mathbb{R}$ is *locally nonsatiated* if for any $x \in X$ and for any $\epsilon > 0$, there exists a $y \in B(x, \epsilon)$ where $B(x, \epsilon) = \{z \in X : | \quad ||z - x|| \leq \epsilon\}$ such that $u(y) > u(x)$.[11]

The classic theorem, due to Afriat (1967), that establishes the logical equivalence between GARP and rationalization by a locally nonsatiated utility function is stated next.

Theorem (Afriat's Theorem (1967))**.** *The following statements are equivalent:*

(i) A data set O^K can be rationalized by a continuous and locally nonsatiated utility function.

(ii) A data set O^K satisfies GARP.

(iii) A data set O^K can be rationalized by a strictly increasing, continuous, and concave utility function.

Afriat's Theorem has formed the basis for testing rationality in different setups. In our empirical application, we use an experimental and a scanner consumption-panel data set to show that the empirical success of rationality is limited.

[10] Equivalently, for all $n \geq 2$, if there is a chain $x^1, x^2, \ldots, x^n$ such that $x^1 \succeq^{R,D} x^2 \succeq^{R,D} \cdots \succeq^{R,D} x^n \succeq^{R,D} x^1$, then it must be the case that $x^s \sim^{R,D} x^t$ for all s, t in the set of indices $\{1, 2, \ldots, n\}$.

[11] 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 that $lim_{n\to\infty} u(x^n) = u(x)$; (ii) *strictly incresing* if $x, y \in X$, $x_l \geq y_l$ for all $l = 1, \cdots, L$ and $x_k \geq y_k$ for some $k \in= 1, \cdots, L$ implies that $u(x) > u(y)$; (iii) and *concave* if for any $x, y \in X$, it follows that $u(x) - u(y) \geq \xi u(x)[x - y]$, for $\xi \in \nabla u(x)$ where $\nabla u(x)$ is the supergradient of the utility.

2.2. Axiomatic Decomposition of GARP

In this subsection, we propose two more primitive behavioral axioms that are logically independent, and which, taken together, are equivalent to GARP. The first condition is new, and we call it Acyclic Enticement (AE). To define AE, we first establish the notion of revealed comparability.

Definition 4. (Revealed Comparability) We say that any two consumption bundles x^t, x^s are comparable if either $x^s \in \{x \in X | p^t x \leq p^t x^t\}$ or $x^t \in \{x \in X | p^s x \leq p^s x^s\}$.

That is, we say that two bundles are comparable if one of them is affordable when the other was chosen.

We are now ready to introduce the cardinal revealed-preference framework. We begin by defining the revealed expenditure premium.

Definition 5. (Revealed Expenditure Premium, REP) The revealed expenditure premium of commodity bundle x^t and prices p^t over x^s is given by $\rho_t(x^s) = p^t[x^t - x^s]$ when x^t is chosen and x^s is affordable at p^t (i.e., $x^s \in \{x \in X | p^t x \leq p^t x^t\}$); otherwise, $\rho_t(x^s) = 0$.

The REP of a bundle over another is a (cardinal) wealth amount. The idea is to define the notion of how enticing a commodity bundle x^t is with respect to another x^s, by comparing the intensity of their REPs.

Definition 6. (Cardinal Revealed Preference/Revealed Enticement) We say that x^t is cardinally and directly revealed preferred to (directly more enticing than) x^s ($x^t \succeq^{E,D} x^s$) whenever the two bundles are comparable and $\rho_t(x^s) \geq \rho_s(x^t)$. We say x^t is cardinally, strictly and directly revealed preferred to (strictly and directly more enticing than) x^s ($x^t \succ^{E,D} x^s$) whenever the two bundles are comparable and $\rho_t(x^s) > \rho_s(x^t)$.

That is, for x^t to be cardinally revealed preferred to x^s, it must be directly revealed preferred; however, in addition, the intensity of that direct revealed preference, measured by the REP, must be at least as strong as the intensity of the reverse REP.

The cardinal revealed-preference relation just defined, by depending upon the REP, takes into account cardinal information revealed by the consumers' expenditure levels. We measure how enticing x^t is relative to x^s, at prices p^t, using the quantity $p^t[x^t - x^s] = w^t - p^t x^s$ (when x^s is affordable), that is, the difference between the actual expenditure w^t incurred to buy x^t and the counterfactual expenditure at price p^t for x^s. The larger this difference, the less enticing x^s is at price p^t. For example, the story might be one in which the consumer has decided to spend w^t when prices are p^t, and consumption plans according to which she ends up spending much less do not suit her planning well. Hence, in addition to the direct revealed preference, x^t being cardinally and directly preferred to x^s means that x^s is less enticing at p^t than x^t is at p^s. We say then that $x^t \succeq^{E,D} x^s$. We define (i) the cardinal preference relation or enticement relation, $\succeq^E$, as the transitive closure of $\succeq^{E,D}$; (ii) the strict cardinal preference relation or strict enticement,$\succ^E$, as the transitive closure of $\succ^{E,D}$; and (iii) the cardinal indifference relation, $\sim^E$, as the transitive closure of $\sim^{E,D}$, defined as $x^t \sim^{E,D} x^s$ whenever $x^t \succeq^{E,D} x^s$ and not $x^t \succ^{E,D} x^s$.

Next, we formulate a new axiom:

Axiom 2. *(Acyclic Cardinal Revealed Preference/Acyclic Enticement, AE) There is no pair $s, t \in \{1, \cdots, K\}$ such that $x^t \succeq^E x^s$, and $x^s \succ^{E,D} x^t$.*[12]

In words, if the consumer finds x^1 (directly or indirectly) more enticing than x^n, then x^n should not be strictly and directly more enticing than x^1.

The second primitive axiom is well known. It is the Weak Axiom of Revealed Preference (WARP).

Axiom 3. *(Weak Axiom of Revealed Preference, WARP) There is no pair $s, t \in \{1, \cdots, K\}$ such that $x^t \succeq^{R,D} x^s$, and $x^s \succ^{R,D} x^t$.*

This version of WARP requires that, if the consumer chooses x^t when x^s is affordable, then, when she chooses x^s, it must be that x^t is at least as expensive (i.e., either it is no longer affordable or both bundles cost the same).

We establish here that a data set O^K satisfies GARP if and only if it satisfies WARP and AE.

Lemma 1. *A data set O^K satisfies WARP and AE if and only if it satisfies GARP.*

With this result in hand, we turn now to prove that WARP and AE are logically independent axioms, by providing cases that fail either the former or the latter.

Example 1. (AE does not imply WARP) By definition, AE can allow for both $x^t \succ^{R,D} x^s$ and $x^s \succ^{R,D} x^t$ simultaneously. Let's say you have $p^t = (1, 2)$, $x^t = (6, 0)$; and $p^s = (1, 3)$, $x^s = (1, 2)$. This violates WARP since $p^t(x^t - x^s) = 1$ ($x^t \succ^{R,D} x^s$) and $p^s(x^s - x^t) = 1$ ($x^s \succ^{R,D} x^t$). However, AE holds, and we can therefore conclude that $x^s \sim^{E,D} x^t$, since $\rho_s(x^t) = \rho_t(x^s)$ with both x^t, x^s being comparable.

Example 2. (WARP does not imply AE) For a data set satisfying WARP, but violating AE, let there be three commodities, and the following three situations: $p^1 = (1, 1, 2)$, $p^2 = (2, 1, 1)$, and $p^3 = (1, 2, 2)$ with corresponding choices $x^1 = (1, 0, 0)$, $x^2 = (0, 1, 0)$, and $x^3 = (0, 0, 1)$, respectively. These choices satisfy WARP: $x^1 \succeq^{R,D} x^2$ and not $x^2 \succ^{R,D} x^1$; $x^2 \succeq^{R,D} x^3$ and not $x^3 \succ^{R,D} x^2$; and $x^3 \succeq^{R,D} x^1$,and not $x^1 \succ^{R,D} x^3$. Furthermore, $x^1 \succeq^{E,D} x^2$, $x^2 \succeq^{E,D} x^3$ and $x^3 \succ^{E,D} x^1$. Hence, it is not the case that $x^1 \sim^{E,D} x^3$ and, thus, AE fails.

Here we provide an example of a three-point data set that allows one to distinguish WARP and GARP. The example still fails both, yet satisfies AE. The example graphically illustrates the usefulness of AE for inferring preferences from seemingly incoherent binary comparisons.

Example 3. (Making sense of a data set that violates GARP and WARP) Let $p^1 = (\frac{5}{2}, \frac{9}{8})$, $p^2 = (\frac{135}{128}, \frac{5}{4})$, $p^3 = (2, 2)$, $x^1 = (2, 1)$, $x^2 = (\frac{258}{1985}, \frac{1341}{397})$, and $x^3 = (1, \frac{39}{160})$. This data set violates GARP; in fact, it has two violations of WARP and one violation of GARP (that is not a violation of WARP). Clearly, this data set presents a difficult situation for the standard revealed-preference analysis: there are 2 cycles of size 2 (i.e., $x^1 \succ^{R,D} x^2$, $x^2 \succ^{R,D} x^1$; and $x^2 \succ^{R,D} x^3$, $x^3 \succ^{R,D} x^2$), so

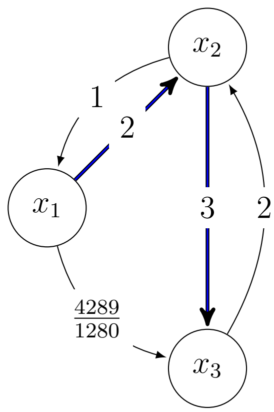

Figure 1 – Violation of GARP

the consumer fails even a simple pairwise-consistency requirement. Moreover, there is a cycle of size 3 (i.e., $x^1 \succ^{R,D} x^3 \succ^{R,D} x^2 \succ^{R,D} x^1$) which means that the consumer also fails transitivity.

We will show that AE can make sense of this data set. It does so by using the information on relative expenditure premia to recover preference information. But first we will visualize the GARP violations. The expenditure premiums are $\rho_1(x^2) = 2$, $\rho_2(x^1) = 1$, $\rho_2(x^3) = 3$, $\rho_3(x^2) = 2$, $\rho_1(x^3) = \frac{4289}{1280}$, and $\rho_3(x^1) = 0$. The expenditure premium information can be summarized in a matrix ρ with entry $\rho_{ij} = \rho_i(x^j)$ for $i, j = 1, 2, 3$. We can represent ρ using a directed graph, where each good is a vertex $V = \{x^1, x^2, x^3\}$ and there is a directed edge $\{x^t \to x^s\}$ if $\rho_t(x^s) > 0$. The data set in this example is visualized in figure 1. Using a graphic representation of ρ, we can see that each violation of GARP is equivalent (for this example with no revealed indifference) to the graph of ρ having a cycle. Cycles involving 2 vertices violate WARP and cycles involving 3 or more violate GARP (and not WARP).

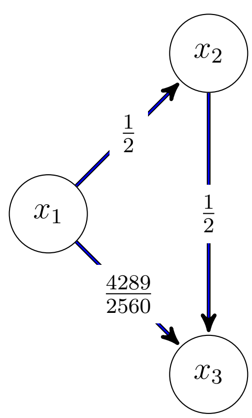

Figure 2 – AE consistent data set.

It turns out that the AE can be seen as a cyclical consistency condition on the skew-symmetric part of the matrix ρ. We define the skew-symmetric part of the matrix ρ as $\rho^{skew} = \frac{1}{2}[\rho - \rho']$ with entry $\rho^{skew}_{ij} = \frac{1}{2}[\rho_i(x^j) - \rho_j(x^i)]$ for $i, j = 1, 2, 3.$. Relying on a graphic representation of ρ^{skew} and

[12] Equivalently, for all $n \geq 2$, if there is a chain $x^1, x^2, \ldots, x^n$ such that $x^1 \succeq^{E,D} x^2 \succeq^{E,D} \cdots \succeq^{E,D} x^n \succeq^{E,D} x^1$, then it must be the case that $x^s \sim^{E,D} x^t$ for all $s, t \in \{1, \cdots, n\}$.

using rules analogous to the case of ρ, we note that AE is equivalent to the fact that the associated graph has no cycle (for the current example without revealed indifference). For the same data set, this is in fact the case as seen in figure 2, since we have an acyclic graph. The AE allows us to infer that x^1 is more enticing than either x^2 and x^3, and x^2 is more enticing than x^3. We show in our application that this example is in fact empirically relevant, since WARP and GARP are routinely violated in experimental and scanner data sets while AE is not.

We are ready to state our first result. The following theorem is a direct consequence of Lemma 1:

Theorem 1. *The following statements are equivalent:*

(i) A data set O^K can be rationalized by a locally nonsatiated utility function.

(ii) A data set O^K satisfies GARP.

(iii) A data set O^K can be rationalized by a strictly increasing, continuous, and concave utility function.

(iv) A data set O^K satisfies AE and WARP.

The proof is an obvious consequence of Lemma 1 and Afriat's Theorem.

3. Price-Dependent Utility Rationalization and Acyclic Enticement

In what follows, we focus on investigating the consequences of relaxing GARP using the more primitive conditions of AE and WARP. The main theoretical benefit of providing a finer characterization of rationalization in terms of axioms more primitive than GARP is that it permits the investigation of the empirical and theoretical robustness of such rationalizations. In particular, we find that if one relaxes WARP, the data set is rationalized by a price-dependent utility function. Similarly, if one relaxes AE, the data set is rationalized by a preference function that captures pairwise-consistent comparisons.

Next, we present the notion of price-dependent rationalization.

Definition 7. (Data price-dependent rationalization) We say that a data set O^K is (weakly) rationalized by a **price-dependent** utility function $u : X \times P \mapsto \mathbb{R}$ whenever, for each $k = 1, \cdots, K$, it follows that $u(x, p^k) \leq u(x^k, p^k)$ for $x \in X$, $p^k x \leq p^k x^k$.

Price-dependent utilities are very natural in that the consumer may change her tastes depending on the context or on a reference point. In our notion of price-dependent rationalization we implicitly assume that the consumer maximizes her price-dependent utility function by fixing the price to that of the choice set, which means that for a given choice set the consumer is maximizing a "usual" utility function. However, when she is presented with new prices, her preferences may change. Without additional assumptions any data set O^K can be rationalized by a price-dependent utility. In this regard, one needs additional restrictions on the price-dependent utilities, to obtain testable implications. Ideally, we would like these restrictions to be natural and have good empirical performance. We show that AE provides such a restriction.

We consider the following property of a price-dependent utility function:

Definition 8. (Expenditure-premium monotonicity) We say that a price-dependent utility $u : X \times P \mapsto \mathbb{R}$ that rationalizes data set O^K is monotonic in the expenditure premium whenever for all $t, s \in \{1, \cdots, K\}$, when x^t, x^s are comparable, $u(x^t, p^t) \geq u(x^s, p^s)$ if and only if $\rho_t(x^s) \geq \rho_s(x^t)$.

The rationalization of a data set by means of maximizing a price-dependent utility function that satisfies expenditure-premium monotonicity implies AE trivially. The argument resembles the reasoning used to show how the traditional rationalization by means of maximizing a price-independent utility function (with local nonsatiation) implies GARP (i.e., the price-independent utility satisfies that $u(x^t) \geq (>) u(x^s)$ whenever $p^t[x^t - x^s] \geq (>)0$).

We say that a price-dependent utility function $u(\cdot, \cdot)$ is continuous, if for a sequence $(z^r)_{r \in \mathbb{N}_+}$ such that $z^r = (x^r, p^r)$ for each $r \in \mathbb{N}_+$, $x^r \in X$ and $p^r \in P$ and $lim_{r \to \infty} z^r = z \in X \times P$, then $lim_{r \to \infty} u(z^r) = u(z)$.

We say that a price-dependent utility function $u(\cdot, \cdot)$ is strictly increasing in its first argument whenever $u(x, \cdot) > u(y, \cdot)$ if $x_l \geq y_l$ for all $l = 1, \ldots, L$ with at least one strict inequality.

The fact that AE implies the existence of a continuous price-dependent utility with the property of expenditure-premium monotonicity (and strict monotonicity in the first argument) is less obvious; so this is the subject of our next result.

Theorem 2. *The following statements are equivalent:*

(i) A data set O^K satisfies Acyclic Enticement.

(ii) A data set O^K can be rationalized by a continuous price-dependent utility function that is monotonic in the expenditure premium and strictly increasing in its first argument.

The expenditure-premium monotonicity implies a form of increasing differences. It restricts price-dependent utilities by connecting observed cardinal revealed preferences with latent indirect utilities, thereby allowing for the possibility of meaningful welfare comparisons for a single individual facing different prices. Without expenditure-premium monotonicity, it is not clear whether we can extract welfare information in the absence of WARP. A violation of WARP implies that consumers are reversing their choices. This implies that a commodity bundle x^t is (ordinally) revealed preferred to x^s at prices p^t, and x^s is (ordinally) revealed preferred to x^t at prices p^s. In a situation like this, the traditional revealed-preference welfare analysis (à la Varian (1983)) is mute. Our price-dependent utility opens the door to extracting welfare information from these seemingly "incoherent" choices. After all, AE says that as long as the consumer acts coherently with respect to the cardinal revealed preference relation, we can conclude that the indirect utility associated with x^t is greater than that associated with x^s.

3.1. Forecast and Welfare Analysis under AE

In this subsection, we elaborate on the preceding arguments, making the point that AE alone –i.e., in the absence of WARP– may suffice for meaningful welfare and forecast analysis to be conducted.

Forecast.– The forecast problem in revealed preferences is to predict a set of consumption bundles x^{K+1} for a given new, out-of-sample price p^{K+1}. We follow Varian (1983) in defining the counterfactual demand set for the new AE condition. The counterfactual demand set for GARP and WARP are well known.

Definition 9. Varian-AE Counterfactual Demand Set. For a new price $p^{K+1} \in P$, the counterfactual demand set $S^{AE}(p^{K+1}) \subseteq X$ is defined as:

$$S^{AE}(p^{K+1}) = \{x \in X | \{p^t, x^t\}_{t=1}^{K+1} \text{ satisfies AE when } x^{K+1} = x\}.$$

One important advantage of our new AE condition is that we can directly adapt the tools used for GARP, as AE preserves the latter's structure, while exploiting the cardinal information in the traditional consumer environment. In fact, we can compute the Varian-AE counterfactual demand set by adapting the same algorithms for GARP, with the new definition of cardinal revealed preferences.

Supporting Sets and Recoverability.– Now we turn our attention to the recoverability of preferences. First, we define the supporting set as the set of prices such that, for a new commodity bundle, they will make the extended data consistent with AE.

Definition 10. Varian-AE Price Support Set. For a new commodity bundle set $\{x^{K+i}\}_{i=1}^{I} \in X$ for $I \geq 1$, the price support set $S^{AE}(\{x^{K+i}\}_{i=1}^{I}) \subseteq P$ is defined as:

$$S^{AE}(\{x^{K+i}\}_{i=1}^{I}) = \{\{p^i\}_{i=1}^{I} \in P | \{p^t, x^t\}_{t=1}^{K+I} \text{ satisfies AE when } p^{K+i} = p^i \forall i = 1, \cdots, I\}.$$

The Varian-AE price support set allows us to ask allows us to inquire into the recoverability of preferences in the sense of Afriat. Given a finite data set O^K that satisfies AE, and two new commodity bundles x^{K+1} and x^{K+2}, is there a way to describe the entire set of price-dependent utility functions that are consistent with the data?

If every price pair $(p^{K+1}, p^{K+2}) \in S^{AE}(\{x^{K+1}, x^{K+2}\})$ is such that $x^{K+1} \succeq^E x^{K+2}$ in the extended data set, then we conclude that x^{K+1} (i) is cardinally revealed preferred to x^{K+2}, and (ii) yields a better outcome according to the price-dependent utility function. To make the latter point more precise, we now study welfare.

Welfare.– In this subsection we propose a way to do welfare analysis when violations of WARP are present. In particular, we use AE and our price-dependent utility function representation to provide a consistent methodology for doing welfare analysis. Such a consistent welfare analysis should be transitive.

Definition 11. Indirect price-dependent utility. We say that $v : P \times \mathbb{R}_{++} \to \mathbb{R}$ is an indirect (price-dependent) utility whenever there exists a price-dependent utility $u(\cdot, \cdot)$ such that $v(p, w) = max_{q \geq 0} u(q, p)$, subject to $pq \leq w$.

In particular, due to the continuity of the price-dependent utility and the properties of the linear constraint, $v(p, w)$ is a continuous mapping. We also observe that $u(x^t, p^t) = v(p^t, p^t x^t)$ for all $t = 1, \cdots, K$. Now we state some properties of the indirect utility associated with the price-dependent utility maximization.

Lemma 2. *The indirect price-dependent utility $v(p, w)$ satisfies the following properties:*

(i) v is continuous in $\{p, w | p >> 0, w > 0\}$.

(ii) v is strictly increasing in w.

(iii) v is monotone in the expenditure premium, i.e., $x^t \succeq^{E,D} x^s$ if and only if $v(p^t, p^t x^t) \geq v(p^s, p^s x^s)$.

The proof is omitted since it follows trivially from our previous results, or by analogy to the rational case, which is well known in the literature. The continuity and strict monotonicity properties allow us to define an expenditure function by means of an inverse mapping. The key property is the monotonicity in the expenditure premium, which provides a welfare (ordinal) interpretation of the cardinal revealed-preference information. This welfare information is obtained even if WARP and GARP fail. We can now define an expenditure function.

Definition 12. The expenditure function $e(p, \overline{u})$ is the solution to $\overline{u} = v(p, e(p, \overline{u}))$ for a fixed $\overline{u}$.

The expenditure function allows us to obtain the wealth required to obtain a fixed level of utility, given a certain price. Notice that, with the help of the expenditure function that we just defined, we can now define the compensating and equivalent variations. However, duality no longer works.[13] Nonetheless, we can still calculate the expenditure function from finding the cheapest bundle that provides a fixed level of utility.

Lemma 3. *The expenditure function is the value of $e(p, \overline{u}) = min_{q \geq 0} pq$ subject to $u(q, p) \geq \overline{u}$.*

The proof is trivial and therefore suppressed. In fact, notice that for a fixed price $p \in P$, the above problem is equivalent to the classical expenditure-minimization problem with a continuous and strictly increasing utility function $u(\cdot, p)$.

Now we can define the money-metric version of our price-dependent utility.

[13] In particular, due to the dependence on prices of the utility function we cannot claim that the Walrasian demand mapping can be obtained using the Hicksian demand mapping. Note in fact that by the Envelope's theorem (assuming for a moment differentiability) will give us: $\nabla_p e(p, \overline{u}) = q + \nabla_p u(q, p)$, which has an additional term that is zero in the classical problem with price independent utility functions.

Definition 13. Money-metric utility. We define the money-metric price-dependent utility as $m(x, p) = e(p, u(x, p))$, for a given price-dependent utility $u(\cdot, \cdot)$.

We follow Varian (1983) in proposing an upper bound for the money-metric utility, for the price-dependent utility function associated with an observed data set. First we define the cardinally revealed upper contour set for any bundle $x^0 \in X$:

$$RE(x^0) = \{x \in X | \forall p \in S^{AE}(\{x^0, x\}); \quad x \succ^E x^0\}.$$

The relation $\succeq^E$ is defined in the extended data set. Now, we define the upper bound of the money-metric function:

$$m^+(p, x^0) = min_{t=1,\cdots,K} p^t z, \text{ such that } z \in RE(x^0).$$

Notice that even if these constructions are analogous to those from Varian (1983), our analysis can still be done even under violations of GARP and, in particular, of WARP. Using the money-metric utility, we can compute bounds for the compensating and equivalent variation.

4. Preference Functions and the Weak Axiom of Revealed Preference

In this section, our goal is to provide a simple characterization of WARP. This will allow us to better understand the meaning of the traditional utility maximization framework as the intersection of the price-dependent utility maximization and a model of consistent pairwise comparisons. To do this, we first introduce the notion of rationalization of the data by a preference function:

Definition 14. (Data Preference-Function Rationalization) A data set O^K is rationalized by a preference function $r : X \times X \mapsto \mathbb{R}$ whenever, for every x^t in the data, $r(x^t, x^t) = 0$ and $r(x^t, x) \geq 0$ if $x \in X$ is such that $p^t x^t \geq p^t x$.

Next, we introduce a property for preference functions that turns out to be key in the characterization of WARP.

Definition 15. (Sign-Asymmetric Preference Function) We say that a preference function $r : X \times X \mapsto \mathbb{R}$ is *sign-asymmetric* if $r(x, y) \geq 0$ implies that $r(y, x) \leq 0$ for all $x, y \in X$.

The next definition is also for a property of preference functions:

Definition 16. (Locally Nonsatiated Preference Function) We say that a preference function $r : X \times X \mapsto \mathbb{R}$ is *locally nonsatiated* if for any $x, y \in X$ such that $r(x, y) = 0$ and for any $\epsilon > 0$, there exists a $y' \in B(y, \epsilon)$ where $B(y, \epsilon) = \{z \in X : | \quad ||z - y|| \leq \epsilon\}$ such that $r(x, y') < 0$.

This property rules out thick indifference curves. If we define the indifference set associated with $x \in X$ by $\{z \in X | r(x,z) = 0\}$, then the local nonsatiation property requires that we can find an item close to any item in the indifference set that is dominated by it.

The main result in this section is the following:

Theorem 3. *The following statements are equivalent:*

(i) The data set O^K satisfies WARP.

(ii) The data set O^K is rationalized by a preference function $r : X \times X \mapsto \mathbb{R}$ that is sign-asymmetric and locally nonsatiated.

Theorem 3 can be seen as a variant of the work of Quah (2006) for finite data sets and for demand correspondences. Quah (2006) establishes the rationalization of a stronger form of WARP in infinite data sets and demand functions, by means of a preference function that is sign-asymmetric. However, he imposes additional properties such as skew-symmetry and continuity, among other regularity conditions.[14] We compare our result and that of Quah (2006) in more detail, in section 7. The rationalization by a sign-asymmetric preference function is equivalent to consistent pairwise comparisons. However, the consumer's preference may violate transitivity. Consistent choice with pairwise comparisons has also been explored in Shafer (1974) and Gerasimou (2010) as models of bounded rationality.

As a consequence of our results, we can make the following observation in figure 3.

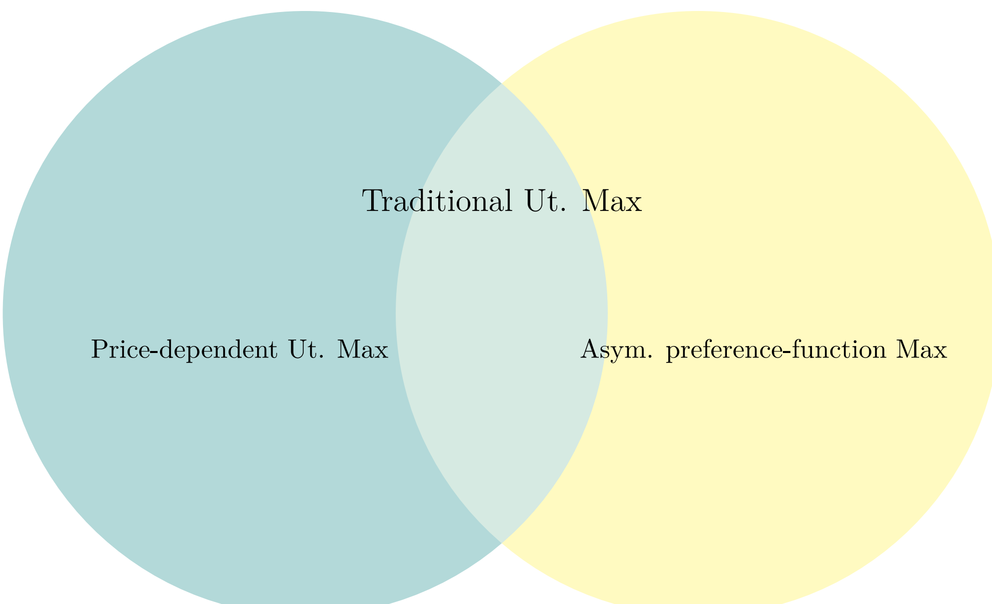

Figure 3 – Utility Maximization Characterization.

That is, traditional utility maximization is at the intersection of (i) the model of price-dependent utility maximization satisfying monotonicity in the expenditure premium, and (ii) the model of preference-function maximization satisfying pairwise consistency. In fact, the traditional model

[14]Skew-symmetry means that for any x^t, x^s, it must be the case that $r(x^t, x^s) = -r(x^s, x^t)$.

requires that (i) the price-dependent utility be constant for any fixed commodity $x \in X$ (i.e., $u(x,p) = u(x)$ for all $p \in P$), and (ii) that the preference-function be additive in its arguments (i.e., $r(x,y) = u(x) - u(y)$). In what follows, we explore which of these two models is more robust and empirically successful.

5. Application of Predictive Success: Relative Performance of AE versus WARP

Since the seminal work of Selten (1991), and Beatty and Crawford (2011) (henceforth BC) on predictive success in the context of the revealed-preference methodology, we know that the mere fact that a data set is consistent with behavioral conditions such as the new AE condition, WARP, or GARP, does not necessarily mean that the model has predictive power. Simply put, the apparent success rate of a behavioral condition may be due to the fact that it is nondemanding empirically. Selten (1991) proposes that the predictive success of a behavioral condition is measured as the difference between the hit rate and the area:

predictive success=hit rate - area.

As explained next, it is a pass/fail index for any condition, say condition m, corrected for the ability to find rejections of the condition.

We follow the BC methodology for predictive success based on the Selten index, using the formulation presented by Demuynck (2015). Say that an experiment $\mathcal{D}$ is a collection of n data sets generated by $i = 1, \cdots, n$ individual consumers, such that $\mathcal{D} = \{O_i^K\}_{i=1}^n$. Let $\Omega = (\mathbb{R}^L_{++} \times X)^K$ denote the set of all possible data sets (i.e., K dimensional arrays of prices and consumption bundles). Define the set $S^m \subseteq \Omega$ as the subset of the outcome space which is consistent with condition $m \in \{\text{AE, WARP, GARP}\}$. Consider the indicator function $I : \Omega \to \{0,1\}$ such that $I(O) = 1$ if and only if $O \in S^m$.

The hit rate of an experiment $\mathcal{D}$ is given by the fraction of individual data sets that are contained in S^m:

$$r^m = \frac{1}{n}\sum_{i=1}^{n} I(O_i^K).$$

The area is defined relative to the set of all possible outcomes (Ω). To define the area associated with model m, a^m, we need some additional definitions. Fix an individual data set $O_i^K \in \Omega$ and define a random array $\mathbf{O}$ with conditional support $\Omega | O_i^K$ and with conditional c.d.f. $\mathbb{F}_i$. The area associated with the individual data set $O_i^K$is given by $\rho(O_i^K)$ taking values in $[0,1]$:

$$\rho^m(O_i^K) = \int I(O)\mathbb{F}_i(dO).$$

The larger this number, the more difficult it is to reject condition m given the data set; if it equals 1, it is impossible to reject it.

The relative size of the area of model m, a^m, is given by the mean of the areas of each data set in the experiment $\mathcal{D}$:

$$a^m = \frac{1}{n}\sum_{i=1}^{n} \rho(s_i).$$

The Selten (1991) predictive-success index for behavioral condition m, in a given experiment with n individuals is given by:

$$\beta^m = r^m - a^m.$$

The predictive-success measure is essentially (up to an affine transformation) the only measure that satisfies the following axioms: (i) Monotonicity, which requires that a behavioral condition that is extremely demanding and consistent with the data should be judged as better than another one that is inconsistent with the data and not demanding; (ii) Equivalence, which requires that a situation in which there are no restrictions and a situation in which nothing is ruled out are considered equally uninformative; and (iii) Aggregability, which requires that the measure be additive over heterogeneous consumers.

In our application, we propose a modification of the random uniform-choice benchmark in BC and Bronars (1987), in order to handle the presence of a high proportion of zeros in the observed consumption. In particular, we consider a uniform distribution on possible nonzero consumption outcomes and a positive mass on zero consumption that corresponds to the observed fraction of zeros.

We define $\mathbb{F}_i$ by the following sampling algorithm: (i) Compute the total fraction s_i of zero-consumption entries in all observations in data set O_i^K. (ii) Draw randomly K budget-share vectors $\hat{b}^t \in [0,1]^L$ ($\sum_{l=1}^{L} \hat{b}_l^t = 1$) of dimension L such that the probability of the total fraction of zeros is the empirical median of $(s_i)_{i=1}^{n}$. (iii) Generate a draw from $\mathbb{F}_i$, using $\hat{x}_l^t = (p^t x^t)\hat{b}_l^t / p_l^t$ and $\hat{p}^t = p^t$ for $(p^t, x^t) \in O_i^K$.

As we have already proposed, in step (ii) we modify BC for the cases where we observe an important fraction of zero consumption, such as the case of scanner consumer-panel data sets. Without this modification the probability of observing zero consumption is zero for the standard BC method. Without this change, the BC procedure will overestimate the size of the area of any model m and thereby underestimate its predictive-success measure. In particular, we use the same uniform distribution as in BC; we then truncate negative realizations to zero to obtain well-behaved budget-shares with zero consumption.[15]

[15] The distribution in BC is uniform, we modify its support to induce a fraction s of zero consumption levels that mimics the observed individual data.

Data

We compute r^m, a^m, and β^m in a well-known experimental data set due to Ahn et al. (2014), and in the scanner consumer panel data set used in Echenique et al. (2011). The main reason for using two different data sets is to combine the strengths of the different data sources in order to obtain robust conclusions about the empirical success of the AE condition. Experimental data sets have high price variation and therefore have considerable power to detect model inconsistencies. However, experimental conclusions are usually suspected of lacking external validity. To attenuate this concern, we also consider a scanner consumption panel. In this type of field data, consumers are making real-life decisions with moderate stakes. Nonetheless, scanner data sets are less powerful in terms of the probability of detecting a model inconsistency because they typically have less price variation than experimental data sets do.

Experimental Data.– Here we provide a quick overview of the experiment conducted by Ahn et al. (2014). The data set is the outcome of 154 subjects solving 50 independent portfolio choice problems. The experimental subjects are presented with three states of the world $l \in \{1, 2, 3\}$. For each state, the subjects can buy an Arrow security that pays 1 token in state l and nothing in other states. The probability of state 2 is $\pi_2 = \frac{1}{3}$, while the other states occur with unknown probabilities, $\pi_1 + \pi_3 = \frac{2}{3}$. Let x_l be the demand for each state-contingent Arrow security and p_l its price. The subject must choose any $x \geq 0$ that satisfies Walras' law. The subjects are paid according to the probability of each state at each round. The effective choice set is $X = \mathbb{R}^3_+$. Any given individual produces a sequence of observations: $O^{50} = \{p^k, x^k\}_{k=1}^{50}$ (where we omit the index i for each individual).

Consumer Scanner Data.– We use the Stanford Basket Data Set household-level scanner panel data set which contains grocery expenditure data for 494 urban households from four grocery stores in an urban area of a large U.S. Midwestern city. There are 26 time periods, each of them corresponding to a 4-week interval over a span of 2 years. We use an aggregated version of this data set prepared by Echenique et al. (2011) that contains 374 good categories. The good categories are brand-aggregated weighted consumption of the following food categories: bacon, BBQ sauce, butter, cereal, coffee, cracker, eggs, ice cream, nuts, frozen pizza, salty snacks, sugar, and yogurt. Any given household produces a sequence of observations $O^{26} = \{p^k, x^k\}_{k=1}^{26}$.

Results

Our main finding is that, as measured by the Selten index, AE has better predictive success than either GARP or WARP in both the experimental and the scanner data sets. Evidently, this finding empirically supports the importance of the new AE condition and lends credence to price-dependent utility maximization as a possible model that can rationalize consumer data from both the laboratory and the field.

Notice that even when AE by definition will have a higher hit rate than GARP, its area is

	AE	WARP	GARP
Experimental Data Set			
r^m	0.577922	0.12987	0.12987
a^m	0.000017	0	0
β^m	0.577905	0.12987	0.12987
Scanner Data Set			
r^m	0.870445	0.200405	0.198381
a^m	0.665749	0.0240081	0.0245344
β^m	0.204696	0.176397	0.173846

Table 1 – Predictive Success of AE, WARP, and GARP.

typically bigger that the area corresponding to GARP. Nonetheless, the very high hit rate of AE, close to 58 percent for the experimental data set and 87 percent for the scanner data, more than compensate, for the increase in the area. In particular, for the experimental environment, which by design has a considerable power to detect model inconsistencies, the area of AE, WARP, and GARP is close to 0. In this case, the hit rate is almost equal to the Selten index. The AE does very well with 58 percent Selten index, which is very high when compared to GARP and WARP, with 13 percent Selten index. Perhaps surprisingly, WARP has a predictive success which is exactly equivalent to that of GARP in the experimental data set and which is very close to that of GARP in the scanner data.

Notice that in the case of the scanner data set, the size or the area associated with AE is relatively higher than the area associated with GARP (66.5 percent and 2.4 percent respectively). In this case, the Selten index differs substantially from the hit rate, but the predictive success of AE is still higher than the predictive success of GARP by almost 3 percent.

The lower predictive success of AE in the scanner data in comparison with its performance in the experimental data is partially explained by the fact that price variation in the field is generally smaller than in the laboratory. The fact that the hit rate of AE is still very high in the experimental data set provides a strong hint that the new AE condition is empirically important. The details of the hit rate, area, and Selten index for AE, WARP, and GARP are provided in table 1. However, we must emphasize that GARP is implied by AE and WARP taken together, which means that it is important to decompose the sources of the predictive success of GARP into those for AE and WARP. We do so in the next section and find that AE accounts for the majority of the predictive power of GARP in both the experimental and scanner data sets.

More important, the high hit rate of the AE shows that our new condition could be the benchmark for developing new special cases of the price-dependent utility maximization model that we have studied. These special cases, would have better predictive power by being less permissive than the general model, but with the benefit of allowing the empirically relevant context dependence of choice.

6. Additively Decomposing the Predictive Success of GARP into AE versus WARP

We know that consistency with GARP is equivalent to the joint consistency of AE and WARP. When we use our notation for the measure of predictive success, the subset of possible outcomes predicted by GARP is equivalent to the intersection of the subsets predicted by AE and by WARP (i.e., $S^{GARP} = S^{AE} \cap S^{WARP}$). This means that the Selten index β^{GARP} depends on the Selten indices for both WARP and AE. Given this, we seek to measure the marginal contribution of AE and WARP to the predictive success of GARP.

Definition 17. (Marginal Contribution of the Behavioral Axioms AE and WARP to β^{GARP}) The marginal contributions of the behavioral axiom $m \in \{AE, WARP\}$ are: (i) $\Delta_m \beta^{GARP} = \beta^m$ when no additional condition is imposed, (ii) $\Delta_{m,m'} \beta^{GARP} = \beta^{GARP} - \beta^{m'}$ when $m' \neq m$ and m' is already imposed.

The marginal contribution of AE when no condition was previously imposed is $\Delta_{AE} \beta^{GARP} = \beta^{AE}$, and the marginal contribution when WARP is already imposed is the residual predictive success measure $\Delta_{AE,WARP} \beta^{GARP} = \beta^{GARP} - \beta^{WARP}$. Similarly for WARP, its marginal contributions are: $\Delta_{WARP} \beta^{GARP} = \beta^{WARP}$ and $\Delta_{WARP,AE} \beta^{GARP} = \beta^{GARP} - \beta^{AE}$. We follow Shapley (1953), in order to consolidate the two marginal contributions into a single index, which is an additive decomposition of the predictive success of GARP into that for AE and WARP.

Definition 18. (Shapley Value) The Shapley value of the behavioral axiom $m \in \{AE, WARP\}$ is given by $\text{Sh}^m = \frac{1}{2}[\Delta_m \beta^{GARP} + \Delta_{m,m'} \beta^{GARP}]$ for $m \in \{AE, WARP\}$.

The Shapley value of the behavioral axiom $m \in \{AE, WARP\}$ is the average marginal contribution of the axiom to the success measure of GARP. This index is the only one that satisfies the following set of desirable properties. (i) Monotonicity in marginal contributions: if there are two data sets each associated with β_1^{GARP} and β_2^{GARP}, and $\Delta_m \beta_1^{GARP} \geq \Delta_m \beta_2^{GARP}$ and $\Delta_{m',m} \beta_1^{GARP} \geq \Delta_{m',m} \beta_2^{GARP}$, then $\text{Sh}_1^m \geq \text{Sh}_2^m$ for all $m \in \{AE, WARP\}$. (ii) Equal treatment, which requires that if $\Delta_{WARP} \beta^{GARP} = \Delta_{AE} \beta^{GARP}$, then $\text{Sh}^{WARP} = \text{Sh}^{AE}$. (iii) Adding up, which requires that β^{GARP} be additively decomposable into the two values of the behavioral conditions $\beta^{GARP} = \text{Sh}^{AE} + \text{Sh}^{WARP}$. For our particular application, we obtain the following decomposition:

$$\text{Experimental data: } \beta^{GARP} = \text{Sh}^{AE} + \text{Sh}^{WARP} = 0.2889525 - 0.159082 = 0.12987,$$

$$\text{Scanner data: } \beta^{GARP} = \text{Sh}^{AE} + \text{Sh}^{WARP} = 0.101073 + 0.0727734 = 0.173846.$$

The marginal contributions of AE and WARP are provided in table 2.

For the experimental data set, the average marginal contribution of AE to the 13 percent predictive success of GARP is 29 percent. In stark contrast, the average marginal contribution of

m	AE	WARP
Experimental Data Set		
$\Delta_m \beta^{GARP}$	0.577905	0.12987
$\Delta_{m,m'} \beta^{GARP}$	0	-0.448035
Sh^m	0.2889525	-0.159082
Scanner Data Set		
$\Delta_m \beta^{GARP}$	0.204696	0.176397
$\Delta_{m,m'} \beta^{GARP}$	-0.00255085	-0.0308502
Sh^m	0.101073	0.0727734

Table 2 – Marginal Contributions to β^{GARP}, AE and WARP.

WARP is -16 percent. In fact, our analysis says that WARP on average decreases the predictive success of GARP by a considerable amount. Our results suggest, strikingly, that from a purely positive approach that focuses on increasing the predictive success of a consumer theory, we are better off by dropping WARP and using AE alone. For the scanner data set, the average marginal contribution made by AE to the 17 percent predictive success of GARP is 10 percent, while WARP contributes 7 percent. In this case, AE accounts for almost 60 percent of the amount of GARP's predictive-power.

7. Relationship with Other Models

7.1. Price-Dependent Utility Functions

Price-dependent utilities were studied early by Samuelson (1968) to model the demand for money (Basmann et al., 1987). However, it quickly became apparent that any data set can be rationalized by a price-dependent utility without further restrictions (Pollak, 1977, Shafer, 1974). Basmann et al. (1988) and Bagwell and Bernheim (1996) use variants of price-dependent utilities to explain conspicuous consumption, in the sense of Veblen. For this case, we may have a consumer choosing to pay a higher price for an equally functional object to signal her socioeconomic status. The dependence of the direct utility on price can allow the modeller to capture secondary utility effects of a change of price. There is also a small literature on general equilibrium with price-dependent preferences. The main work in this area is Balasko (2003), which not only proves the existence of equilibrium under mild regularity conditions, but also establishes that the main properties of general equilibrium regarding qualitative comparative statics, remain the same as in the price-independent case under some conditions.[16] Finally, price-dependent utility models have been shown to explain

[16]Some key assumptions are: continuity, strict monotonicity, and quasiconcavity in the first argument, and dependence on normalized prices (not relative prices). Our model shares the first two assumptions. Also, our price-dependent utility is concave in the first argument for the budget set. However, the monotonicity in the expenditure premium is a new property, and studying its role for general equilibrium economies would certainly be

violations of the law of demand, when consumers use prices to infer quality (Pollak, 1977, Martin, 1986). Our work contributes to this literature by providing a testable restriction on price-dependent utilities, the AE condition, which turns out to have a higher predictive success than GARP.

7.2. Menu and Reference-Dependence in Choice

A model with (exogenous) reference dependence is one in which the choices of a consumer are affected by the presence of a default alternative. All these models are special cases of Salant and Rubinstein (2008).[17] They study choice in the presence of frames in a very general setup. Their framework presents a situation in which there is an enriched data set, consisting of a choice set and a frame. The frame does not change preferences but does influence choice.[18] Their main result is that any choice with a frame model (without any structure), is equivalent to rational choice when the frame is not observable.

Our framework certainly uses additional information present in the traditional consumer environment, in particular, prices and expenditures. (That is, to obtain the reduced data set, we compute the budget constraint using the observed price and choice $B^t \subseteq X$ such that $B^t = \{x \in X | p^t x \leq p^t x^t\}$; we then preserve only the information $\{B^t, x^t\}_{t=1}^{K}$.) However, our framework is not a special case of the choice with frames rule considered in Salant and Rubinstein (2008), because even if we eliminate the price and expenditure information, we still would observe violations of WARP, if the data set is generated by a price-dependent utility. The cardinal information in our environment is of a different nature altogether with respect to the frames considered in their work.

Alternatively, the presence of some alternative in the choice set may alter the decision rule that the consumer follows; we usually call this the default alternative. Our price-dependent utility model differs conceptually from this class of models in that there is no single consumption bundle that acts as a reference. Rather, preferences change continuously in prices. In this regard, our price-dependent utility model is closer to the menu-dependent decision making procedures. However, our price-dependent utility model may predict that preferences are different even if the choice set is the same. Consider the case in which we have two choice sets or budget constraints given by $B_1 = \{x \in X | p^1 x \leq w^1\}$ and $B_2 = \{x \in X | p^2 x \leq w^2\}$. Furthermore, assume that the second price-wealth pair is a scaled-up version of the first pair: $(p^2, w^2) = (\lambda p^1, \lambda w^1)$ for some $\lambda > 0$. Then, our model allows the predicted choice to be different in each of these two experiments; in particular, our model does not imply that demand be homogeneous of degree zero. The reason is that in general we do not require that $u(x, p^1) = u(x, p^2)$; this means that even if the choice set is the same in both situations, preferences can change due to the scaling of prices. This also means

an interesting avenue for future research.

[17]See the work of Kahneman and Tversky (1979) on prospect theory, Masatlioglu and Ok (2005) on the role of status-quo bias on consumption, and recently DellaVigna et al. (2017) on the role of reference-dependence on job search.

[18]For example, in a search and satisficing framework, the frame will be the list from which the decision maker is choosing her satisficing alternative according to a threshold rule that depends on her preferences and the order of the list.

that our model introduces a different type of dependence on the menu, since it allows the consumer to change her preferences depending on the price alone.

On the other hand, there are conceptual similarities between the different works studying (endogenous) reference dependence, and in particular between the work of Ok, Ortoleva and Riella (2015) and our work. Ok et al. (2015) present a theory of (endogenous) *revealed reference dependence* to explain violations of WARP. They define a new notion of revealed preference that works even in the presence of violations of WARP. They then require an acyclicity condition on this relation to characterize the empirical content of their theory. Our own definition of AE resembles this approach. Moreover, their work and ours both achieve the goal of making welfare analysis possible whenever WARP fails. On the other hand, their work differs from ours both in the environment and in the definition of revealed preferences. In particular, they do not use any cardinal information because their environment contains none. Not that their definition of revealed preferences is different from our AE condition.[19] Their representation further differs from ours in that theirs has a utility function, a reference map, and a correspondence that produces a set of items to which the consumer is attracted if she chooses a reference. In contrast, our representation is a price-dependent utility. In this sense, our work complements their treatment in that we also provide a framework for doing welfare analysis in the presence of violations of WARP for the classical consumer environment.[20]

Another important treatment of endogenous reference dependence in the consumer environment (for discrete choice sets) is provided by Kőszegi and Rabin (2006). In their model, the reference point is formed as a self-equilibrium for a consumer with rational expectations about her consumption. Interestingly, their model predicts that the willingness to pay for an alternative is increasing in the (expected) prices, thus connecting choice with price dependence. Moreover, this means that their model can accommodate violations of WARP. Despite this conceptual overlap, our model, environment, and methodological approach to the problem of price-dependence are very different.

Finally, in marketing, a model of consumption close to our price-dependent utility maximization is the model of Putler (1992) on consumption with reference price effects. That model predicts *kinked* demand behavior (for discrete choice) that changes their elasticity depending on a reference price, where the kink is produced. Cornelsen et al. (2016) develop a variant of the model by Putler (1992), for the classical consumer environment. This variant is still a special case of the rationality benchmark, since the kinked demands are related to a Slutsky substitution matrix that satisfies all the regularity properties required for integrability.

7.3. Quah's (2006) Regular Preferences, Binary Consistency and WARP

Quah (2006) shows that any demand function that satisfies a stronger version of WARP, which we call WARP-2 here, can be rationalized by preference functions that are "regular."

[19]Their definition of revealed preference says that if there is an alternative $y \in X$ such that $y = \mathbf{c}(\{x, y\})$ and $x = \mathbf{c}(\{x, y, z\})$ then z is revealed preferred to x, where $\mathbf{c} : 2^X \setminus \emptyset \to X$ is a choice function.

[20]A generalized treatment of endogenous reference dependence is provided in Masatlioglu and Ok (2013), they provide applications to the classic consumer environment but they do not exploit any cardinal information therein.

Axiom 4. *(Weak Axiom of Revealed Preference 2, WARP-2) A data set O^K satisfies WARP-2 if there is no pair s, t such that $x^t \succeq^{R,D} x^s$ and $x^s \succeq^{R,D} x^t$.*

WARP-2 rules out (strong) cycles of size 2; in particular, it rules out indifference, and therefore excludes demand correspondences. It is apparent that WARP-2 is stronger than WARP. Quah (2006) shows that a data set that satisfies WARP-2 can be rationalized by a preference function $r : X \times X \to \mathbb{R}$ that is "regular," when the data set O^K is infinite (i.e., it is generated by a demand function). We define $\boldsymbol{x} : \mathbb{R}^L_{++} \times \mathbb{R}_{++} \to X$ as the demand function associated with O^K such that $x^t = \boldsymbol{x}(p^t, p^t x^t)$ for all $t = 1, \cdots, K$.

Definition 19. (Regular Preference Function) We say that a preference function $r : X \times X \to \mathbb{R}$ is "regular," whenever $r(x, y) = min_{q \in Q(x)}[yq - 1] + max_{q^* \in Q(y)}[1 - xq^*]$ with $Q(x) = \{p \in \mathbb{R}^L_{++} : \boldsymbol{x}(p, 1) = x\}$.

It is evident that regular preference functions are sign-asymmetric; in fact, they are skew-symmetric, $r(x, y) = -r(y, x)$, a property pointed out by Shafer (1974) for representing preference relations which are not necessarily transitive and which may be incomplete. Our Theorem 3 provides a finite-data version of this result for WARP.

Other related work that imposes binary consistency but not transitivity is Gerasimou (2010); it shows that incomplete preferences can be represented by a form of preference functions and price-dependent utilities simultaneously. (Of course, theirs differs from ours, in that theirs will not satisfy expenditure monotonicity generically.)[21] Their axiomatization corresponds to a condition close to WARP for correspondences and equivalent to WARP-2 for demand functions.

Kihlstrom et al. (1976) show that in the case of infinite data sets (demand functions), WARP is equivalent to the negative semidefiniteness of the Slutsky substitution matrix. They conjecture also that there exists a nontransitive consumer that rationalizes such a demand function. Our contribution answers that conjecture for finite data sets in the affirmative, while Quah (2006) does so for the case of infinite data sets and for WARP-2.

John (2001) and Keiding and Tvede (2013) provide a finite data set rationalization notion by means of maximizing a preference function that is (i) sign-asymmetric, (ii) skew-symmetric, (iii) concave in the first argument, (iv) and convex in the second argument. Their notion of rationalization corresponds to a stronger condition than the WARP, called the Weighted Law of Demand. Keiding and Tvede (2013) provides an example in which a finite data set satisfies WARP and at the same time cannot be rationalized by means of maximizing a preference function that is has all properties described above. In the light of our result, we see that WARP alone is not enough to guarantee concavity in the first argument, convexity in the second argument, and skew-symmetry of the preference function in finite data sets.

[21] In particular, if we assume continuity of the preference function, we can obtain a price-dependent utility representation of our preference function for WARP by Gerasimou (2010).

7.4. Ville Axiom of Revealed Preference

Hurwicz and Richter (1979) provide the only previous work that proposes an axiomatic decomposition of GARP (and the Strong Axiom of Revealed Preference, SARP). (For infinite data sets or for a demand function where GARP and SARP coincide.) They propose an axiom called the Ville Axiom of Revealed Preference (VARP) which is logically independent from WARP, but that taken together with WARP, implies SARP for demand functions. Briefly, VARP is equivalent to the symmetry of the Slutsky substitution matrix function associated with the demand function. (That is, the typical entry of the Slutsky matrix function is $S_{ij}(p,w) = \frac{\partial \boldsymbol{x}_i(p,w)}{\partial p_j} + \frac{\partial \boldsymbol{x}_i(p,w)}{\partial w}\boldsymbol{x}_k(p,w)$ for $i = 1, \cdots, L$.) We use an example by Hurwicz and Richter (1979) to show that a demand function that satisfies VARP can be used to generate a finite data set that fails AE.

Example 4. Consider the demand function $\boldsymbol{x}_1(p,w) = \frac{p_1 w}{p_1^2+p_2^2}$, $\boldsymbol{x}_2(p,w) = \frac{p_2 w}{p_1^2+p_2^2}$. This demand system satisfies Walras' law, $p_1\boldsymbol{x}_1(p,w) + p_2\boldsymbol{x}_2(p,w) = w$. The demand function also satisfies VARP, which is equivalent to a symmetric Slutsky substitution matrix function $S(p,w) \in \mathbb{R}^{L\times L}$. Indeed, the Slutsky matrix is symmetric:

$$S(p,w) = \begin{pmatrix} \frac{p_2^2 w}{(p_1^2+p_2^2)^2} & -\frac{p_1 p_2 w}{(p_1^2+p_2^2)^2} \\ -\frac{p_1 p_2 w}{(p_1^2+p_2^2)^2} & \frac{p_1^2 w}{(p_1^2+p_2^2)^2} \end{pmatrix}.$$

However, we can generate a finite data set O^K such that $x^t = \boldsymbol{x}(p^t, w^t)$ for a fixed wealth level for each $t = 1, \cdots, K$. The generated price data set is: $p^1 = (26, 89)$, $p^2 = (68, 72)$, and $p^3 = (13, 9)$. The corresponding demand data is $x^1 = (\frac{1612}{8597}, \frac{5518}{8597})$, $x^2 = (\frac{289}{613}, \frac{306}{613})$, and $x^3 = (\frac{78}{125}, \frac{54}{125})$ (with the wealth levels implied by Walras' law). We have $x^2 \succeq^{E,D} x^1$, $x^1 \succeq^{E,D} x^3$, and $x^3 \succ^{E,D} x^2$, which is a violation of AE. Note that the demand function violates GARP as well.[22] In fact, this same data set fails WARP; for instance, it follows that $x^1 \succeq^{R,D} x^2$ and $x^2 \succ^{R,D} x^1$.

Thus, VARP is not comparable to AE. The previous example makes the simple point that imposing VARP on a demand function does not imply the AE property for all finite data sets generated by it. Conversely, the fact that AE is satisfied does not imply that VARP also holds, since AE does not imply the differentiability which is needed for VARP to hold.

The previous observations make it clear that our decomposition of GARP based on the AE condition differs from the decomposition of GARP (SARP) in Hurwicz and Richter (1979). Note also that Aguiar and Serrano (2017) perform a quantification of departures from rationality into WARP and VARP, finding evidence against VARP in experimental data sets. Here, in contrast, we provide evidence that AE is empirically more successful than WARP.

Other related works such as Sen (1971) and Nosratabadi (2017) are interested in decomposing axiomatically WARP, in a full data environment where all possible menus or choice sets are observed. In the classical consumer environment, WARP is different from GARP because we cannot observe

[22] This can be concluded from observing that $S(p,w)$ is not negative semidefinite; in particular, its eigenvalues are $\lambda_1(p,w) = 0$, and $\lambda_2(p,w) = \frac{w}{p_1^2+p_2^2} > 0$. This means that there is a finite data set generated by this demand function for which WARP also fails (Kihlstrom et al., 1976).

all possible menus. Note that a WARP decomposition is a GARP decomposition only in the full data set environment.[23] Nosratabadi (2017), decomposes the WARP into three conditions and then relates some of these conditions to reference dependent choice. His decomposition does not use any cardinal notion as the environment he considers has no information on prices and wealth. In addition, it is conceptually different since in his model, context dependent choice may depend on a referential alternative, while in our framework context effects are related to prices.

7.5. Price Revealed Preference

Deb et al. (2017) propose the price revealed-preference condition. They define the direct price revealed-preference relation in the following manner: $p^s \succeq_p (\succ_p) p^t$ if $p^s x^t \leq (<) p^t x^t$. If a bundle is cheaper at prices p^s than at prices p^t, the consumer prefers the former to the latter. Then they impose the following condition analogous to GARP, called the Generalized Axiom of Price Preferences (GAPP).

Axiom 5. *(Generalized Axiom of Price Preferences, GAPP) For all $n \geq 2$, if there is a chain $p^1, p^2, \ldots, p^n$ such that $p^1 \succeq^{E,D} p^2 \succeq^{E,D} \cdots \succeq^{E,D} p^n \succeq^{E,D} p^1$ it must be the case that $p^s \sim^{E,D} p^t$ for all $s, t \in \{1, \cdots, n\}$.*

First, we use an example from Deb et al. (2017) to show that AE does not imply GAPP.

Example 5. (AE does not imply GAPP) Assume we have two observations $p^1 = (2, 1)$, $x^1 = (4, 0)$ and $p^2 = (1, 2)$, $x^2 = (0, 1)$. By definition they cannot fail AE (it does not fail GARP either). However, it is clear that we have a violation of GAPP . In particular $p^2 x^1 < p^1 x^1$ $(p^2 \succ_p p^1)$ but at the same time $p^2 x^2 > p^1 x^2$ $(p^1 \succ_p p^2)$.

Now we show that GAPP does not imply AE.

Example 6. (GAPP does not imply AE) Consider the following three observations: $p^1 = (2, 1)$, $x^1 = (2, 1)$; $p^2 = (1, 4)$, $x^2 = (0, 2)$; and $p^3 = (1, 1)$, $x^3 = (2, 1)$. It satisfies GAPP since (i) $x^1(p^1 - p^2) = -1$ (not $p_2 \succ_p p_1$); (ii) $x^2 \cdot (p^2 - p^1) = 6$ $(p^1 \succ_p p^2)$; (iii) $x^2[p^2 - p^3] = 6$ $(p^3 \succ_p p^2)$; and (iv) $x^1[p^1 - p^3] = 2$ $(p^3 \succ_p p^1)$. However, this data set violates AE because: (i) $(p^1 + p^2)(x^1 - x^2) = 1$ $(x^1 \succ^E x^2)$; (ii)$(p^2 + p^3)(x^2 - x^3) = 1$ $(x^2 \succ^E x^3)$; and (iii) $(p^3 + p^1)(x^3 - x^1) = 0$ $(x^3 \sim^E x^1)$. This is a violation of AE since $x^1 \succ^E x^2 \succ^E x^3$ but $x^3 \sim^E x^1$.

Both examples illustrate that GAPP and AE are logically independent axioms. Moreover, AE is strictly weaker than GARP, while GAPP is not necessarily weaker than GARP.[24] Deb et al. (2017) find that GAPP is logically equivalent to a form of rationalization of finite data sets by means of maximizing an expenditure-dependent utility function. This new rationalization concept has some similarities with our price-dependent utility rationalization, but the empirical

[23] Budget sets are only a subset of all possible choice sets.

[24] In particular, for the experimental data set considered in our empirical application, GAPP and GARP coincide since the wealth of each experiment is 1.

content of each model is clearly different. Their work is focused more on relaxing the intrinsic separability of (i) the consumption within a certain category of commodities and, (ii) the rest of the unobserved consumption that GARP imposes, while our work is focused on relaxing the context-free decision-making assumption that GARP implies.

Deb et al. (2017) find support for a random-utility version of their model in a survey data set. In this regard, our empirical findings are not strictly comparable to theirs since we apply a deterministic test to experimental and scanner consumer-panel data set, which lends strong empirical support for AE (in the form of a high predictive success Selten index).

Other related work, by Cosaert, considers the possibility that a consumer not only cares about the quantity of a product but also cares about its value (which he calls *diamondness*). Cosaert provides a notion of expenditure-dependent rationalization. However, Cosaert does not provides an axiomatic characterization of his expenditure-dependent notion of rationalization. Instead, he provides Afriat-like inequalities for his model, and the model depends on a parameter which allows one to rationalize any data set. Cosaert chooses that parameter on the basis of the Selten index. Because of this feature, the work of Cosaert is not directly comparable to ours, since our AE condition provides an axiomatic characterization of our price-dependent utility notion of rationalization.

8. Conclusion

We have axiomatically decomposed GARP into two more primitive conditions, AE and WARP. Moreover, we have established that AE is logically independent from WARP. AE is equivalent to the rationalization of a data set by a price-dependent utility that is monotonic in the expenditure premium. On the other hand, WARP is equivalent to the rationalization of a data set by a sign-asymmetric and locally non satiated preference function. We have compared the absolute and relative empirical success of AE and WARP with the empirical success of GARP in both experimental and scanner consumer panel data sets. We find that AE is significantly more successful than both GARP and WARP.

References

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

Aguiar, V. H. and Serrano, R. (2017). Slutsky matrix norms: The size, classification, and comparative statics of bounded rationality. *Journal of Economic Theory*, 172:163 – 201.

Ahn, D., Choi, S., Gale, D., and Kariv, S. (2014). Estimating ambiguity aversion in a portfolio choice experiment. *Quantitative Economics*, 5(2):195–223.

Andreoni, J. and Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):737–753.

Bagwell, L. S. and Bernheim, B. D. (1996). Veblen effects in a theory of conspicuous consumption. *The American Economic Review*, 86:349–373.

Balasko, Y. (2003). Economies with price-dependent preferences. *Journal of economic theory*, 109(2):333–359.

Basmann, R. L., Molina, D. J., and Slottje, D. J. (1987). Price-dependent preferences and the fechner-thurstone direct utility function: An exposition. *Journal of Institutional and Theoretical Economics*, 143(4):568–594.

Basmann, R. L., Molina, D. J., and Slottje, D. J. (1988). A note on measuring veblen's theory of conspicuous consumption. *The Review of Economics and Statistics*, 70:531–535.

Beatty, T. K. and Crawford, I. A. (2011). How demanding is the revealed preference approach to demand? *American Economic Review*, 101(6):2782–95.

Blundell, R., Browning, M., and Crawford, I. (2008). Best nonparametric bounds on demand responses. *Econometrica*, 76(6):1227–1262.

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

Cherchye, L., Demuynck, T., and De Rock, B. (2017). Transitivity of preferences: when does it matter? *Theoretical Economics*.

Choi, S., Fisman, R., Gale, D. M., and Kariv, S. (2007). Revealing preferences graphically: an old method gets a new tool kit. *The American economic review*, 97(2):153–158.

Cornelsen, L., Mazzocchi, M., and Smith, R. (2016). Between preferences and references: On the estimation of asymmetric price elasticities. *Mimeo*.

Cosaert, S. Revealed preferences for diamond goods. *American Economic Journal: Microeconomics*, 7.

Dean, M. and Martin, D. (2016). Measuring rationality with the minimum cost of revealed preference violations. *Review of Economics and Statistics*, 98(3):524–534.

Deb, R., Kitamura, Y., Quah, J. K.-H., and Stoye, J. (2017). Revealed price preference: Theory and stochastic testing. *Mimeo*.

DellaVigna, S., Lindner, A., Reizer, B., and Schmieder, J. F. (2017). Reference-dependent job search: evidence from hungary. *The Quarterly Journal of Economics*, 132(4):1969–2018.

Demuynck, T. (2015). Statistical inference for measures of predictive success. *Theory and Decision*, 79(4):689–699.

Echenique, F., Lee, S., and Shum, M. (2011). The money pump as a measure of revealed preference violations. *Journal of Political Economy*, 119(6):1201–1223.

Fostel, A., Scarf, H. E., and Todd, M. J. (2004). Two new proofs of afriat's theorem. *Economic Theory*, 24(1):211–219.

Gerasimou, G. (2010). Consumer theory with bounded rational preferences. *Journal of Mathematical Economics*, 46(5):708–714.

Hurwicz, L. and Richter, M. K. (1979). Ville axioms and consumer theory. *Econometrica*, 47:603–619.

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

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, 47.

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*, 44:971–978.

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

Martin, R. E. (1986). On judging quality by price: price dependent expectations, not price dependent preferences. *Southern Economic Journal*, pages 665–672.

Masatlioglu, Y. and Ok, E. A. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, 121(1):1–29.

Masatlioglu, Y. and Ok, E. A. (2013). A canonical model of choice with initial endowments. *Review of Economic Studies*, 81(2):851–883.

Nosratabadi, H. (2017). Referential revealed preference theory.

Ok, E. A., Ortoleva, P., and Riella, G. (2015). Revealed p-reference theory. *American Economic Review*, 105(1):299–321.

Pollak, R. A. (1977). Price dependent preferences. *The American Economic Review*, 67(2):64–75.

Putler, D. S. (1992). Incorporating reference price effects into a theory of consumer choice. *Marketing science*, 11(3):287–309.

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

Salant, Y. and Rubinstein, A. (2008). Choice with frames. *The Review of Economic Studies*, 75(4):1287–1296.

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

Samuelson, P. A. (1968). What classical and neoclassical monetary theory really was. *The Canadian Journal of Economics*, 1(1):1–15.

Selten, R. (1991). Properties of a measure of predictive success. *Mathematical Social Sciences*, 21(2):153–167.

Sen, A. K. (1971). Choice functions and revealed preference. *The Review of Economic Studies*, 38(3):307–317.

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

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28):307–317.

Varian, H. R. (1983). Non-parametric tests of consumer behaviour. *The review of economic studies*, 50(1):99–110.

9. Appendix

9.1. Proof of Lemma 1

Proof. Suppose a data set O^K satisfies GARP. Clearly, letting $n = 2$ in the length of a chain, it satisfies WARP as well.

Furthermore, taking an arbitrary chain satisfying that $x^1 \succeq^{E,D} x^2 \succeq^{E,D} \cdots \succeq^{E,D} x^n \succeq^{E,D} x^1$, we also have that $x^1 \succeq^{R,D} x^2 \succeq^{R,D} \cdots \succeq^{R,D} x^n \succeq^{R,D} x^1$.

Then, by GARP, $x^s \sim^{R,D} x^t$ for all s, t in the set of indices $\{1, 2, \ldots, n\}$. This implies that $p^s[x^s - x^t] = 0$, for all indices in the set $\{1, 2, \ldots, n\}$. By GARP (in fact by WARP) it cannot be that $x^t \succ^{R,D} x^s$, then $p^t[x^t - x^s] \leq 0$ for all the set of indices $\{1, 2, \ldots, n\}$. This implies that $\rho_s(x^t) = 0$, and $\rho_t(x^s) = 0$, for all the set of indices $\{1, 2, \ldots, n\}$. Then, by definition it must be that $x^s \sim^{E,D} x^t$ for any s, t in the set $\{1, 2, \ldots, n\}$. That is, the data set satisfies Acyclic Enticement.

Conversely, to show that WARP and Acyclic Enticement imply GARP, we argue by contradiction. If the data set O^K violates GARP, there exists a finite chain $x^1, x^2, \ldots, x^n$ such that $x^1 \succeq^{R,D} x^2 \succeq^{R,D} \cdots \succeq^{R,D} x^n \succ^{R,D} x^1$. Now, $x^1 \succeq^{R,D} x^2$ means $p^1x^1 \geq p^1x^2$. By WARP, we know that $p^2x^2 \leq p^2x^1$. Hence, $p^1[x^1 - x^2] - p^2[x^2 - x^1] \geq 0$, which implies that $x^1 \succeq^{E,D} x^2$. Repeating the same argument for the other links in the chain, we can prove that $x^2 \succeq^{E,D} x^3$, and so on, all the way to $x^{n-1} \succeq^{E,D} x^n$, and given that $p^n[x^n - x^1] > 0$ because $x^n \succ^{R,D} x^1$ and $p^1[x^1 - x^n] \leq 0$ by WARP, it follows that $x^n \succ^{E,D} x^1$, but then Acyclic Enticement would be violated. ■

9.2. Proof of Theorem 2

Proof. As asserted before the statement, the implication (2) $\implies$ (1) can be trivially established. Thus, we shall prove now that (1) $\implies$ (2).

First we need some preliminary definitions.

Definition 20. (Cyclical Consistency) A square matrix A of dimension n is cyclically consistent if $a_{jj} = 0$ for every $j \in \{1, \ldots, n\}$, and for every chain $\{k, j, l, \ldots, m\} \subset \{1, \ldots, n\}$, $a_{kj} \leq 0$, $a_{jl} \leq 0, \ldots, a_{mk} \leq 0$ implies that $a_{kj} = a_{jl} = \ldots = a_{mk} = 0$.

We build a matrix A with entry $a_{t,s} = max(p^s[x^s - x^t], 0) - p^t[x^t - x^s]$ for $t, s \in \{1, \ldots, n\}$, notice that $a_{t,s} = \rho_s(x^t) - \rho_t(x^s)$ when $\rho_t(x^s) \geq 0$ ($x^t \succeq^{R,D} x^s$) which ensures comparability. Moreover, this implies that if $a_{t,s} \leq 0$ it must be that $a_{t,s} = \rho_s(x^t) - \rho_t(x^s)$.

Lemma 4. *A data set O^n with n data points satisfies Acyclic Enticement if and only if the matrix A of order (n, n) is cyclically consistent.*

Proof. Suppose that A is cyclically consistent and the data set O^n fails Acyclic Enticement. This implies that there are indices $\{k, j, l, \ldots, h, m\}$ such that $x^k \succeq^{E,D} x^j, x^j \succeq^{E,D} x^l, \cdots, x^h \succeq^{E,D} x^m, x^m \succeq^{E,D} x^k$ hold, and that $x^m \succ^{E,D} x^k$.

From here, we know that $a_{kj} \leq 0, a_{jl} \leq 0, \ldots, a_{hm} \leq 0$. To see this, note that $x^k \succeq^{E,D} x^j$ implies that $a_{kj} = \rho_j(x^k) - \rho_k(x^j) \leq 0$ when x^k, x^j are comparable, and the same is true for the other adjacent pairwise comparisons in the set of indices. Since Acyclic Enticement fails, $\rho_m(x^k) > \rho_k(x^m)$, with x^k, x^m comparable.

Thus, $\{k, j, l, \cdots, h, m\}$ would be a chain satisfying the conditions of cyclical consistency, but $a_{mk} = \rho_k(x^m) - \rho_m(x^k) < 0$, and hence cyclical consistency of A is violated. We have therefore shown that the data set O^n satisfies Acyclic Enticement if A is cyclically consistent.

Conversely, if the data set O^n satisfies Acyclic Enticement, we construct the matrix A as before. First we notice that $a_{jj} = 0$ by definition for $j \in \{1, \cdots, n\}$. Second, take a chain $\{k, j, l, \cdots, h, m\}$ such that $a_{kj} \leq 0, a_{jl} \leq 0, \cdots, a_{hm} \leq 0, a_{mk} \leq 0$. For any two adjacent elements in the chain say m, k we have $x^m \succeq^E x^k$, by Acyclic Enticement this means that there is a chain $x^k \succeq^{E,D} x^j \succeq^{E,D} x^l \cdots x^h \succeq^{E,D} x^m \succeq^{E,D} x^k$, it must be that $x^t \sim^{E,D} x^s$ for all $t, s \in \{k, j, l, \cdots, h, m\}$, by definition this implies that $p^t[x^s - x^t] \leq 0$ and $\rho_t(x^s) = \rho_s(x^t)$, hence

$a_{t,s} = 0$ for all $t, s \in \{k, j, l, \cdots, h, m\}$. We conclude that A is cyclically consistent when the data set O^n satisfies Acyclic Enticement. ■

Now we present a technical lemma which was proved by Fostel, Scarf, and Todd (2004) (Section 2, 3).

Lemma 5. *If a square matrix A of dimension n is cyclically consistent, there exist positive real numbers* $(V^t, \lambda^t)_{t=1}^n$*, such that for all* $t, s = 1, \cdots, n$*,*

$$V^t \leq V^s + \lambda^s a_{st}.$$

We are now ready to prove that (1) $\implies$ (2) in the statement of the theorem.

Using the numbers in the preceding lemma, we define the following price-dependent utility function $u : X \times P \mapsto \mathbb{R}$:

$$u(x, p) = min_i\{V^i + \lambda^i a_i(x, p)\},$$

where $i = 1, \cdots, n$, $a_i(x, p) = max(p[x - x^i], 0) - p^i[x^i - x]$.

Now we prove that the utility so defined generates the data set O^n and that is monotonic on the expenditure premium.

First we establish the following Claim. If O^n satisfies Acyclic Enticement, for all $t = 1, \cdots, n$ it must be that for the price-dependent utility defined above $u(x^t, p^t) = V^t$.

Proof. We prove that $min_i\{V^i + \lambda^i a_i(x^t, p^t)\} < V^t$ is impossible. Indeed, notice that for the observation k that achieves the minimum, we have $V^k + \lambda^k a_{k,t} < V^t$, because, by construction, $a_k(x^t, p^t) = a_{k,t}$. but this inequality contradicts Acyclic Enticement (by Lemma 5).

This means that $min_i\{V^i + \lambda^i a_i(x^t, p^t)\} = V^t$ because at $i = t$ $a_t(x^t, p^t) = 0$. ■

Now we prove that the price-dependent utility defined above generates the data set O^n (i.e., $u(x^t, p^t) \geq u(x, p^t)$ for all $p^t x^t \geq p^t x$).

Lemma 6. *If* O^n *satisfies Acyclic Enticement,* $u(x, p^t) = min_i\{V^i + \lambda^i a_i(x, p^t)\} \leq u(x^t, p^t)$ *if* $p^t x \leq p^t x^t$ *(i.e,* O^n *is rationalized by the price-dependent utility* $u(\cdot, \cdot)$ *defined above). Moreover,* $u(\cdot, \cdot)$ *is continuous, monotone in the expenditure premium, and strictly increasing in the first argument.*

Proof. By construction, $u(x, p^t) \leq V^t + \lambda^t a_t(x, p^t) = V^t + \lambda^t[p^t x - p^t x^t]$ when $p^t x^t \geq p^t x$.

We note that $a_t(x, p^t) \leq 0$ if $p^t[x - x^t] \leq 0$. This follows by the definition of $a_i(x, p)$:

$a_t(x, p^t) = max(p^t[x - x^t], 0) - p^t[x^t - x] = p^t[x - x^t]$ if $p^t[x - x^t] \leq 0$.

Then, given that $\lambda^t > 0$, it must be that $u(x, p^t) \leq V^t$. By Claim 9.2, we conclude that $u(x, p^t) = min_i\{V^i + \lambda^i a_i(x, p^t)\} \leq u(x^t, p^t)$ if $p^t x \leq p^t x^t$. Then O^n is rationalized by the price-dependent utility $u(\cdot, \cdot)$ defined above.

The expenditure-premium monotonicity of the price-dependent utility follows immediately from Claim 9.2 and the fact that $u(\cdot, \cdot)$ rationalizes a data set O^n that satisfies Acyclic Enticement. That is, if $\rho_t(x^s) \geq \rho_s(x^t)$ and x^t, x^s are comparable, $V^t \geq V^s$, or equivalently, $u(x^t, p^t) \geq u(x^s, p^s)$ (the same holds for strict inequalities).

Finally, notice that

$$u(x, \overline{p}) = min_i\{f_i(x, \overline{p})\},$$

where for all $i = 1, \cdots, n$, $f_i(x, \overline{p}) = V^i + \lambda^i a_i(x, \overline{p})$, where each $f_i(x, \overline{p})$ is a continuous and strictly increasing function on the argument x for a fixed $\overline{p}$. Then, the minimum of a set of continuous and strictly increasing functions is also continuous and strictly increasing in the argument x. ■

■

9.3. Proof of Theorem 3

Proof. We begin with the simpler direction: (2) $\implies$ (1).

If a data set O^n is rationalized by a sign-asymmetric preference function $r : X \times X \mapsto \mathbb{R}$, it must be that $r(x^t, x) \geq 0$ for all $x \in X$ such that $p^t x^t \geq p^t x$. If for any x^t, x^s $p^t[x^t - x^s] \geq 0$ the rationalization of the data set by a sign-asymmetric preference function implies that (i) $r(x^t, x^s) \geq 0$, and (ii) $r(x^s, x^t) \leq 0$.

Assume WARP is violated. Then, $p^s[x^s - x^t] > 0$. This implies by rationalization by means of a preference function that $r(x^s, x^t) \geq 0$. If $r(x^s, x^t) > 0$, we find a contradiction to sign-asymmetry. If $r(x^s, x^t) = 0$, we also find a contradiction, because then by local nonsatiation, we could find $y \in B(x^t, \epsilon)$ for some small $\epsilon > 0$ such that $p^s x^s > p^s y$ with $r(x^s, y) < 0$, contradicting that the preference function rationalizes the choices.

First, we define a matrix A, with entry $a_{ij} = min(p^i(x^i - x^j), 0)$ if x^i and x^j are comparable. Let $a_{ij} = 0$ otherwise. Note that the comparability relation is symmetric, hence if $a_{ij} = min(p^i(x^i - x^j), 0)$ then $a_{ji} = min(p^j(x^j - x^i), 0)$.

We move to the opposite implication: (1) $\implies$ (2).

First, we begin with the following claim: If a data set O^n satisfies WARP, then there are real numbers w_{ij} such that: (i) $w_{ij} + w_{ji} \leq 0$, (ii) if $w_{ij} \geq 0$ then $w_{ji} \leq 0$, and:

$$w_{ij} \geq a_{ij}, \quad \forall i, j \in \{1, \ldots, n\}.$$

Proof. Let $w_{ij} = min(p^i(x^i - x^j), 0)$ if x^i and x^j are comparable, i.e., either $p^i x^i \geq p^i x^j$ or $p^j x^j \geq p^j x^i$. Let $w_{ij} = 0$ otherwise. Note that the comparability relation is symmetric, hence if $w_{ij} = min(p^i(x^i - x^j), 0)$ then $w_{ji} = min(p^j(x^j - x^i), 0)$.

If x^i, x^j are comparable, and without loss of generality, if $p^i(x^i - x^j) \geq 0$, we have that: (i) $w_{ij} = min(p^i(x^i - x^j), 0) = 0$, and by WARP, (ii) $p^j(x^j - x^i) \leq 0$. This means that the condition that requires that, if $w_{ij} \geq 0$ then $w_{ji} \leq 0$, holds. Note also that $w_{ij} + w_{ji} = p^j(x^j - x^i) \leq 0$.

If x^i, x^j are not comparable, we know that $w_{ij} = 0$, $w_{ji} = 0$, satisfy all the properties trivially. ■

Second, we extend the logic of the previous claim to bundles that have not been observed. This is done by building a matrix function $r_{ij} : X \times X \mapsto \mathbb{R}$ for all $i, j \in \{1, \ldots, n\}$. First define the function $a_i : X \mapsto \mathbb{R}$ for all $i \in \{1, \ldots, n\}$, $a_i(x) = -min(p^i(x^i - x), 0)$ if $x = x^j$ for some $j \in \{1, \cdots, n\}$ and such that x^i, x^j are comparable. Let $a_i(x) = 0$ if $x = x^j$ for some $j \in \{1, \cdots, n\}$ but such that x^i, x^j are not comparable. Finally, $a_i(x) = p^i(x - x^i)$, whenever $x \neq x^j$ for any $j \in \{1, \cdots, n\}$.

Now, we define $r_{ij}(x, y) = w_{ij} + a_i(x) - a_j(y)$, observe that $r_{ji}(y, x) = w_{ji} + a_j(y) - a_i(x)$. Then, indeed: If O^n satisfies WARP, then for any $i, j \in \{1, \ldots, n\}$, one has that if $r_{ij}(x, y) \geq 0$, $r_{ji}(y, x) \leq 0$. Moreover, $r_{ij}(x, y) + r_{ji}(y, x) \leq 0$.

Proof. If $r_{ij}(x, y) = w_{ij} + a_i(x) - a_j(y) \geq 0$, it follows that $r_{ji}(y, x) = w_{ji} + a_j(y) - a_i(x) \leq w_{ji} + w_{ij} \leq 0$, where the last inequality follows because O^n satisfies WARP, by Claim 9.3.

Furthermore, by definition, $r_{ij}(x, y) + r_{ji}(y, x) = w_{ij} + w_{ji} \leq 0$, where again the inequality follows from Claim 9.3. ■

Third, we build a candidate preference function $\hat{r} : X \times X \mapsto \mathbb{R}$. First let $\lambda, \mu \in \Delta$ be elements of the $n - 1$ unit simplex, and let $\hat{r}(x, y) = min_{\lambda \in \Delta} max_{\mu \in \Delta} \sum_{i,j} \lambda_i \mu_j r_{ij}(x, y) = max_{\mu} min_{\lambda} \sum_{i,j} \lambda_i \mu_j r_{ij}(x, y)$, (where the last equality is a consequence of the minimax Theorem in linear programming). If O^n satisfies WARP, then $\hat{r}$ is such that if $\hat{r}(x, y) \geq 0$ then $\hat{r}(y, x) \leq 0$.

Proof. If $\hat{r}(x, y) \geq 0$, then $-\hat{r}(x, y) = min_{\mu} max_{\lambda} \sum_{i,j} \mu_j \lambda_i (-r_{ij}(x, y)) \leq 0$. Using the "moreover" part of Claim 9.3, the matrix function $r_{ji}(y, x) \leq -r_{ij}(x, y)$.

By definition and the above observation, we must have $\hat{r}(y, x) = min_{\mu} max_{\lambda} \sum_{i,j} \mu_j \lambda_i r_{ji}(y, x) \leq min_{\mu} max_{\lambda} \sum_{i,j} \mu_j \lambda_i - r_{ij}(x, y) \leq 0$, as desired. ■

Fourth, we verify that if O^n satisfies WARP, then it is rationalized by the constructed preference function $\hat{r}$. If O^n satisfies WARP, then it must be that $\hat{r}(x^t, y) \geq 0$ for all $y \in X$ such that $p^t(y - x^t) \leq 0$ for all $t \in \{1, \ldots, n\}$.

Proof. By definition, for any $t \in \{1, \ldots, n\}$:

$$\hat{r}(x^t, y) = max_{\mu} min_{\lambda} \sum_{i,j} \lambda_i \mu_j r_{ij}(x^t, y)$$

$$\geq min_{\lambda} \sum_{i,j} \lambda_i \mu_j^t r_{ij}(x^t, y) = min_{\lambda} \sum_{i} \lambda_i r_{it}(x^t, y),$$

where $\mu^t \in \Delta$ is the element of the $n - 1$ simplex such that $\mu_j^t = 0$ if $j \neq t$ and $\mu_t^t = 1$. We want to show that $r_{it}(x^t, y) \geq 0$ for all $i \in \{1, \ldots, n\}$.

First, $a_t(y) = -min(p^t(x^t - y), 0) = 0$ whenever $y = x^j$ for some $j \in \{1, \ldots, n\}$ because $p^t x^t - p^t y \geq 0$. Second, $a_t(y) = p^t(y - x^t) \leq 0$ for $y \neq x^j$ for any $j \in \{1, \ldots, n\}$. Third, we compute $a_i(x^t) = -min(p^i(x^i - x^t), 0) = -w_{it} \geq 0$ if x^i, x^t are comparable, and $a_i(x^t) = 0 = -w_{it}$ when x^t, x^i are not comparable.

It follows that $r_{it}(x^t, y) = w_{it} + a_i(x^t) - a_t(y) \geq 0$ for all $i \in \{1, \ldots, n\}$ because of the inequalities in Claim 9.3, and because $-a_t(y) \geq 0$.

This implies that $\hat{r}(x^t, y) \geq min_\lambda \sum_i \lambda_i r_{it}(x^t, y) \geq 0$. ■

Finally, by the fact that $a_i(x)$ when x is not any of the bundles in the data is strictly increasing in x, the preference function $\hat{r}$ is also locally nonsatiated. Indeed, recall that, by definition, $\hat{r}(x, y) = min_{\lambda \in \Delta} max_{\mu \in \Delta} \sum_{i,j} \lambda_i \mu_j r_{ij}(x, y)$, by definition of $a_i(x)$ There is a $y' \in B(y, \epsilon)$ such that $y' \neq x^t$ for all $t \in \{1, \ldots, n\}$ such that $r_{ij}(x, y') < r_{ij}(x, y)$ for all $i, j \in \{1, \ldots, n\}$. Then $\hat{r}(x, y) < 0$. ■