Stata Journal

Peer-reviewed and accepted version

Testing Axioms of Revealed Preference in Stata

Marcos Demetry, Per Hjertstrand and Matthew Polisson

Published version:
https://doi.org/10.1177/1536867X221106374

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

Testing axioms of revealed preference in Stata

Marcos Demetry
Linnaeus University
Växjö, Sweden
marcos.demetry@lnu.se

Per Hjertstrand
Research Institute of Industrial Economics
Stockholm, Sweden
per.hjertstrand@ifn.se

Matthew Polisson
University of Bristol
Bristol, United Kingdom
matthew.polisson@bristol.ac.uk

Abstract. The revealed preference approach in economics is central to the empirical analysis of consumer behavior. This paper introduces the *Stata* commands `checkax`, `aei`, and `powerps` as a bundle within the package `rpaxioms`. The first command allows a user to test whether consumer expenditure data satisfy a number of revealed preference axioms; the second command calculates measures of goodness-of-fit when the data violate these axioms; and the third command calculates power against uniformly random behavior as well as predictive success for each axiom. The commands are illustrated using individual-level experimental data and household-level aggregate consumption data.

Keywords: rpaxioms, checkax, aei, powerps, revealed preference, GARP, Afriat efficiency index, power, predictive success

1 Introduction

The econometrics of consumer demand is central to economic analysis—it involves testing economic theories, making out-of-sample predictions, and drawing welfare comparisons across different environments and policy regimes. As such, the research program on empirical consumer demand has held a central position within the economics literature for many decades (see, for example, McFadden (1973) and Deaton and Muellbauer (1980)). Much of the emphasis within this literature has been on consistently estimating consumer preferences, the inherently unobservable primitive from which tests, predictions, and welfare statements can be derived.

The analogue to this research program within the context of finite data is known as *revealed preference*. The revealed preference approach involves checking whether a finite set of price and demand observations made on an individual consumer is compatible with economic rationality, i.e., *rationalizable* by some form of utility maximization. Revealed preference is fully *nonparametric*, in the sense that it does not impose any auxiliary functional form or distributional assumptions, only the basic primitives of utility maximization. Famously, Afriat's (1967) theorem states that a data set is rationalizable by the maximization of a well-behaved utility function *if and only if* it obeys an intuitive no-cycling condition on the data. This property is more commonly referred to as the *generalized axiom of revealed preference* (GARP), and there are efficient algorithms for checking this axiom in empirical practice (Varian 1982). Other notions (special cases) of rationalizability can also be characterized in terms of observables, and the approach has given rise to a suite of revealed preference tests in the tradition of Afriat (1967) that can be used in applied empirical work.

In general, revealed preference tests are "sharp", in the sense that they deliver a binary response as to whether observed expenditure data are compatible with an underlying behavioral model. However, given sufficiently rich data, an outright failure of even fairly permissive notions of rationalizability should not come as unexpected, and it may well be that the data are in fact very close to rationalizability. As Varian (1990) notes, "for most purposes, 'nearly optimizing' behavior' is just as good as 'optimizing' behavior". Afriat (1973) proposes to test for nearly optimizing behavior by allowing a part of the consumer's expenditure to be "wasted". The fraction of expenditure that is not being wasted by the consumer is usually referred to as the *efficiency level* of the test. Varian's (1982) original formulation of GARP implicitly assumes an efficiency level of 1, i.e., the consumer is not allowed to waste any part of her expenditure.

Varian (1982) introduces a simple combinatorial algorithm to test whether consumer expenditure data obey GARP. This algorithm can be easily adapted to test GARP at any efficiency level. Our first command, `checkax`, implements Varian's algorithm to test whether a data set satisfies GARP at any efficiency level specified by the user. The command also allows a user to test whether the data obey the following revealed preference axioms at any efficiency level: the strong axiom of revealed preference (SARP), the weak generalized axiom of revealed preference (WGARP), the weak axiom of revealed preference (WARP), the symmetric generalized axiom of revealed preference (SGARP), the homothetic axiom of revealed preference (HARP), and cyclical monotonicity (CM). These axioms and their behavioral implications are described in detail in Section 2.5.

Afriat (1973) proposes an upper limit on the efficiency level at which a data set satisfies GARP, or the *critical* cost efficiency, as a measure of approximate rationalizability. Hence, this index, called the Afriat efficiency index (AEI, also known as the critical cost efficiency index, CCEI), measures the severity of violations as the minimal expenditure adjustment that is required in order for the data to comply with GARP. As such, Varian (1990) interprets (and extends) this measure as a "goodness-of-fit" criterion. The approach can also be applied to other axioms, and our second command, `aei`, implements the AEI for each of the following seven axioms: GARP, SARP, WGARP, WARP, SGARP, HARP, and CM. The AEI is discussed in more detail in Section 2.3.

In addition to goodness-of-fit, the outcome of a revealed preference test in many empirical applications is often reported alongside some measure of power. The power of a revealed preference test, say for GARP, is defined as the probability of rejecting GARP, given that the data were generated from some type of "irrational" consumption behavior. Bronars (1987) proposes a power index where the irrational behavior is based on Becker's (1962) uniformly random consumption model. Thus, for this widely used power index, the choices generated by an irrational consumer are uniformly distributed on the frontiers of the budget sets. Our third command, `powerps`, implements the Bronars power index for any of the axioms above at any efficiency level. This command also reports a measure of "predictive success" originally introduced by Selten (1991) and adapted to the revealed preference framework by Beatty and Crawford (2011). This measure is motivated by the idea that if the data satisfy a given revealed preference axiom, then any robust conclusion on rationalizability should, at a minimum, require the test to have high power against uniformly random behavior. As such, the predictive success measure combines the pass rate of the revealed preference test with Bronars power index. Power and predictive success are further discussed in Section 2.4.

Finally, for applied practitioners it is imperative for revealed preference methods to be easily implementable and reproducible. To this end, we present the *Stata* package `rpaxioms`. We illustrate our three commands—`checkax`, `aei`, and `powerps`—on two types of data that are commonly used in empirical applications of revealed preference. First, using experimental data collected by Andreoni and Miller (2002), we test whether the social allocations selected by individual subjects are compatible with utility maximization taking a number of different forms. Second, using aggregate household consumption data on four food categories from Poi (2002), we test whether these data can be rationalized by preferences that are common across all households.[1]

2 Revealed preference

Suppose that there are T observations of the prices and quantities of K goods. At each observation $t = 1, \ldots, T$, the price vector is denoted by $\mathbf{p}^t = (p_1^t, \ldots, p_K^t) \gg 0$ and the quantity bundle by $\mathbf{x}^t = (x_1^t, \ldots, x_K^t) \geqslant 0$. We assume that all prices are strictly positive, and that all quantities are non-negative (note that some but not all quantities at any given observation may be equal to zero, i.e., all expenditures are strictly positive). The T observations of $(\mathbf{p}^t, \mathbf{x}^t)$ then form the finite data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$.

2.1 Rationalizability

The data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ is said to be *rationalizable* by utility maximization if there is a utility function $U : \mathbb{R}_+^K \to \mathbb{R}$, such that, at every observation $t = 1, \ldots, T$,

$$U(\mathbf{x}^t) \geqslant U(\mathbf{x}) \text{ for any } \mathbf{x} \in \{\mathbf{x} \in \mathbb{R}_+^K : \mathbf{p}^t \cdot \mathbf{x} \leqslant \mathbf{p}^t \cdot \mathbf{x}^t\}.$$

In words, rationalizability means that we can find a utility function defined on the consumption space which assigns (weakly) higher utility to the quantity bundle $\mathbf{x}^t$ than to any other bundle $\mathbf{x}$ which is affordable at the prevailing prices $\mathbf{p}^t$. Without any further restrictions on U, *any* data set $\mathcal{D}$ is rationalizable since U could simply assign the same level of utility to every bundle in the consumption space. For the question to be meaningful, we require further structure on the utility function U.

Afriat (1967) was the first to show that a finite data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ is rationalizable by a *locally nonsatiated* utility function U if and only if it obeys an intuitive

1. Poi (2002) uses the same data to illustrate the estimation of parametric demand systems in *Stata*.

property now known as the generalized axiom of revealed preference (GARP).[2] A data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ obeys GARP so long as the preference cycles it reveals are weak rather than strict, i.e., for any cycle represented by

$$\mathbf{p}^t \cdot \mathbf{x}^t \geqslant \mathbf{p}^t \cdot \mathbf{x}^u, \ \mathbf{p}^u \cdot \mathbf{x}^u \geqslant \mathbf{p}^u \cdot \mathbf{x}^v, \ \ldots \ , \ \mathbf{p}^w \cdot \mathbf{x}^w \geqslant \mathbf{p}^w \cdot \mathbf{x}^t,$$

the inequalities cannot be strict. The intuition of GARP as a no-cycling condition on the data ought to be strong, and it should also come as no surprise that GARP is necessary (or implied by) the maximization of a locally nonsatiated utility function. Afriat's (1967) theorem shows that GARP is also *sufficient* for rationalizability by a locally nonsatiated utility function U.[3] The importance of the result is that GARP completely characterizes the content of utility maximization in terms of observables, and can therefore be used as an empirical test for rationalizability.

2.2 Approximate rationalizability

In a sufficiently rich empirical setting, it is unlikely that *any* data set is exactly rationalizable, and so we require some notion of its *distance to* or *departure from* exact rationalizability. Loosely speaking, one could think of this as allowing for "error", which has long been the convention in the econometrics of consumer demand (see, for example, Deaton and Muellbauer (1980) and McFadden (1973)). To this end, the convention in the revealed preference literature is to accommodate error through *cost inefficiency*, an idea first developed by Afriat (1972, 1973).

According to Afriat's (1973) formulation, the consumer "has a definite structure of wants" and "programs at a level of cost-efficiency e", which is tantamount to relaxing the definition of rationalizability. Consider any efficiency level $e \in (0, 1]$. The data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ is said to be *rationalizable at efficiency level* e if there is a utility function $U : \mathbb{R}_+^K \to \mathbb{R}$, such that, at every observation $t = 1, \ldots, T$,

$$U(\mathbf{x}^t) \geqslant U(\mathbf{x}) \ \text{ for any } \ \mathbf{x} \in \{\mathbf{x} \in \mathbb{R}_+^K : \mathbf{p}^t \cdot \mathbf{x} \leqslant e\, \mathbf{p}^t \cdot \mathbf{x}^t\}.$$

When $e = 1$, this definition corresponds to *exact* rationalizability, and for any $e < 1$, to *approximate* rationalizability. Afriat (1973) shows that a data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ is rationalizable at efficiency level e if and only if, for any cycle represented by

$$e\, \mathbf{p}^t \cdot \mathbf{x}^t \geqslant \mathbf{p}^t \cdot \mathbf{x}^u, \ e\, \mathbf{p}^u \cdot \mathbf{x}^u \geqslant \mathbf{p}^u \cdot \mathbf{x}^v, \ \ldots \ , \ e\, \mathbf{p}^w \cdot \mathbf{x}^w \geqslant \mathbf{p}^w \cdot \mathbf{x}^t,$$

the inequalities cannot be strict. The equivalent condition known as eGARP is necessary and sufficient for approximate rationalizability by a locally nonsatiated utility function U.[4] When $e = 1$, eGARP and GARP coincide.

2. To be precise, Afriat (1967) refers to this property as *cyclical consistency*, which Varian (1982) shows is equivalent to GARP.
3. Furthermore, Afriat (1967) shows that such a U can be chosen to be continuous, strictly increasing, and concave; these properties, within the current context, have no empirical content. See also Diewert (1973) and Varian (1982) for proofs of this seminal result.
4. Once again, U can be chosen to be continuous, strictly increasing, and concave. See also Halevy et al. (2018) for a proof of this result.

The operationalization of eGARP as an empirical test is straightforward. Consider any efficiency level $e \in (0,1]$. For any pair of observations (t,s), we say that $\mathbf{x}^t$ is *directly revealed preferred to* $\mathbf{x}^s$ *at efficiency level* e, written $\mathbf{x}^t R_e^D \mathbf{x}^s$, if $e\,\mathbf{p}^t \cdot \mathbf{x}^t \geqslant \mathbf{p}^t \cdot \mathbf{x}^s$. This means that $\mathbf{x}^t$ is chosen even though the cost of the bundle $\mathbf{x}^s$ (at the prevailing prices $\mathbf{p}^t$) does not exceed $e\,\mathbf{p}^t \cdot \mathbf{x}^t$. Analogously, we say that $\mathbf{x}^t$ is *strictly directly revealed preferred to* $\mathbf{x}^s$ *at efficiency level* e, written $\mathbf{x}^t P_e^D \mathbf{x}^s$, if $e\,\mathbf{p}^t \cdot \mathbf{x}^t > \mathbf{p}^t \cdot \mathbf{x}^s$. We say that $\mathbf{x}^t$ is *revealed preferred to* $\mathbf{x}^s$ *at efficiency level* e, written $\mathbf{x}^t R_e \mathbf{x}^s$, if there exists a sequence of observations $(t,u,v,\ldots,w,s)$ such that $\mathbf{x}^t R_e^D \mathbf{x}^u, \mathbf{x}^u R_e^D \mathbf{x}^v, \ldots, \mathbf{x}^w R_e^D \mathbf{x}^s$. Hence, R_e is the transitive closure of R_e^D. When $e = 1$, these relations reduce to the usual revealed preference relations (Varian 1982).

A data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ satisfies eGARP if $\mathbf{x}^t R_e \mathbf{x}^s$ implies $e\,\mathbf{p}^s \cdot \mathbf{x}^s \leqslant \mathbf{p}^s \cdot \mathbf{x}^t$.

The eGARP condition can be tested at any efficiency level e by slightly modifying the algorithm proposed by Varian (1982). First, the relations R_e^D and P_e^D are formed by constructing the $T \times T$ matrices RD and PD, where the elements RD_{ts} and PD_{ts} are equal to 1 if $e\,\mathbf{p}^t \cdot \mathbf{x}^t \geqslant \mathbf{p}^t \cdot \mathbf{x}^s$ and $e\,\mathbf{p}^t \cdot \mathbf{x}^t > \mathbf{p}^t \cdot \mathbf{x}^s$, respectively, and 0 otherwise. Second, the relation R_e is formed by calculating the transitive closure of the matrix RD, which gives a $T \times T$ matrix RT with element RT_{ts} equal to 1 if $\mathbf{x}^t R_e \mathbf{x}^s$, and 0 otherwise. Varian (1982) suggests calculating RT using Warshall's (1962) algorithm. Finally, eGARP is violated whenever $RT_{ts} = 1$ and $PD_{st} = 1$ for any pair of observations (t,s). The total number of violations is given by the number of pairs (t,s), with $t \neq s$, such that $RT_{ts} = 1$ and $PD_{st} = 1$. Therefore, with T observations, the total possible number of eGARP violations is $T(T-1)$, and the fraction of violations is given by the ratio of the number of violations to $T(T-1)$.

Our first command, `checkax`, constructs RD, PD, and RT at any efficiency level e specified by the user. The latter is constructed using a vectorized version of Warshall's algorithm. The command `checkax` then reports to the user whether or not the data satisfy eGARP, as well as the number and fraction of violations.[5]

2.3 The Afriat efficiency index

It is clear that *any* data set is approximately rationalizable at some sufficiently small efficiency level $e \in (0,1]$. Afriat (1973) defines the *critical cost efficiency index* (CCEI) or the *Afriat efficiency index* (AEI) as the maximal value of e (the supremum, to be precise) such that a data set obeys eGARP.[6] Varian (1990) interprets the AEI as a measure of "goodness-of-fit" in terms of wasted expenditure: if a consumer has an AEI of $e^* < 1$, then she could have obtained the same level of utility by spending only the fraction e^* of what she actually spent. This is the sense in which the consumer is

5. Swofford and Whitney (1987) originally suggests using the number of violations as a goodness-of-fit measure, while Famulari (1995) proposes a related metric, which can roughly be interpreted as the fraction of violations.
6. Note that Varian (1990) extends Afriat's (1973) approach by allowing the amount of cost inefficiency to vary across observations, so that $\mathbf{e} = (e^1, \ldots, e^T) \in (0,1]^T$ represents a *vector* of efficiency levels; Halevy et al. (2018) discusses the different *aggregators* of $\mathbf{e}$, which include the Afriat (1973), Varian (1990), and Houtman and Maks (1985) indices.

exhibiting cost inefficiency, and in many applications the AEI is interpreted as a measure of “rationality” or “decision making quality” (see, for example, Choi et al. (2014)).[7]

Our second command, `aei`, calculates the AEI by implementing the binary search algorithm described in Varian (1990).

2.4 Power and predictive success

Within applications of revealed preference, alongside goodness-of-fit it is important to know something about the *power* of the test/empirical environment. To this end, the convention in the applied revealed preference literature is to test against an alternative behavioral model, which is typically random choice and interpreted as “naive” or “irrational”. The notion of “irrationality” which underpins the Bronars (1987) power index is based on a model of uniformly random consumption, in which all feasible consumption allocations (i.e., frontier bundles) are equally likely to be chosen.

Bronars (1987) suggests implementing the index using Monte Carlo methods. The first step consists of generating artificial budget shares that are consistent with uniformly random consumption. At each observation, this involves generating K random variables drawn from the Dirichlet distribution with all parameters (characterizing this distribution) set equal to one. By construction, at each observation, these random variables are uniformly distributed on the $(K-1)$-dimensional unit simplex, and consequently, can be interpreted as budget shares in the uniformly random model. The second step solves for each uniformly random consumption quantity (denoted by q_k^t) from the budget share equation given by $w_k^t = p_k^t q_k^t / \mathbf{p}^t \cdot \mathbf{x}^t$, where each w_k^t denotes an artificial budget share generated in the first step. (Notice that $\mathbf{p}^t$ and $\mathbf{x}^t$ are given in the original data set). Thus, the first two steps generate a synthetic data set across K goods and T observations that is compatible with uniformly random behavior. The third step repeats the first two steps many times, and for each repetition checks whether the synthetic data set of prices and uniformly random quantities satisfy eGARP at a given efficiency level e. The power measure is the fraction of these synthetic data sets which would then violate eGARP.

Our third command, `powerps`, calculates power at any efficiency level e and for any number of repetitions specified by the user. Moreover, the command allows the user to set the random seed in the generation of the Dirichlet random variables (in the first step of the procedure) in order to ensure that power calculations are perfectly replicable. The command `powerps` also reports Beatty and Crawford’s (2011) revealed preference measure of “predictive success”. For a given data set, this measure is defined as the difference between the pass/fail indicator and one minus the Bronars’ power index, where the pass/fail indicator takes the value 1 if the original data obey eGARP at a given efficiency level e, and 0 otherwise, and where the power index corresponding eGARP is calculated at the same efficiency level e. This measure of predictive success can then be straightforwardly aggregated across individual data sets.

7. For a precise description of the relationship between (approximate) rationalizability and (approximate) cost-rationalizability, which in turn motivates the AEI, see Polisson and Quah (2022).

2.5 Other axioms

Our commands are also implementable for other revealed preference axioms that characterize a number of the common special cases of utility maximization. The default axiom in every command is eGARP (with $e = 1$), but each command can also be executed for six other revealed preference axioms any efficiency level e specified by the user.

- (Strong axiom of revealed preference) A data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ satisfies the *strong axiom of revealed preference at efficiency level* e, abbreviated eSARP, if $\mathbf{x}^t R_e \mathbf{x}^s$ implies $e\,\mathbf{p}^s \cdot \mathbf{x}^s < \mathbf{p}^s \cdot \mathbf{x}^t$ whenever $x^t \neq x^s$. Matzkin and Richter (1991) shows that SARP ($e = 1$) is necessary and sufficient for rationalizability by a continuous, strictly increasing, and *strictly* concave utility function. Notice that the difference between GARP and SARP is that GARP allows for "flat spots" of indifference (demand correspondences versus demand functions). Like eGARP, there can be up to $T(T-1)$ violations of eSARP.

- (Weak generalized axiom of revealed preference) A data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ satisfies the *weak generalized axiom of revealed preference at efficiency level* e, abbreviated eWGARP, if $\mathbf{x}^t R_e^D \mathbf{x}^s$ implies $e\,\mathbf{p}^s \cdot \mathbf{x}^s \leqslant \mathbf{p}^s \cdot \mathbf{x}^t$. Aguiar et al. (2020) shows that WGARP ($e = 1$) is necessary and sufficient for rationalizability by a continuous, strictly increasing, piecewise concave, and skew-symmetric preference function (see Aguiar et al. (2020) for the definition of a preference function and its properties). Banerjee and Murphy (2006) shows that WGARP and GARP are equivalent when $K = 2$ (when there are two goods). The total possible number of violations of eWGARP is $T(T-1)/2$.

- (Weak axiom of revealed preference) A data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ satisfies the *weak axiom of revealed preference at efficiency level* e, abbreviated eWARP, if $\mathbf{x}^t R_e^D \mathbf{x}^s$ implies $e\,\mathbf{p}^s \cdot \mathbf{x}^s < \mathbf{p}^s \cdot \mathbf{x}^t$ whenever $x^t \neq x^s$. Aguiar et al. (2020) shows that WARP ($e = 1$) is necessary and sufficient for rationalizability by a continuous, strictly increasing, piecewise *strictly* concave, and skew-symmetric preference function. The difference between WARP and WGARP is analogous to the difference between SARP and GARP. Furthermore, Rose (1958) shows that WARP and SARP are equivalent when $K = 2$. Like eWGARP, there can be up to $T(T-1)/2$ violations of eWARP.

- (Symmetric generalized axiom of revealed preference) For any (t, s), we can modify the definition of R_e^D so that $\mathbf{x}^t R_e^D \mathbf{x}^s$ if $e\,\mathbf{p}^t \cdot \mathbf{x}^t \geqslant \mathbf{p}^t \cdot \mathbf{y}^s$, where $\mathbf{y}^s$ is *any permutation* of $\mathbf{x}^s$,[8] and where the transitive closure R_e of R_e^D follows accordingly. A data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ satisfies the *symmetric generalized axiom of revealed preference at efficiency level* e, abbreviated eSGARP, if $\mathbf{x}^t R_e \mathbf{x}^s$ implies $e\,\mathbf{p}^s \cdot \mathbf{x}^s \leqslant \mathbf{p}^s \cdot \mathbf{y}^t$ (where $\mathbf{y}^t$ is any permutation of $\mathbf{x}^t$). Nishimura et al. (2017) shows that eSGARP is necessary and sufficient for rationalizability by a continuous, strictly increasing, concave, and *symmetric* utility function. Polisson et al.

8. For example, if $\mathbf{x}^s = (3, 1, 2)$, then there are six permutations of $\mathbf{x}^s$: $(1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(2, 3, 1)$, $(3, 1, 2)$, and $(3, 2, 1)$.

(2020) implements eSGARP in the context of symmetric risk, i.e., the utility function must also obey first order stochastic dominance (FOSD). The total possible number of violations of eSGARP is T^2.

- (Homothetic axiom of revealed preference) A data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ satisfies the *homothetic axiom of revealed preference at efficiency level* e, abbreviated eHARP, if for all distinct sequences of observations $(s, t, u, \ldots, v)$, it must be the case that $(\mathbf{p}^t \cdot \mathbf{x}^s)(\mathbf{p}^s \cdot \mathbf{x}^u) \cdots (\mathbf{p}^v \cdot \mathbf{x}^t) \geqslant (e\,\mathbf{p}^t \cdot \mathbf{x}^t)(e\,\mathbf{p}^s \cdot \mathbf{x}^s) \cdots (e\,\mathbf{p}^v \cdot \mathbf{x}^v)$. Varian (1983) shows that HARP ($e = 1$) is necessary and sufficient for rationalizability by a continuous, strictly increasing, concave, and *homothetic* utility function. Heufer and Hjertstrand (2019) provide a more general characterization ($e < 1$), and refer to e^* in this case as the homothetic efficiency index (HEI). The command `checkax` implements eHARP as described in Varian (1983) using the Floyd-Warshall algorithm. The total possible number of violations of eHARP is T.

- (Cyclical monotonicity) A data set $\mathcal{D} = \{(\mathbf{p}^t, \mathbf{x}^t)\}_{t=1}^T$ satisfies a *cyclical monotonicity condition at efficiency level* e, abbreviated eCM, if for all distinct sequences of observations $(s, t, u, \ldots, v)$, it must be the case that $\mathbf{p}^t \cdot (\mathbf{x}^s - e\,\mathbf{x}^t) + \mathbf{p}^s \cdot (\mathbf{x}^u - e\,\mathbf{x}^s) + \cdots + \mathbf{p}^v \cdot (\mathbf{x}^t - e\,\mathbf{x}^v) \geqslant 0$. Brown and Calsamiglia (2007) shows that CM ($e = 1$) is necessary and sufficient for rationalizability by a continuous, strictly increasing, concave, and *quasilinear* utility function. The command `checkax` implements eCM in a similar manner to eHARP using the Floyd-Warshall algorithm. Like eHARP, there can be up to T violations of eCM.

We conclude this section with two comments.

First, notice that in general a data set is approximately rationalizable if it could have arisen from the maximization of *some* utility/preference function subject to a *modified* budget set. Explicit theoretical support for these notions of rationalizability have been developed in the case of eGARP, eSGARP, and eHARP, but not for the other axioms.

Second, we note that smoothness/differentiability has no material empirical content once cost inefficiency has been taken into account. For example, Chiappori and Rochet (1987) shows that Strong SARP (SSARP) is necessary and sufficient for rationalizability by an infinitely differentiable, strictly increasing, and strictly concave utility function. Suppose that a data set obeys SARP, but fails SSARP, which amounts to the same consumption bundle being chosen at two or more distinct price vectors. If we set the efficiency level to $1 - \varepsilon$, for some $\varepsilon > 0$ arbitrarily small, then we could always find a smooth rationalization. Since the CCEI is defined as a supremum, the CCEI for SSARP would still be equal to 1. In other words, smoothness/differentiability are "untestable" in a meaningful way. See also the discussion in Polisson et al. (2020).

3 Stata commands

Our commands `checkax`, `aei`, and `powerps` do not require any additional *Stata* packages. The commands are freely available as a bundle within the package `rpaxioms` on

SSC (Statistical Software Components) provided through Boston College and RePEc (Research Papers in Economics), and can be installed by entering "ssc install rpaxioms" in the Stata command prompt. The example dataset used in this section can be downloaded by entering "net get rpaxioms". Note that this downloads the dataset to your current working directory. All three commands take as their two main (required) arguments the $T \times K$ price and quantity matrices:

`price(`*string*`)` specifies a $T \times K$ price matrix, where each row corresponds to an observation t and each column to a good k. All prices are required to be strictly positive. If any of the elements in the price matrix are non-positive (or if the price and quantity matrices have different dimensions), the commands return an error message.

`quantity(`*string*`)` specifies a $T \times K$ quantity matrix, where each row corresponds to an observation t and each column to a good k. All quantities are required to be non-negative. Some (but not all) quantities at a given observation may be equal to zero. If the quantity matrix violates these conditions (or if the price and quantity matrices have different dimensions), the commands return an error message.

3.1 checkax

The syntax of `checkax` is:

`checkax, price(`*string*`) quantity(`*string*`)` [`axiom(`*string*`) efficiency(`#`)`]

Options

`axiom(`*string*`)` specifies the axiom(s) that the user would like to test. The default option is `axiom(eGARP)`. There are seven axioms that can be tested: `eGARP`, `eSARP`, `eWGARP`, `eWARP`, `eSGARP`, `eHARP`, and `eCM`. The user may also test all axioms simultaneously by specifying `axiom(all)`.

`efficiency(`#`)` specifies the efficiency level at which the user would like to test the axiom(s). The default option is `efficiency(1)`. The efficiency level must be strictly positive, and no greater than one.

Output and stored results

Running `checkax` produces a table with the following entries and return list:

`Axiom` returns the axiom(s) being tested. Given as the macro `r(AXIOM)` in `return list`.

`Pass` is a binary number indicating whether the data satisfy the axiom or not: `Pass=1` if the data satisfy the axiom and `Pass=0` if the data do not satisfy the axiom. Given as the scalar `r(PASS_`*axiom*`)` in `return list`.

`#vio` is the number of violations. Note that `#vio>0` if `Pass=0`, and `#vio=0` if `Pass=1`. Given as the scalar `r(NUM_VIO_`*axiom*`)` in `return list`.

`%vio` is the fraction of violations. Note that `%vio>0` if `Pass=0`, and `%vio=0` if `Pass=1`. Given as the scalar `r(FRAC_VIO_`*axiom*`)` in `return list`.

`Goods` is the number of goods. Given as the scalar `r(GOODS)` in `return list`.

`Obs` is the number of observations. Given as the scalar `r(OBS)` in `return list`.

`Eff` is the efficiency level of the test. Given as the scalar `r(EFF)` in `return list`.

Examples

The following examples illustrate the command `checkax` using a data set of 20 observations on the prices and quantities of five goods. The prices of goods 1 to 5 are `p1`, `p2`, `p3`, `p4`, and `p5`. The quantities are `x1`, `x2`, `x3`, `x4`, and `x5`. The price and quantity matrices are `P` and `X`, respectively. The first example runs `checkax` using its default options, i.e., for eGARP at the efficiency level $e = 1$. The second example runs `checkax` for eGARP and eHARP at the efficiency level $e = 0.95$.

```
. net get rpaxioms
. use rpaxioms_example_data.dta, clear
. mkmat p1-p5, matrix(P)
. mkmat x1-x5, matrix(X)
. checkax, price(P) quantity(X)
                    Number of obs           =      20
                    Number of goods         =       5
                    Efficiency level        =       1

          Axiom |       Pass       #vio       %vio
        --------+----------------------------------
          eGARP |          0        161      .4237

. checkax, price(P) quantity(X) axiom(eGARP eHARP) efficiency(0.95)
                    Number of obs           =      20
                    Number of goods         =       5
                    Efficiency level        =     .95

          Axiom |       Pass       #vio       %vio
        --------+----------------------------------
          eGARP |          0        104      .2737
          eHARP |          0         20          1
```

3.2 aei

The syntax of `aei` is:

`aei, price(`*string*`) quantity(`*string*`)` [`axiom(`*string*`) tolerance(`#`)`]

Options

axiom(*string*) is the same as in the command checkax.

tolerance(#) sets the tolerance level of the termination criterion 10^{-n} by specifying the integer n. For example, tolerance(10) sets the tolerance level to 10^{-10}. The default option is tolerance(6), which gives the default tolerance level 10^{-6}. The integer n in the termination criterion 10^{-n} cannot be less than 1 or greater than 18.

Output and stored results

Running aei produces a table with the following entries and return list:

Axiom, Goods, and Obs are the same as in the command checkax.

AEI is the AEI. Given as the scalar r(AEI_*axiom*) in return list.

Tol is the tolerance level of the termination criterion for the AEI calculation. Given as the scalar r(TOL) in return list.

Examples

The following examples illustrate the command aei using the same data as above. The first example runs aei using its default options, i.e., for *e*GARP with a tolerance level of 10^{-6}. The second example runs aei for *e*GARP and *e*HARP with the tolerance level set to 10^{-10}, and shows that the command quietly can be used to suppress the output.

```
. aei, price(P) quantity(X)
    Number of obs           =         20
    Number of goods         =          5
    Tolerance level         =    1.0e-06

          Axiom |        AEI
          eGARP |   .9055848

. quietly aei, price(P) quantity(X) axiom(eGARP eHARP) tolerance(10)
```

3.3 powerps

The syntax of powerps is:

powerps, price(*string*) quantity(*string*) [axiom(*string*) efficiency(#) simulations(#) seed(#) aei tolerance(#) progressbar]

Options

axiom(*string*) and efficiency(#) are the same as in the command checkax.

simulations(#) specifies the number of repetitions of the simulated uniformly random data. The default number of repetitions is simulations(1000).

seed(#) specifies the random seed in the generation of the Dirichlet random numbers. The default random seed is seed(12345).

aei specifies whether the user wants to compute the AEI for each simulated data set and specified axiom. The default option is that aei is *not* specified. Note that including this option may increase computation times substantially.

tolerance(#) sets the tolerance level of the termination criterion 10^{-n} by specifying the integer n when computing the AEI. See Section 3.2 for a more detailed description. This option is only useful in combination with the aei option.

progressbar specifies if the user wants to display the number of repetitions that have been executed. The default is that progressbar is *not* specified.

Output and stored results

Running powerps produces tables with the following entries and return list:

Axiom, Goods, and Obs are the same as in the command checkax.

Power is the power. Given as the scalar r(POWER_*axiom*) in return list.

PS is the predictive success. Given as the scalar r(PS_*axiom*) in return list.

PASS is a binary number indicating whether the actual data satisfy the axiom or not: Pass=1 if the actual data satisfy the axiom and Pass=0 if the actual data do not satisfy the axiom. Given as the scalar r(PASS_*axiom*) in return list.

AEI is the AEI corresponding to the actual data. Given as the scalar r(AEI_*axiom*) in return list.

Sim is the number of repetitions of the simulated uniformly random data. Given as the scalar r(SIM) in return list.

Eff is the efficiency level at which power and predictive success are computed. Given as the scalar r(EFF) in return list.

For each axiom being tested, the command also produces a table containing summary statistics over all simulated data with the following entries and return list:

#vio gives the mean (Mean), standard deviation (Std. Dev.), minimum (Min), first quartile (Q1), median (Median), third quartile (Q3), and maximum (Max) of the number of violations. Given as the matrix r(SUMSTATS_*axiom*) in return list.

%vio gives the mean (Mean), standard deviation (Std. Dev.), minimum (Min), first quartile (Q1), median (Median), third quartile (Q3), and maximum (Max) of the fraction of violations. Given as the matrix `r(SUMSTATS_`*axiom*`)` in `return list`.

`AEI` gives the mean (`Mean`), standard deviation (`Std. Dev.`), minimum (`Min`), first quartile (`Q1`), median (`Median`), third quartile (`Q3`), and maximum (`Max`) of the AEI. Given as the matrix `r(SUMSTATS_`*axiom*`)` in `return list`. This is only displayed if the option `aei` is specified. The tolerance level of the termination criterion in the AEI calculation is given as the scalar `r(TOL_`*axiom*`)` in `return list`.

For each axiom being tested, the matrix `r(SIMRESULTS_`*axiom*`)` in `return list` contains, for every simulated uniformly random data set, the number of violations, the fraction of violations, and the AEI (only if the option `aei` is specified).

Examples

The following examples illustrate the command `powerps` using the same data as above. The first example runs `powerps` for the axioms eGARP and eHARP. All other options are set to their defaults. The second example also runs `powerps` for the axioms eGARP and eHARP but now includes the option `aei`, which calculates the AEI for both axioms for each of the 1,000 simulated data sets. Note that including the `aei` option increases computation time substantially.

```
. powerps, price(P) quantity(X) axiom(eGARP eHARP)
                                   Number of obs          =        20
                                   Number of goods        =         5
                                   Simulations            =      1000
                                   Efficiency level       =         1

       Axioms |      Power         PS       Pass        AEI
--------------+-----------------------------------------------
        eGARP |       .995      -.005          0   .9055848
        eHARP |          1          0          0   .8449683

Summary statistics for simulations:

        eGARP |      #vio       %vio
--------------+----------------------
         Mean |    47.339   .1245762
    Std. Dev. |  29.45589   .0775135
          Min |         0          0
           Q1 |        24      .0632
       Median |        45      .1184
           Q3 |      68.5     .18025
          Max |       143      .3763

        eHARP |      #vio       %vio
--------------+----------------------
         Mean |        20          1
    Std. Dev. |         0          0
          Min |        20          1
           Q1 |        20          1
```

Median	20	1
Q3	20	1
Max	20	1

```
. powerps, price(P) quantity(X) axiom(eGARP eHARP) aei
                      Number of obs          =        20
                      Number of goods        =         5
                      Simulations            =      1000
                      Efficiency level       =         1
```

Axioms	Power	PS	Pass	AEI
eGARP	.995	-.005	0	.9055848
eHARP	1	0	0	.8449683

Summary statistics for simulations:

eGARP	#vio	%vio	AEI
Mean	47.339	.1245762	.842074
Std. Dev.	29.45589	.0775135	.0814885
Min	0	0	.5616641
Q1	24	.0632	.7924724
Median	45	.1184	.8516641
Q3	68.5	.18025	.9015746
Max	143	.3763	1

eHARP	#vio	%vio	AEI
Mean	20	1	.7268926
Std. Dev.	0	0	.0760639
Min	20	1	.4819741
Q1	20	1	.6767941
Median	20	1	.7307339
Q3	20	1	.7845821
Max	20	1	.8955998

4 Empirical illustrations

This section illustrates how to implement our commands using two types of data that are common in many revealed preference applications. The first type of data set contains the individual choices of experimental subjects. Such controlled environments are desirable from the perspective of empirical testing because relative prices can be calibrated across observations in order to engineer a sufficiently powerful test of, say, utility maximization. In our empirical illustration, we analyze the budgetary data collected in Andreoni and Miller (2002); other prominent examples of experiments involving budgetary designs include Choi et al. (2007, 2014), Andreoni and Sprenger (2012), and Halevy et al. (2018). The second type of data set contains annual household food consumption within broad categories. Aggregate household-level data have long been used to estimate parametric demands systems (see, e.g., Deaton and Muellbauer (1980), Banks et al. (1997), and Lewbel and Pendakur (2009)), and moreover, Poi (2002) makes use of the same data set in order to illustrate the estimation of parametric demand systems in *Stata*.

4.1 Experimental data

Andreoni and Miller (2002) tests whether the social choices of experimental subjects are rational, employing a dictator game in which one subject (the dictator) allocates token endowments between himself and another subject (the beneficiary) according to some rate of transfer. The payoffs of the dictator and the beneficiary are essentially two distinct goods and the transfer rates are the price ratios. The experiment contains two parts, where 142 subjects (Group 1) face $T = 8$ decision rounds, and where 34 subjects (Group 2) face $T = 11$ rounds. In this illustration, we focus on subjects in Group 1.

Andreoni and Miller (2002) finds that 13 subjects in Group 1 violate rationality, and for each of these 13 subjects reports the AEI (for GARP) and the number of violations of eGARP, eSARP, and eWARP at the efficiency level $e = 1$ (see Table II in Andreoni and Miller (2002)). Banerjee and Murphy (2006) complements this analysis and reports the number of violations of eWGARP at the efficiency level $e = 1$ (see Table 1 in Banerjee and Murphy (2006)). Using the commands `checkax` and `aei`, the following code replicates these results:

```
. local axioms eGARP eWGARP eSARP eWARP
.
. forvalues subject = 1/142 {
.         foreach axiom of local axioms {
.                 checkax, price(P) quantity(Q`subject´) axiom(`axiom´)
.          }
.         aei, price(P) quantity(Q`subject´) axiom(eGARP)
}
```
(*output omitted*)

The results from the preceding code are reported in Table 1. In Figure 1, we plot the fraction of the 142 subjects satisfying eGARP, eSGARP, eHARP, and eCM for values of e between 0.85 and 1 in an equally spaced grid with increments of 0.01. The results used to generate Figure 1 are obtained by looping over all subjects, axioms, and efficiency levels in the grid, and evaluating the command `checkax` for each subject, axiom, and efficiency level. The following line of code illustrates one such evaluation:

```
checkax, price(P) quantity(Q`subject´) efficiency(0.85)
```
(*output omitted*)

Since subjects are choosing from among bundles of two goods, eGARP (eSARP) and eWGARP (eWARP) are equivalent, and must by construction deliver identical results in terms of pass rates (but not in terms of the number and fraction of violations). Furthermore, while theoretically possible, the empirical differences between eGARP (eWGARP) and eSARP (eWARP) are negligible, implying that distinctions between demand correspondences and demand functions are not of first order importance within

Table 1: Replication of results in Andreoni and Miller (2002, Table II) and Banerjee and Murphy (2006, Table 1)[†]

	Number (fraction) of violations				
Subject	*e*GARP	*e*WGARP	*e*SARP	*e*WARP	AEI (GARP)
3	2 (0.036)	1 (0.036)	4 (0.071)	1 (0.036)	1.000*
38	8 (0.143)	2 (0.071)	9 (0.161)	2 (0.071)	0.917
40	8 (0.143)	3 (0.107)	11 (0.196)	3 (0.107)	0.833
41	1 (0.018)	1 (0.036)	2 (0.036)	1 (0.036)	1.000*
47	1 (0.018)	1 (0.036)	2 (0.036)	1 (0.036)	1.000*
61	4 (0.071)	1 (0.036)	5 (0.089)	1 (0.036)	0.917
72	1 (0.018)	1 (0.036)	2 (0.036)	1 (0.036)	1.000*
87	1 (0.018)	1 (0.036)	2 (0.036)	1 (0.036)	1.000*
90	2 (0.036)	1 (0.036)	2 (0.036)	1 (0.036)	0.975
104	1 (0.018)	1 (0.036)	3 (0.054)	1 (0.036)	1.000*
126	1 (0.018)	1 (0.036)	4 (0.071)	1 (0.036)	1.000*
137	1 (0.018)	1 (0.036)	2 (0.036)	1 (0.036)	1.000*
139	1 (0.018)	1 (0.036)	2 (0.036)	1 (0.036)	1.000*

[†] The number (and fraction) of violations are reported at the efficiency level $e = 1$. The symbol (*) indicates that an ε-change in choices eliminates all GARP violations.

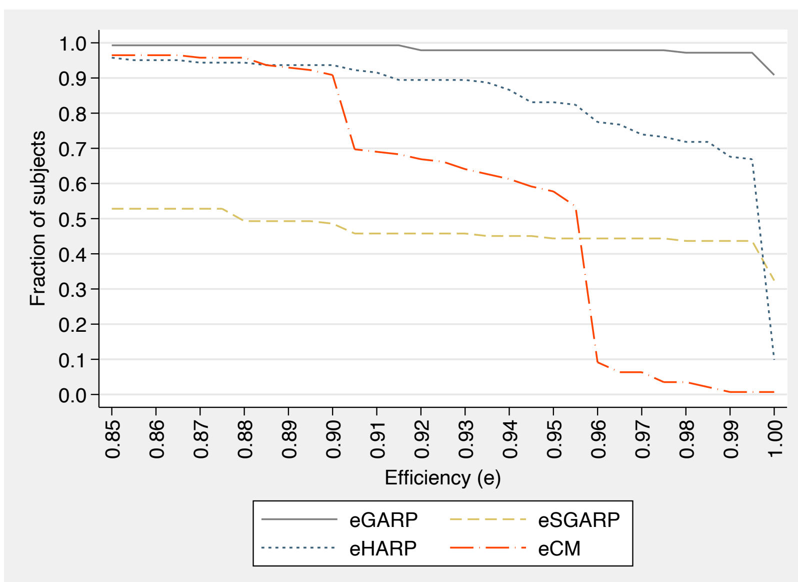

Figure 1: AEI distributions for *e*GARP, *e*SGARP, *e*HARP, and *e*CM

Table 2: Summary statistics for eSGARP, eHARP, and eCM[†]

	Number (fraction) of violations			AEI		
Statistic	eSGARP	eHARP	eCM	SGARP	HARP	CM
Mean	16.47 (0.257)	6.29 (0.789)	7.68 (0.960)	0.745	0.976	0.935
Std. dev.	16.80 (0.262)	2.90 (0.362)	1.03 (0.129)	0.288	0.049	0.035
Min	0 (0.000)	0 (0.000)	0 (0.000)	0.333	0.707	0.800
Q1	0 (0.000)	5 (0.625)	8 (1.000)	0.333	0.966	0.905
Median	8 (0.125)	8 (1.000)	8 (1.000)	0.875	1.000	0.957
Q3	37 (0.578)	8 (1.000)	8 (1.000)	1.000	1.000	0.957
Max	41 (0.641)	8 (1.000)	8 (1.000)	1.000	1.000	1.000

[†] The number (and fraction) of violations are reported at the efficiency level $e = 1$.

Table 3: Power summary statistics for eSGARP, eHARP, and eCM[†]

	Number (fraction) of violations			AEI		
Statistic	eSGARP	eHARP	eCM	SGARP	HARP	CM
Mean	17.53 (0.274)	7.96 (0.995)	7.93 (0.992)	0.693	0.763	0.761
Std. dev.	12.29 (0.192)	0.47 (0.0583)	0.65 (0.081)	0.181	0.120	0.124
Min	0 (0.000)	0 (0.000)	0 (0.000)	0.335	0.358	0.358
Q1	8 (0.125)	8 (1.000)	8 (1.000)	0.551	0.684	0.675
Median	15 (0.234)	8 (1.000)	8 (1.000)	0.667	0.773	0.769
Q3	27 (0.422)	8 (1.000)	8 (1.000)	0.840	0.856	0.859
Max	53 (0.828)	8 (1.000)	8 (1.000)	1.000	1.000	1.000

[†] The number (and fraction) of violations are reported at the efficiency level $e = 1$.

these data. Since neither Andreoni and Miller (2002) nor Banerjee and Murphy (2006) reports any results for eSGARP, eHARP, or eCM, we give these axioms more attention: we calculate the mean, standard deviation, minimum, first quartile (Q1), median, third quartile (Q3), and maximum of the number (and fraction) of violations and of the AEIs corresponding to SGARP, HARP, and CM. The results are displayed in Table 2.

Finally, we turn to power and predictive success. By looping over all subjects, axioms, and values of e between 0.4 and 1.0, we calculate the power and predictive success for every subject, axiom, and efficiency level in the grid. The following line of code illustrates one such evaluation:

```
powerps, price(P) quantity(Q`subject´) efficiency(0.4)
  (output omitted)
```

We summarize the results in three different ways. First, Figure 2 plots the power of eGARP, eSGARP, eHARP, and eCM for every efficiency level in the grid. Note that since all subjects face the same budgets, the power of each test is identical across sub-

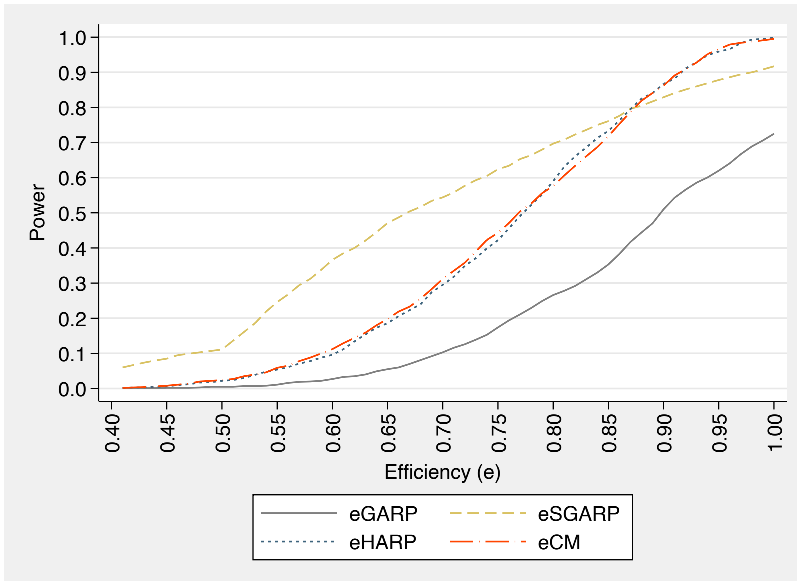

Figure 2: Power of eGARP, eSGARP, eHARP, and eCM

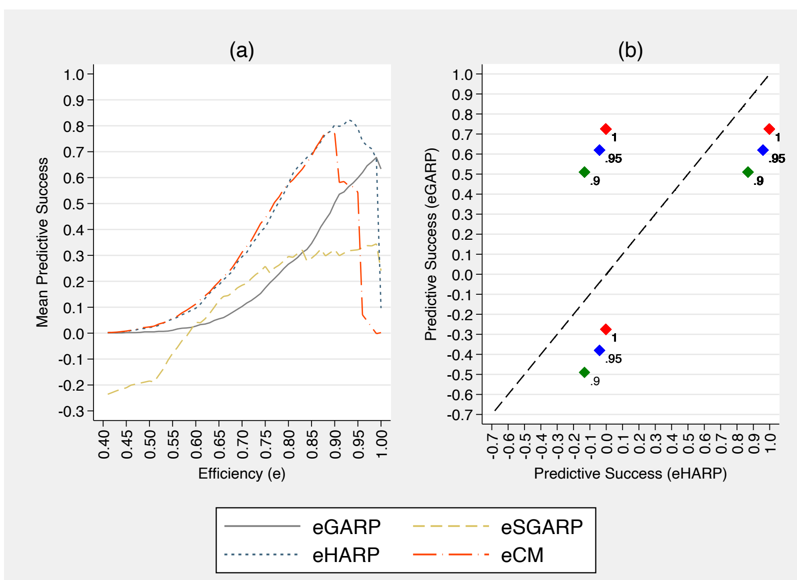

Figure 3: (a) Mean predictive success for eGARP, eSGARP, eHARP, and eCM; (b) Scatterplot of eHARP versus eGARP. In panel (b), the dashed line is the 45° line, and the marker numbers refer to efficiency levels.

jects. Second, Table 3 gives the mean, standard deviation, minimum, first quartile (Q1), median, third quartile (Q3), and maximum of the number (and fraction) of violations and of the AEIs corresponding to SGARP, HARP, and CM, over all repetitions in the simulated uniformly random data. Third, Figure 3(a) plots the mean predictive success across all subjects at each efficiency level in the grid, and Figure 3(b) is a subject-level scatterplot of *e*HARP versus *e*GARP at selected efficiency levels.

4.2 Aggregate household consumption data

In the second empirical illustration, we use aggregate household consumption data from the 1987-1988 Nationwide Food Consumption Survey conducted by the United States Department of Agriculture. This data set is used by Poi (2002) in order to illustrate how Stata's `ml` command can be used to fit the quadratic almost ideal demand system (QUAIDS). This data set is named `food.dta` in the repository "data sets for Stata Base Reference Manual, Release 16" (https://www.stata-press.com/data/r16/r.html), and contains budget shares and prices for the following four aggregated food categories: meats, fruits and vegetables, breads and cereals, and miscellaneous. As in Poi (2002), we use a sample of 4,048 households.

To test whether the data can be rationalized by preferences that are common across all households, we compute the AEI for GARP and WGARP:

```
. use http://www.stata-press.com/data/r16/food.dta, clear
. mkmat p1 p2 p3 p4, matrix(P)
. forvalues i = 1(1)4 {
.         gen x`i´ = w`i´* expfd/p`i´
. }
. mkmat x1 x2 x3 x4, matrix(X)
. aei, price(P) quantity(X)
   Number of obs          =       4048
   Number of goods        =          4
   Tolerance level        =    1.0e-06

      Axiom          AEI
      eGARP      .459821

. aei, price(P) quantity(X) axiom(eWGARP)
   Number of obs          =       4048
   Number of goods        =          4
   Tolerance level        =    1.0e-06

      Axiom          AEI
     eWGARP      .459821
```

We find that testing for *e*GARP takes considerably longer than testing for *e*WGARP, which suggests that the main computational burden in testing for *e*GARP is associated with the calculation of the transitive closure of the revealed preference relation. Interestingly, we find identical values of the AEI for GARP and WGARP, indicating that none of the violations of GARP can be attributed to violations of transitivity.

Finally, because eWGARP is considerably faster to test than eGARP, we calculate the power of eWGARP at an efficiency level equal to the AEI for WGARP:

```
. aei, price(P) quantity(X) axiom(eWGARP)
    Number of obs           =       4048
    Number of goods         =          4
    Tolerance level         =    1.0e-06

         Axiom |        AEI
---------------+-----------
        eWGARP |    .459821

. return list
scalars:
                r(TOL) =  1.00000000000e-06
              r(GOODS) =  4
                r(OBS) =  4048
         r(AEI_eWGARP) =  .4598209857940674
macros:
              r(AXIOM) : " eWGARP"

. powerps, price(P) quantity(X) axiom(eWGARP) efficiency(`r(AEI_eWGARP)´)
                       Number of obs           =       4048
                       Number of goods         =          4
                       Simulations             =       1000
                       Efficiency level        =        .46

        Axioms |      Power         PS       Pass        AEI
---------------+--------------------------------------------
        eWGARP |       .423       .423          1   .4598211

Summary statistics for simulations:

        eWGARP |       #vio       %vio
---------------+----------------------
          Mean |       .832          0
     Std. Dev. |   1.790246          0
           Min |          0          0
            Q1 |          0          0
        Median |          0          0
            Q3 |          1          0
           Max |         39          0
```

5 Conclusions

In this paper, we have presented new *Stata* commands to test whether observed data on prices and quantities can be rationalized by different notions of utility maximization.

The commands are implementations of nonparametric revealed preference restrictions that can be formulated as combinatorial algorithms. An important property of such algorithms is that they converge in a finite number of steps, and consequently, can be implemented on rather large data sets. Although the commands are implementations of characterizations that are intrinsically deterministic (in the sense that they lack stochastic components), they also allow the user to calculate diagnostic measures such as goodness-of-fit, power, and predictive success of the underlying behavioral model.

The package `rpaxioms` contains implementations of perhaps the most basic concepts in the empirical revealed preference literature. Two natural extensions left for future work come to mind. The first is to provide implementations of revealed preference characterizations of other behavioral models, including special cases on preferences which amount to different forms of separability (e.g., expected utility under risk and exponential discounted utility over time). The second is to provide implementations of more disaggregated measures of goodness-of-fit and power. Although some of these models and measures can be implemented by solving (mixed-integer) linear programming problems, this is not a trivial task and the computational complexity of doing so crucially depends on the algorithms used to solve such problems.

6 Acknowledgments

We thank Glenn Nielsen and John Quah for helpful comments and useful suggestions. Hjertstrand thanks Torsten Söderbergs stiftelse for financial support.

7 References

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

———. 1972. Efficiency estimation of production functions. *International Economic Review* 13: 568–598.

———. 1973. On a system of inequalities in demand analysis: An extension of the classical method. *International Economic Review* 14: 460–472.

Aguiar, V. H., P. Hjertstrand, and R. Serrano. 2020. A Rationalization of the Weak Axiom of Revealed Preference. Bravo Working Paper 2020-016, Brown University and IFN Working Paper No. 1321.

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

Andreoni, J., and C. Sprenger. 2012. Estimating time preferences from convex budgets. *American Economic Review* 102: 3333–3356.

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

Banks, J., R. Blundell, and A. Lewbel. 1997. Quadratic Engel curves and consumer demand. *Review of Economics and Statistics* 79: 527–539.

Beatty, T. K. M., and I. A. Crawford. 2011. How demanding is the revealed preference approach to demand? *American Economic Review* 101: 2782–2795.

Becker, G. S. 1962. Irrational behavior and economic theory. *Journal of Political Economy* 70: 1–13.

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

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

Chiappori, P.-A., and J.-C. Rochet. 1987. Revealed preferences and differentiable demand. *Econometrica* 55: 687–691.

Choi, S., R. Fisman, D. Gale, and S. Kariv. 2007. Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review* 97: 1921–1938.

Choi, S., S. Kariv, W. Muller, and D. Silverman. 2014. Who is (more) rational? *American Economic Review* 104: 1518–1550.

Deaton, A., and J. Muellbauer. 1980. An almost ideal demand system. *American Economic Review* 70: 312–326.

Diewert, W. E. 1973. Afriat and revealed preference theory. *Review of Economic Studies* 40: 419–425.

Famulari, M. 1995. A household-based, nonparametric test of demand theory. *Review of Economics and Statistics* 77: 372–383.

Halevy, Y., D. Persitz, and L. Zrill. 2018. Parametric recoverability of preferences. *Journal of Political Economy* 126: 1558–1593.

Heufer, J., and P. Hjertstrand. 2019. Homothetic efficiency: theory and applications. *Journal of Business & Economic Statistics* 37: 235–247.

Houtman, M., and J. A. H. Maks. 1985. Determining All Maximal Data Subsets Consistent with Revealed Preference. *Kwantitatieve Methoden* 6(19): 89–104.

Lewbel, A., and K. Pendakur. 2009. Tricks with Hicks: The EASI demand system. *American Economic Review* 99: 827–863.

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

McFadden, D. 1973. Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics*, ed. P. Zarembka, 105–142. New York: Academic Press.

Nishimura, H., E. A. Ok, and J. K.-H. Quah. 2017. A comprehensive approach to revealed preference theory. *American Economic Review* 107: 1239–1263.

Poi, B. P. 2002. From the help desk: Demand system estimation. *Stata Journal* 2: 403–410.

Polisson, M., and J. K.-H. Quah. 2022. Rationalizability, Cost-Rationalizability, and Afriat's Efficiency Index. IFS Working Papers 22/04.

Polisson, M., J. K.-H. Quah, and L. Renou. 2020. Revealed preferences over risk and uncertainty. *American Economic Review* 110: 1782–1820.

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

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

Swofford, J. L., and G. A. Whitney. 1987. Nonparametric tests of utility maximization and weak separability for consumption, leisure and money. *Review of Economics and Statistics* 69: 458–464.

Varian, H. R. 1982. The nonparametric approach to demand analysis. *Econometrica* 50: 945–974.

———. 1983. Non-parametric tests of consumer behavior. *The Review of Economic Studies* 50: 99–110.

———. 1990. Goodness-of-fit in optimizing models. *Journal of Econometrics* 46: 125–140.

Warshall, S. 1962. A theorem on Boolean matrices. *Journal of the American Association of Computing Machinery* 9: 11–12.

About the authors

Marcos Demetry is a PhD Student at Linnaeus University in Växjö, Sweden, and an affiliated doctoral student at the Research Institute of Industrial Economics (IFN) in Stockholm, Sweden.

Per Hjertstrand is an Associate Professor and Research Fellow at the Research Institute of Industrial Economics (IFN) in Stockholm, Sweden.

Matthew Polisson is a Senior Lecturer at the University of Bristol, United Kingdom, and a Research Fellow at the Institute for Fiscal Studies (IFS) in London, United Kingdom.