IFN Working Paper No. 1342, 2020

Testing Axioms of Revealed Preference in Stata

Marcos Demetry, Per Hjertstrand and Matthew Polisson

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
Research Institute of Industrial Economics
Stockholm, Sweden
marcos.demetry@ifn.se

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

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

Abstract. This paper introduces the *Stata* commands `checkax`, `aei`, and `powerps` as a bundle within the package `rpaxioms`. The first command allows the user to test whether consumer demand data satisfy a number of revealed preference axioms at a given efficiency level; 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 at any given efficiency level. The commands are illustrated using individual-level experimental data and aggregated household-level consumption data.

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

1 Introduction

The standard way to check whether a finite set of consumer demand observations is compatible with economic rationality, i.e., the hypothesis of utility maximization, is to apply a revealed preference test. Such a procedure checks whether the data, which consist of observed prices and quantities for a set of consumer goods, satisfy a given revealed preference axiom, e.g., the generalized axiom of revealed preference (GARP). Varian's (1982) formulation of Afriat's (1967) well-known theorem states that consumer demand data obey GARP *if and only if* there exists a continuous, strictly increasing, and concave utility function which rationalizes the data.

In general, revealed preference tests are 'sharp', in the sense that they deliver a binary response as to whether the observed demand data are compatible with the underlying behavioral model. However, given sufficiently rich data, an outright failure of even fairly permissive notions of rationalizability should not be unexpected, and it may well be that these same data are in fact very close to rationalizability. As noted by Varian (1990), 'nearly optimizing' behavior is often 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) proposes a simple combinatorial algorithm to test whether consumer demand 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). All axioms and their behavioral implications are described in more detail in Section 2.4.

Afriat (1973) proposes that an upper limit on the efficiency level at which the data satisfy GARP, or the *critical* cost efficiency, is 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) later extends and interprets 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 2.2.

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 from 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 and 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 2.3.

We illustrate these three commands—`checkax`, `aei`, and `powerps`—on two types of data sets 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 subjects are compatible with basic utility maximization and several different variants of this model. Second, using aggregated 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]

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

2 Revealed preference

Suppose that there are T observations of the prices and quantities of K goods. At observation $t = 1, \ldots, T$, the prices and quantities are denoted by $\mathbf{p}^t = (p_1^t, \ldots, p_K^t)$ and $\mathbf{x}^t = (x_1^t, \ldots, x_K^t)$, respectively. We assume that all prices are strictly positive, and that all quantities are non-negative (i.e., some but not all quantities at any given observation may be equal to zero).

2.1 The Generalized Axiom of Revealed Preference at efficiency e

Consider any number $0 < e \leq 1$. 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 $e\mathbf{p}^t \cdot \mathbf{x}^t \geq \mathbf{p}^t \cdot \mathbf{x}^s$. This means that $\mathbf{x}^t$ is chosen even though the cost of the bundle $\mathbf{x}^s$ (at 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 the direct revealed preferred relation R_e^D. The number e can be interpreted as a level of cost efficiency. When $e = 1$, which we refer to as 'full' efficiency, these relations reduce to the usual revealed preference relations (Varian 1982).

A data set $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ satisfies the *generalized axiom of revealed preference* at efficiency level e, abbreviated eGARP, if $\mathbf{x}^t R_e \mathbf{x}^s$ implies $e\mathbf{p}^s \cdot \mathbf{x}^s \leq \mathbf{p}^s \cdot \mathbf{x}^t$.

Varian's (1982) standard version of the generalized axiom of revealed preference (GARP) can be obtained by setting $e = 1$. It is well known that GARP is necessary and sufficient for a data set to be rationalized by a continuous, strictly increasing, and concave utility function (Afriat 1967; Diewert 1973; Varian 1982). The eGARP axiom 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 $e\mathbf{p}^t \cdot \mathbf{x}^t \geq \mathbf{p}^t \cdot \mathbf{x}^s$ 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} that is equal to 1 if $\mathbf{x}^t R_e \mathbf{x}^s$, and 0 otherwise. Varian (1982) suggests calculating RT using Warshall's algorithm (Warshall 1962). The command `checkax` constructs RT using a vectorized version of Warshall's algorithm. Third, eGARP is violated if $RT_{ts} = 1$ 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, in a data set of 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 $T\,(T-1)$. At any user-specified efficiency level e, the command `checkax` reports whether or not the data satisfy eGARP, as well as the number and fraction of violations.[2]

2. 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.

2.2 The Afriat Efficiency Index

The Afriat (1973) efficiency index (AEI) is defined as the maximal value of e (the supremum, to be precise) such that the data obey eGARP. 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. The command `aei` calculates the AEI by implementing the binary search algorithm described in Varian (1990).

2.3 Power and predictive success

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., bundles along the frontiers of the budget sets) are equally likely to be chosen. Bronars (1987) suggests implementing the index using Monte Carlo methods, which are executed in the command `powerps` across three steps. 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, say, eGARP at a given efficiency level e. The power measure is the fraction of these synthetic data sets which would then violate eGARP.

The command `powerps` allows the user to simultaneously calculate the power corresponding to several axioms at once, in order to simplify power comparisons across axioms. The command also allows the user to choose the efficiency level and to specify the number of repetitions involved in the third step. Moreover, it also allows the user to set the random seed in the generation of the Dirichlet random variables in the first step, in order to make any power calculations 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 some axiom at a given efficiency level, and 0 otherwise, and where the power index corresponding to that axiom is calculated at the same efficiency level. This measure of predictive success can then be straightforwardly aggregated across individual data sets.

2.4 Other axioms

The empirical content of utility maximization is entirely captured by eGARP. Varian (1982) gives a revealed preference characterization of the utility maximization model under full efficiency, i.e., when $e = 1$. Halevy et al. (2018) extends these results and provides a characterization under partial efficiency, i.e., when $e < 1$.

Our commands are also implementable for other revealed preference axioms that characterize a number of the common variants of basic 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 at any user-specified efficiency level.

- (Strong axiom of revealed preference) A data set $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ 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 (full efficiency) is necessary and sufficient for a data set to be rationalized 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, which means that GARP can accommodate demand correspondences while SARP requires demand functions. Like eGARP, there can be up to $T(T-1)$ violations of eSARP.

- (Weak generalized axiom of revealed preference) A data set $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ 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 $e\mathbf{p}^s \cdot \mathbf{x}^s \leq \mathbf{p}^s \cdot \mathbf{x}^t$. Aguiar et al. (2020) shows that WGARP (full efficiency) is necessary and sufficient for a data set to be rationalized by a continuous, strictly increasing, piecewise concave, and skew-symmetric preference function (see Aguiar et al. (2020) for the definitions of a preference function and the relevant properties pertaining to preference functions). Banerjee and Murphy (2006) shows that WGARP and GARP are equivalent when the consumer chooses from among bundles of two goods, i.e., when $K = 2$. The total possible number of violations of eWGARP is $T(T-1)/2$.

- (Weak axiom of revealed preference) A data set $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ 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 (full efficiency) is necessary and sufficient for a data set to be rationalized by a continuous, strictly increasing, piecewise *strictly* concave, and skew-symmetric preference function. The difference between WGARP and WARP is analogous to the difference between GARP and SARP. 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 $e\mathbf{p}^t \cdot \mathbf{x}^t \geq \mathbf{p}^t \cdot \mathbf{y}^s$, where $\mathbf{y}^s$ is *any permuta-*

tion of $\mathbf{x}^s$,[3] and where the transitive closure R_e of R_e^D follows accordingly. A data set $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ satisfies the *symmetric generalized axiom of revealed preference* at efficiency level e, abbreviated eSGARP, if $\mathbf{x}^t R_e \mathbf{x}^s$ implies $e\mathbf{p}^s \cdot \mathbf{x}^s \leq \mathbf{p}^s \cdot \mathbf{y}^t$ (where once again $\mathbf{y}^t$ is any permutation of $\mathbf{x}^t$). Nishimura et al. (2017) shows that eSGARP is necessary and sufficient for a data set to be rationalized by a continuous, strictly increasing, concave, and *symmetric* utility function. Polisson et al. (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 $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ satisfies the *homothetic axiom of revealed preference* at efficiency level e, abbreviated eHARP, if for all distinct $(s, t, u, \ldots, v)$ we have $(\mathbf{p}^t \cdot \mathbf{x}^s)(\mathbf{p}^s \cdot \mathbf{x}^u) \cdots (\mathbf{p}^v \cdot \mathbf{x}^t) \geq (e\mathbf{p}^t \cdot \mathbf{x}^t)(e\mathbf{p}^s \cdot \mathbf{x}^s) \cdots (e\mathbf{p}^v \cdot \mathbf{x}^v)$. Varian (1983) shows that HARP (full efficiency) is necessary and sufficient for a data set to be rationalized by a continuous, strictly increasing, concave, and *homothetic* utility function. Heufer and Hjertstrand (2019) provide a characterization under partial efficiency, 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 $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ satisfies a *cyclical monotonicity* condition at efficiency level e, abbreviated eCM, if for all distinct $(s, t, u, \ldots, v)$ it must be the case that $\mathbf{p}^t \cdot (\mathbf{x}^s - e\mathbf{x}^t) + \mathbf{p}^s \cdot (\mathbf{x}^u - e\mathbf{x}^s) + \cdots + \mathbf{p}^v \cdot (\mathbf{x}^t - e\mathbf{x}^v) \geq 0$. Brown and Calsamiglia (2007) shows that CM (full efficiency) is necessary and sufficient for a data set to be rationalized 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 relaxations of rationalizability has 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 a data set to be rationalized 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 $1 - \epsilon$, for some $\epsilon > 0$ arbitrarily small, then we could always find a smooth rationalization. Since the CCEI is defined as a supremum, the

3. 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 $(3, 2, 1)$.

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 available on SSC and can be installed by entering 'SSC install rpaxioms' at the Stata command prompt. 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 Syntax of checkax

The syntax of `checkax` is as follows:

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

The optional arguments are:

`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 positive, and no greater than one.

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`.

The following examples illustrate `checkax`. The price and quantity matrices are `P` and `X`, respectively, where both data matrices are in Excel (`.xls`) format. 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 $e = 0.95$. The output also contains the return list for the second example.

```
. import excel using "${datadir}/prices.xls", clear
. mkmat A B C D E, matrix(P)
. import excel using "${datadir}/quantities.xls", clear
. mkmat A B C D E, 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       42.37
----------------------------------------------------

. 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       27.37
          eHARP |        0         20         100
----------------------------------------------------

. return list
scalars:
        r(FRAC_VIO_eHARP) =  100
         r(NUM_VIO_eHARP) =  20
            r(PASS_eHARP) =  0
                   r(EFF) =  .95
                 r(GOODS) =  5
                   r(OBS) =  20
        r(FRAC_VIO_eGARP) =  27.37
         r(NUM_VIO_eGARP) =  104
            r(PASS_eGARP) =  0

macros:
                 r(AXIOM) : " eGARP eHARP"
```

3.2 Syntax of aei

The syntax of `aei` is as follows:

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

The optional arguments are:

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

`tolerance(`#`)` sets the tolerance level of the termination criterion 10^{-n} by specifying the integer n. For example, `tolerance(6)` sets the tolerance level to 10^{-6}. The default option is `tolerance(12)`, which gives a default tolerance level 10^{-12}. The integer n in the termination criterion 10^{-n} cannot be less than 1 or greater than 18.

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

`Axiom`, `Goods`, and `Obs` are the same as in `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`.

The following examples illustrate `aei` using the same data as above. The first example runs `aei` using its default options, i.e., for eGARP with a tolerance level set to 10^{-12}. The second example runs `aei` for eGARP and eHARP with the tolerance level set to 10^{-6}. The output also contains the return list for the second example.

```
. aei, price(P) quantity(X)
    Number of obs          =         20
    Number of goods        =          5
    Tolerance level        =    1.0e-12

          Axiom |      AEI
  --------------+-----------
          eGARP |   .9055851

. quietly aei, price(P) quantity(X) axiom(eGARP eHARP) tolerance(6)
. return list
scalars:
                r(TOL) =  1.00000000000e-06
              r(GOODS) =  5
                r(OBS) =  20
           r(AEI_eHARP) =  .844968318939209
           r(AEI_eGARP) =  .9055848121643066
macros:
              r(AXIOM) : " eGARP eHARP"
```

3.3 Syntax of powerps

The syntax of `powerps` is as follows:

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

The optional arguments are:

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

`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.

Running `powerps` 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`.

`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`.

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

`Obs` is the number of goods. Given as the scalar `r(OBS)` 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).

The following examples illustrate 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 tests the same axioms but also includes the option aei, which calculates the AEI for both *e*GARP and *e*HARP for each of the 1,000 simulated data sets. The output also contains the return list for the second example.

```
. 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	.9055851
eHARP	1	0	0	.8449687

Summary statistics for simulations:

eGARP	#vio	%vio
Mean	47.339	12.45762
Std. Dev.	29.45589	7.751351
Min	0	0
Q1	24	6.32
Median	45	11.84
Q3	68.5	18.025
Max	143	37.63

eHARP	#vio	%vio
Mean	20	100
Std. Dev.	0	0
Min	20	100

Q1	20	100
Median	20	100
Q3	20	100
Max	20	100

```
. 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	.9055851
eHARP	1	0	0	.8449687

Summary statistics for simulations:

eGARP	#vio	%vio	AEI
Mean	47.339	12.45762	.842074
Std. Dev.	29.45589	7.751351	.0814885
Min	0	0	.5616643
Q1	24	6.32	.7924721
Median	45	11.84	.8516639
Q3	68.5	18.025	.9015748
Max	143	37.63	1

eHARP	#vio	%vio	AEI
Mean	20	100	.7268926
Std. Dev.	0	0	.0760639
Min	20	100	.4819745
Q1	20	100	.6767944
Median	20	100	.7307337
Q3	20	100	.7845822
Max	20	100	.8955996

```
. return list

scalars:
        r(POWER_eHARP) =  1
           r(PS_eHARP) =  0
         r(PASS_eHARP) =  0
          r(AEI_eHARP) =  .844968688899371
                r(SIM) =  1000
          r(TOL_eharp) =  12
        r(POWER_eGARP) =  .995
           r(PS_eGARP) =  -.005
         r(PASS_eGARP) =  0
          r(AEI_eGARP) =  .9055851063826594
          r(TOL_eGARP) =  12
                r(EFF) =  1
              r(GOODS) =  5
                r(OBS) =  20
```

```
macros:
              r(AXIOM) : " eGARP eHARP"
matrices:
       r(SUMSTATS_eHARP) :  7 x 3
     r(SIMRESULTS_eHARP) :  1000 x 3
       r(SUMSTATS_eGARP) :  7 x 3
     r(SIMRESULTS_eGARP) :  1000 x 3
```

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. Aggregated 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*)

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

	Number of violations (fraction)				
Subject	eGARP	eWGARP	eSARP	eWARP	AEI (GARP)
3	2 (3.57)	1 (3.57)	4 (7.14)	1 (3.57)	1.000*
38	8 (14.29)	2 (7.14)	9 (16.07)	2 (7.14)	0.917
40	8 (14.29)	3 (10.71)	11 (19.64)	3 (10.71)	0.833
41	1 (1.79)	1 (3.57)	2 (3.57)	1 (3.57)	1.000*
47	1 (1.79)	1 (3.57)	2 (3.57)	1 (3.57)	1.000*
61	3 (5.36)	1 (3.57)	5 (8.93)	1 (3.57)	0.917
72	1 (1.79)	1 (3.57)	2 (3.57)	1 (3.57)	1.000*
87	1 (1.79)	1 (3.57)	2 (3.57)	1 (3.57)	1.000*
90	2 (3.57)	1 (3.57)	2 (3.57)	1 (3.57)	0.975
104	1 (1.79)	1 (3.57)	3 (5.36)	1 (3.57)	1.000*
126	1 (1.79)	1 (3.57)	4 (7.14)	1 (3.57)	1.000*
137	1 (1.79)	1 (3.57)	2 (3.57)	1 (3.57)	1.000*
139	1 (1.79)	1 (3.57)	2 (3.57)	1 (3.57)	1.000*

Notes: The number (and fraction) of violations are reported at the efficiency level $e = 1$. *Indicates that an ε-change in choices eliminates all GARP violations.

The results from the preceding code are reported in Table 1. Next, 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 an increment 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(0.7)
```
(*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 empirical results (except for 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 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),

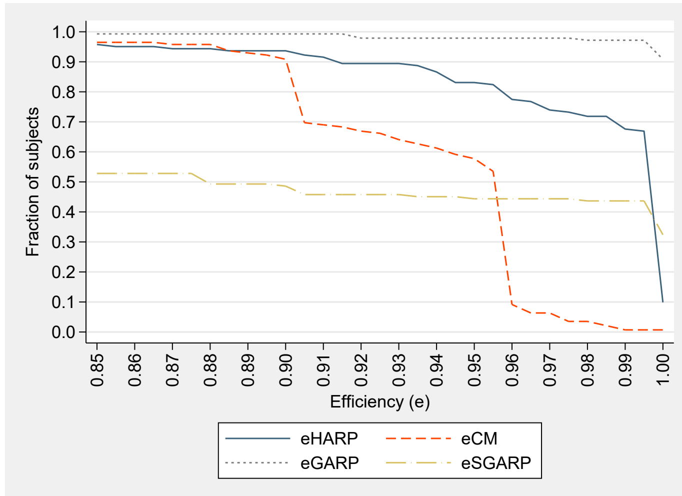

Figure 1: AEI distributions for eGARP, eSGARP, eHARP, and eCM

Table 2: Summary statistics for eSGARP, eHARP, and eCM

	Number of violations (fraction)			AEI		
Statistic	eSGARP	eHARP	eCM	SGARP	HARP	CM
Mean	16.47 (25.74)	6.29 (78.61)	7.68 (96.04)	0.745	0.976	0.935
Std. dev.	16.80 (26.24)	2.90 (36.21)	1.03 (12.92)	0.288	0.049	0.035
Min	0 (0)	0 (0)	0 (0)	0.333	0.707	0.800
Q1	0 (0)	5 (62.50)	8 (100)	0.333	0.966	0.905
Median	8 (12.50)	8 (100)	8 (100)	0.875	1	0.957
Q3	37 (57.81)	8 (100)	8 (100)	1	1	0.957
Max	41 (64.06)	8 (100)	8 (100)	1	1	1

Notes: The number (and fraction) of violations are reported at the efficiency level $e = 1$.

and maximum of the number (and fraction) of violations and 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

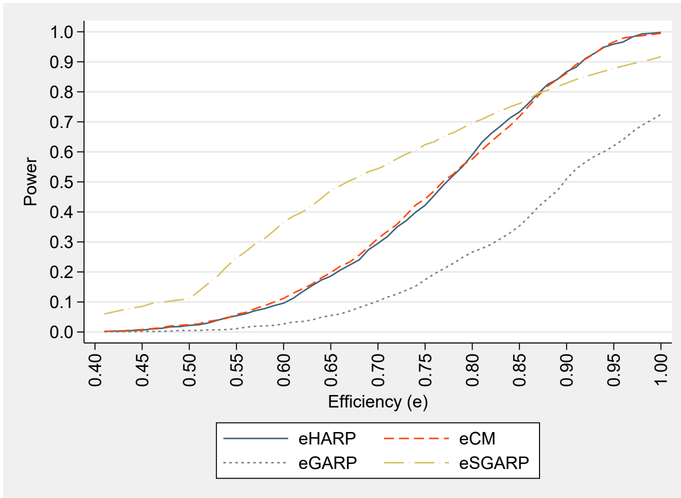

Figure 2: Power of eGARP, eSGARP, eHARP, and eCM

that since all subjects face the same budgets, the power of each test is identical across subjects. 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 for 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 eHARP versus eGARP at selected efficiency levels.

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

	Number of violations (fraction)			AEI		
Statistic	eSGARP	eHARP	eCM	SGARP	HARP	CM
Mean	17.53 (27.39)	7.96 (99.49)	7.93 (99.15)	0.693	0.763	0.761
Std. dev.	12.29 (19.21)	0.47 (5.83)	0.65 (8.10)	0.181	0.120	0.124
Min	0 (0)	0 (0)	0 (0)	0.335	0.358	0.358
Q1	8 (12.50)	8 (100)	8 (100)	0.551	0.684	0.675
Median	15 (23.44)	8 (100)	8 (100)	0.667	0.773	0.769
Q3	27 (42.19)	8 (100)	8 (100)	0.840	0.856	0.859
Max	53 (82.81)	8 (100)	8 (100)	1	1	1

Notes: The number (and fraction) of violations are reported at the efficiency level $e = 1$.

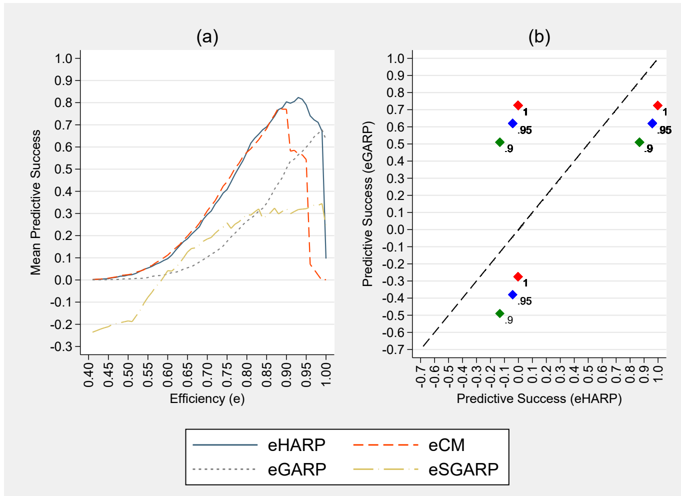

Figure 3: (a) Mean predictive success for eGARP, eSGARP, eHARP, and eCM; (b) Scatterplot of eHARP versus eGARP. In panel (b), the dashed line is the 45° line, and the marker numbers refer to efficiency levels.

4.2 Aggregated household consumption data

In the second empirical illustration, we use aggregated 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 'Datasets 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) tolerance(6)
    Number of obs           =       4048
    Number of goods         =          4
    Tolerance level         =    1.0e-06

─────────────────────────
         Axiom │        AEI
```

```
        eGARP     .459821

. aei, price(P) quantity(X) axiom(eWGARP) tolerance(6)
    Number of obs          =       4048
    Number of goods        =          4
    Tolerance level        =    1.0e-06

        Axiom         AEI

       eWGARP     .459821
```

We have chosen a higher tolerance level for the termination criterion equal to 10^{-6} because of the large number of observations in the data set. At a given efficiency level, we find that testing for eGARP takes considerably longer than testing for eWGARP, which suggests that the main computational burden in testing for eGARP 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) tolerance(6)
    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 Acknowledgments

We thank Glenn Nielsen and John Quah for helpful comments and useful suggestions. Hjertstrand thanks Jan Wallander och Tom Hedelius stiftelse and Marianne och Marcus Wallenberg stiftelse for financial support.

6 References

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

———. 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.

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.

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., 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 Research Assistant 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 in Economics at the University of Bristol and a Research Fellow at the Institute for Fiscal Studies (IFS).