Ruhr Economic Papers

Jan Heufer

A Geometric Measure for the Violation of Utility Maximization

#69

Ruhr Graduate School in Economics
ECON
technische universität dortmund

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstraße 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstraße 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI Essen)
Hohenzollernstrasse 1/3, 45128 Essen, Germany

Ruhr Economic Papers #69

Responsible Editor: Wolfgang Leininger

ISSN 1864-4872 (online) – ISBN 978-3-86788-073-2

Ruhr Economic Papers

#69

Jan Heufer

A Geometric Measure for the Violation of Utility Maximization

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-073-2

Jan Heufer*

A Geometric Measure for the Violation of Utility Maximization

Abstract

Revealed Preference offers nonparametric tests for whether consumption observations can be rationalized by a utility function. If a consumer is inconsistent with GARP, we might need a measure for the severity of inconsistency. One widely used measure is the Afriat efficiency index (AEI). We propose a new measure based on the extent to which the indifference surfaces intersect the budget hyperplanes. The measure is intuitively appealing and, as a cut-off-rule evaluated by Monte Carlo experiments, performs very well compared to the AEI. The results suggest that the new measure is better suited to capture small deviations from utility maximation.

JEL Classification: C14, C60, D11, D12

Keywords: Consumer choice, efficiency index, GARP, nonparametric tests, revealed preference

September 2008

---

* Ruhr Graduate School in Economics and Technische Universität Dortmund. – This paper is drawn from doctoral research done at the Ruhr Graduate School in Economics at the University of Dortmund under the guidance of Wolfgang Leininger. I am grateful for his support and comments. Thanks to James Andreoni and John Miller for access to their data. The work was financially supported by the Paul Klemmer Scholarship of the RWI Essen, which is gratefully acknowledged. – All correspondence to Jan Heufer, Technische Universität Dortmund, Department of Economics and Social Science, Vogelpothsweg 87, 44227 Dortmund, Germany, e-mail: jan.heufer@uni-dortmund.de.

1 Introduction

Revealed Preference methods offer an elegant and unambiguous way of testing whether a set of observations on consumption could have been generated by a single utility maximizing consumer. The test was originally developed by Afriat (1967). Varian (1982) showed that his Generalized Axiom of Revealed Preference (GARP) is equivalent to Afriat's condition of cyclic consistency. Consistency with GARP can be tested very easily.

If a consumer's decisions are inconsistent with GARP we might want to have an idea of how severe this violation of utility maximization is. Alternatively, we would like to have a test for "almost optimizing" behavior. One such measure is the Afriat efficiency index (AEI, Afriat 1972), which is widely used.

We propose a new measure based on the extent to which the upper bound of the indifference surface of a decision intersects the budget on which the decision was made. The idea is to use preference relations that are implicit in a set of observations to construct the set of bundles which are revealed preferred to a consumption choice. The boundary of this set can be interpreted as an upper bound for the indifference surface. If the data violate GARP, some of these sets will intersect the budget hyperplane on which the choice was made. We then compute the area (or volume in higher dimensions) of the intersection of the revealed preferred set and the budget.

We also suggest a procedure to decide whether or not to treat a consumer who violates GARP as "close enough" to utility maximization. It is based on the reduction of the power the test has against random behavior. When testing this procedure with a set of utility maximizing decisions with added stochastic error, our new geometric measure performs very well compared to the AEI.

The remainder is organized as follows: Section 2 first briefly summarizes the revealed preference approach and the AEI. The new geometric measure is introduced and a procedure to decide whether a set of observations should be accepted as utility maximizing is suggested. Section 3 describes the procedure in more detail and compares the new measure and its performance with the AEI. In Section 4 the new measure is applied to experimental data from Andreoni and Miller (2002). Section 5 discusses the advantages and disadvantages of the new measure and concludes.

2 Theory

2.1 Preparations

The following notation is used: For all $x, y \in \mathbb{R}^\ell$ we write $x \geq y$ for $x_i \geq y_i$ for all i, $x > y$ for $x_i \geq y_i$ and $x \neq y$ for all i, and $x \gg y$ for $x_i > y_i$ for all i. We denote $\mathbb{R}^\ell_+ = \{x \in \mathbb{R}^\ell : x \geq 0\}$ and $\mathbb{R}^\ell_{++} = \{x \in \mathbb{R}^\ell : x \gg 0\}$.

A set of observed consumption choices consists of a set of chosen bundles of commodities and the prices and incomes at which these bundles were chosen. Let $X = \mathbb{R}^\ell_+$ be the commodity space, where $\ell \geq 2$ denotes the number of different commodities. The price space is $P = \mathbb{R}^\ell_{++}$, and the space of price-income vectors is $P \times \mathbb{R}_{++}$. Consumers choose bundles $x^i = (x^i_1, \ldots, x^i_\ell)' \in X$ when facing a price vector $p^i = (p^i_1, \ldots, p^i_\ell) \in P$ and an income $w^i \in \mathbb{R}_{++}$. A budget set is then defined by $B^i = B(p^i, w^i) = \{x \in X : p^i x^i \leq w^i\}$. The entire set of n observations on a consumer is denoted as $S = \{(x^i, B^i)\}^n_{i=1}$.

A utility function $u(x)$ *rationalizes* a set of observations S if $u(x^i) \geq u(x)$ for all x such that $p^i x^i \geq p^i x$ for all $i = 1, \ldots, n$.

The following definitions are needed to recover consumer preferences that are implicit in a set of consumption choices:

An observation x^i is
(1) *directly revealed preferred* to x, written $x^i R^0 x$, if $p^i x^i \geq p^i x$;
(2) *revealed preferred* to x, written $x^i R^* x$, if either $x^i R^0 x$ or for some sequence of bundles $(x^j, x^k, \ldots, x^m)$ such that $x^i R^0 x^j, x^j R^0 x^k, \ldots, x^m R^0 x$. In this case R^* is the *transitive closure* of the relation R^0.
(3) *strictly directly revealed preferred* to x, written $x^i P^0 x$, if $p^i x^i > p^i x$.

For consistency with the maximization of a piecewise linear utility function, Varian (1982) introduced the following condition: The set of observations S satisfies the *Generalized Axiom of Revealed Preference* (GARP) if $x^i R^* x^j$ does not imply $x^j P^0 x^i$.

It can then be shown (Afriat 1967, Varian 1982) that if the data satisfies GARP then there exists a concave, monotonic, continuous, non-satiated utility function that rationalizes the data.

The set of bundles that are revealed preferred to a certain bundle x^0 (which does not have to be an observed choice) is given by the convex monotonic hull of all choices revealed preferred to x^0. The interior of the convex monotonic hull is used to compute an approximate overcompensation function by Varian (1982). Knoblauch (1992) shows that the set of bundles revealed preferred to x^0, denoted $RP(x^0)$, is just the convex monotonic hull of all bundles in S that are revealed preferred to x^0:

$$RP\left(x^0\right) = H_{\text{convex}}\left(\{x \in X : x \geq x^i \text{ such that } x^i R^* x^0 \text{ for some } i = 1, \ldots, n\}\right), \quad (1)$$

where $H_{\text{convex}}(A)$ denotes the convex hull of a set of points A, i.e.

$$H_{\text{convex}}(A) = \left\{\sum_{i=1}^{n} \lambda_i a^i : a^i \in A, \lambda_i \in \mathbb{R}_+, \sum_{i=1}^{n} \lambda_i = 1\right\}$$

2.2 Prior Measures

Several goodness-of-fit measures have been proposed. Possibly the most popular measure for the severity of a violation is the Afriat efficiency index (AEI) due to Afriat (1972). Reporting the AEI has become a standard at least for experimental studies.[1] To obtain the AEI, budgets are shifted towards the origin until a set of observations is consistent with GARP. Let e be a number between 0 and 1. Define the relation $R^0(e)$ to be $x^i R^0(e) x^j$ if and only if $ep^i x^i \geq p^i x$, and let $R^*(e)$ be the transitive closure of $R^0(e)$. Define GARP(e) as

$$\text{GARP}(e) \Leftrightarrow \text{ If } x^i R^*(e) x^j \text{ does not imply } ep^j x^j > p^j x^i.$$

Then the AEI is the largest number such that GARP(e) is satisfied.

Other measures include Varian's (1985) minimum perturbation test, based on the minimal movements of the data needed to accept the null hypothesis of utility maximization; Famulari's (1995) violation rate, which is the proportion of combinations that form violations among observations for which violations can be expected; and comparison of the observed number of violations with the maximum number of

[1] See, for example, Sippel (1997), Mattei (2000), Harbaugh et al. (2001), Andreoni and Miller (2002), Février and Visser (2004), Choi et al. (2007).

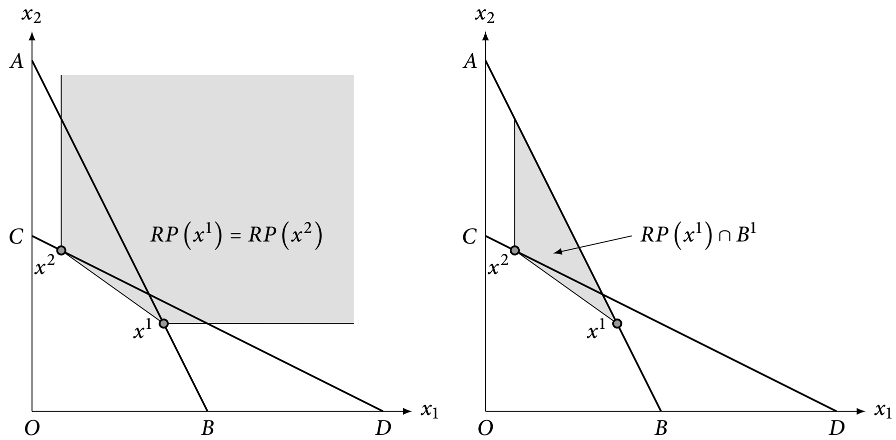

Figure 1: Left: Two observations which violate GARP. The shaded area gives the set of all bundles revealed preferred to x^1 and x^2. Since x^1 and x^2 form a preference cycle the sets are necessarily identical. Right: The intersection of $RP(x^1)$ with the budget line AB on which x^1 was chosen. Note that the maximal possible area of intersection is the area of $ABCA$.

violations possible, as applied by Swofford and Whitney (1987) and McMillan and Amoako-Tuffou (1988). See Gross (1995) for a survey of measures and his own test, which is based on an estimate of wasted expenditure.

2.3 The New Measure

Obviously, if a consumer makes decisions that are incompatible with GARP, then at least for one choice the indifference curve through that point, as implied by the other choices, intersects the budget line he made the choice on.[2] The idea of our measure is to ask, "how much of a given budget did a consumer reveal to prefer to the actual choice he made on the budget?". To answer the question, we take the upper bound of the indifference curve through a choice x^i and compute the area between that curve and the budget line. That is to say, we compute the area of the intersection of the two sets B^i and $RP\left(x^i\right)$. This basic idea is illustrated in Figure 1 and 2.

[2] Note that for illustrative purposes, we occasionally use terms only applicable to the two dimensional case.

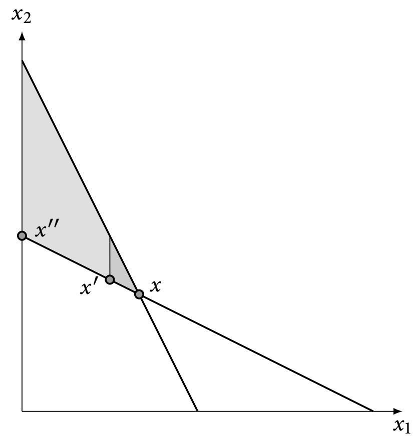

Figure 2: Both $\{x', x\}$ and $\{x'', x\}$ lead to the same Afriat efficiency index of 1 - ε, but have different volume violation indices.

The "size" of an intersection of B^i and $RP(x^i)$ is an area in two dimensions, and a volume in three dimensions. For simplicity, the generalization to arbitrary dimensions (the "hypervolume") will be also be called volume[3] and denoted by vol(Polytope). For example, the volume of an ℓ-dimensional hypercube h with edge length a is $vol(h) = a^\ell$.

Denote the volume of the intersection of a budget B^i and all bundles revealed preferred to x^i by

$$V(x^i) = vol\left(RP\left(x^i\right) \cap B^i\right). \tag{2}$$

Obviously, if S satisfies GARP, $V(x^i) = 0$ for all $i = 1, \ldots, n$.

To compare the extent of violation of GARP between many consumers who all made decisions on the same budgets, $V(x^i)$ does not have to be adjusted. However, if consumers made decisions on different budgets, the magnitude of $V(x^i)$ can be misleading. We therefore normalize the volume in the following way:

[3]The generalization of an area or volume to higher dimensions is also known as the *content*. See Weisstein (2008).

Denote the ratio of $V(x^i)$ to the maximum size of $V(x^i)$ possible as

$$V^m(x^i) = \frac{V(x^i)}{\max V_i}, \tag{3}$$

where

$$\max V_i = vol\left(H_{\text{convex}}\left(\left\{x \in X : p^i x \leq w^i, p^j x = w^j \ \forall\ j \in \{1,\ldots,n\} - \{i\}\right\}\right)\right).$$

Given the $V(x^i)$, we would like an index that aggregates the different intersections. One obvious way to define the index is the mean of all $V(x^i)$.[4] Denote by

$$\text{VI}^{\text{mean}} = \frac{1}{n}\sum_{i=1}^{n} V^m(x^i) \tag{4}$$

an index using the mean of all $V(x^i)$. Because V^m is bound between 0 and 1, we can define the *volume efficiency index* (VEI) as

$$\text{VEI} = 1 - \text{VI}^{\text{mean}} \tag{5}$$

In two dimensions, computation is fairly simply. For higher dimensions, we use the program qhull, which implements the quick hull algorithm for convex hulls (see Barber et al. 1996).

2.4 Power against Random Behavior

Depending on the characteristic of the budget sets, the chance of violating GARP can differ substantially. A completely rational consumer will always be consistent and is not in "danger" of violating GARP. However, even a consumer who makes purely random decision has a chance to satisfy GARP. Bronars (1987) suggests a Monte Carlo approach to determine the power the test has against random behavior. The approximate power of the test is the percentage of random choices which violated GARP. Bronars' first algorithm follows Becker's (1962) example by inducing a uniform distri-

[4] Another option is to take the maximal of all $V(x^i)$, or the median. The results are robust with respect to the aggregation method.

bution across the budget hyperplane. Using Bronars' second algorithm, the random choices are generated by drawing ℓ i.i.d uniform random variables, $z_1, \ldots, z_\ell$, for each price vector, and calculate budget shares $Sh_i = z_i / \sum_{j=1}^{\ell} z_j$. The random demand for commodity x_i is then calculated as $x_i = (Sh_i\, w)/p_i$.

One way to utilize Bronars' power is to compute it based on a range of values for an index that measures the severity of the violations, i.e. the power of the test is the percentage of random choices which do "worse" than the value of the index. For example, for an VEI or AEI of .9, compute the percentage of random choices which have a VEI or AEI lower than .9. This will give an idea of how much power the test loses if we allow consumers to deviate from optimizing behavior. See Section 3 for details.

2.5 Theoretical Considerations

In Section 3 we will evaluate the new measure based on Monte-Carlo experiments. However, it should already be pointed out that while the new measure is quite intuitive, it has a theoretical shortcoming. The AEI can be related to wasted *absolute income*, which is a real magnitude. The volume efficiency index is related to the fraction of the budget which is preferred to the actual decision, and puts equal weight on fractions of the same volume. Neither does it tell us anything about wasted income, nor does it say much about wasted utility.

Consider Figure 2. If we move x just a bit upwards on the steeper of the two budget lines we can find a utility function that rationalizes $\{x, x'\}$ and $\{x, x''\}$. Suppose that the data was collected with a small measurement error and that the consumer's actual decision was indeed a bit to the upper left of the observation x. While the AEI is 1 (or $1 - \varepsilon$) and raises little concern about the rationality of the consumer, the VEI suggests a small but substantial deviation from utility maximization if the data is $\{x, x'\}$ and a relatively large deviation if the data is $\{x, x''\}$.

To understand this unrobust behavior of the volume efficiency index, note that the set of observations $\{x, x'\}$ or $\{x, x''\}$ only implies that the shaded area is revealed preferred to x, but nothing can be said about *how much* it is preferred. Consider Figure 3, and suppose x^1 and x^2 are the actual decisions instead of the observed ones in

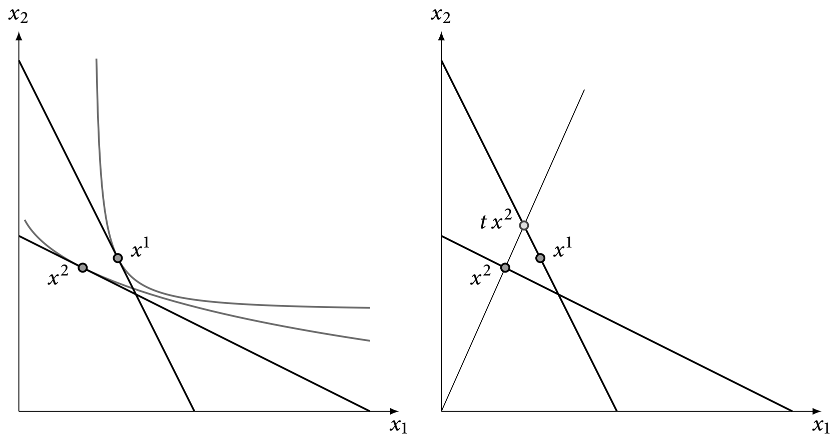

Figure 3: Left: Wasting a large fraction of the budget does not necessarily mean that the achieved utility level could have been a lot higher. Right: The set $\{x^1, x^2\}$ does not satisfy homotheticity, because $t\,x^2$ would then be strictly preferred to x^1. Observations with measurement error are less likely to cause a low volume efficiency index if the underlying preferences are homothetic.

Figure 2. Suppose the two indifference curves represent utility levels which, in absolute terms, are only marginally different. Then no bundle in the part of the budget which is preferred to x in Figure 2 adds much utility to x.

Based on these considerations we can expect the volume efficiency index to be more robust if the underlying preferences are homothetic or "almost" homothetic. In the diagram on the bottom of Figure 3 the bundle x^2 is projected on budget 1. Homotheticity implies that $t\,x^2$ is preferred over all bundles to the right of $t\,x^2$ on budget 1, so the actual decision made on budget 1 would have to be on the left of $t\,x^2$. This implies that if we measure decisions of a homothetic consumer with some measurement error or slight failures in the maximization process it is unlikely that we observe a decision pattern as the one depicted in Figure 2.

Notwithstanding these theoretical concerns, the results of the Monte-Carlo experiments in the next section imply that the volume efficiency index can be useful in applications.

3 Comparison: Power Against Random Behavior

3.1 Procedure

Varian (1990), perhaps in a somewhat playful manner, suggests a 95% Afriat efficiency level as the critical value to decide which GARP-violating sets of observations to accept as utility maximizing, "for sentimental reasons". There is, however, no natural critical value, and the AEI can be difficult to interprete without a reference. We therefore suggest to generate random choices on the budget sets and to recompute Bronars' power for all efficiency levels between 0 and 1. This will give us an idea of how much power the test loses if we accept GARP-violating observations as close enough to utility maximizing. This procedure also allows us to compare the AEI with the VEI.

To evaluate the two indices, we take data from known generating functions and add stochastic error to simulate measurement error. Ideally, we would like to accept all of the thusly obtained sets as utility maximizing without thereby reducing the power of the applied test.

We use a similar procedure as applied in Fleissig and Whitney (2003, 2005). First, we generate data from a five commodity utility function. The first function is a Cobb-Douglas type utility function given by

$$U^{\mathrm{CD}}(x^*) = \prod_{i=1}^{5} x_i^{*\alpha i}, \qquad \text{with } \sum_{i=1}^{5} \alpha_i = 1 \tag{6}$$

The second utility function is a non-homothetic utility function with variable elasticity of substitution:

$$U^{\mathrm{VES}}(x^*) = \sum_{i=1}^{5} \alpha_i \left(x_i^* - \gamma_i\right)^{\beta i}, \qquad \text{with } \sum_{i=1}^{5} \alpha_i = \sum_{i=1}^{5} \beta_i = 1 \tag{7}$$

For the Cobb-Douglas utility function we use a set of fixed parameters taken from Fleissig and Whitney (2003):

$$\alpha: \quad \alpha_1 = .4,\ \alpha_2 = .3,\ \alpha_3 = .15,\ \alpha_4 = .1,\ \alpha_5 = .05. \tag{8}$$

For the VES utility function we use a set of fixed parameters, with α_i as before and

$$\beta: \ \beta_1 = .1,\ \beta_2 = .3,\ \beta_3 = .15,\ \beta_4 = .3,\ \beta_5 = .15 \tag{9}$$

$$\gamma: \ \gamma_1 = -3,\ \gamma_2 = 2,\ \gamma_3 = 0,\ \gamma_4 = -4,\ \gamma_5 = 4 \tag{10}$$

We repeat the computations for both functions with a different set of random parameters each time by drawing each α_i and each β_i from a uniform distribution $\mathcal{U}[.05, .95]$ and then normalizing it such that $\sum_{i=1}^{5} \alpha_i = \sum_{i=1}^{5} \beta_i = 1$. We draw each γ_i from a uniform distribution $\mathcal{U}[-5, 5]$.

For the Monte-Carlo experiment we assume that we observe the demand according to the given utility function with some measurement error that fluctuates by κ% around the true demand; we use $\{\kappa_1, \kappa_2, \kappa_3\} = \{.05, .1, .2\}$.

The datasets have $n_1 = 20$ observations each, repeated with $n_2 = 40$, with expenditure w drawn from a uniform distribution $\mathcal{W} \sim \mathcal{U}[10000, 12000]$. Price vectors are drawn from a uniform distribution $\mathcal{P}_1 \sim \mathcal{U}[95, 100]$. The same steps are repeated with a distribution $\mathcal{P}_2 \sim \mathcal{U}[90, 100]$. These expenditures and prices lead to many intersections of budget sets which can lead to many violations of GARP.

To summarize, we use 48 different settings, each one being an element of $\{\text{CD}, \text{VES}\} \times \{\text{fixed, random parameters}\} \times \{\kappa_1, \kappa_2, \kappa_3\} \times \{n_1, n_2\} \times \{\mathcal{P}_1, \mathcal{P}_2\}$.

The data are generated by the following steps:

A1 Randomly draw $n \in \{n_1, n_2\}$ expenditure observations from a uniform distribution $\mathcal{W}$ and n price vectors for which each element is drawn from a uniform distribution $\mathcal{P} \in \{\mathcal{P}_1, \mathcal{P}_2\}$.

A2 Generate utility maximizing demand for each budget, using the respective functional form and set of parameters. Denote the maximizing decisions x_i^* for $i = 1, \ldots, 5$.

A3 Generate demands with measurement error by multiplying $x_1^*, \ldots, x_5^*$ by a uniform random number, so that $x_i = x_i^*(1 + \varepsilon_i)$ for $i = 1, \ldots, 5$, where $\varepsilon_i \sim \mathcal{U}[-\kappa, \kappa]$ and $\kappa \in \{.05, .1, .2\}$. To keep expenditure constant, normalize the x_i by multiplying each with $\lambda = w/(p \cdot x)$.

A4 Repeat steps A1 – A4 many times, say 10,000.

To approximate the power the GARP test has if we allow deviations from utility maximization, we need to generate random choices on the budget sets:

B1 Generate budgets as in step A1.

B2 Generate random choices on the budget sets of step B1, following Bronars' first and second algorithm: (1) Draw a random point *SP* from the 4-simplex using a simplex point picking algorithm. The random demand for commodity x_i is then calculated as $x_i = (SP_i w)/p_i$. (2) Draw five i.i.d uniform random variables, $z_1, \ldots, z_5$, for each price vector, and calculate budget shares $Sh_i = z_i / \sum_{i=1}^{5} z_j$. The random demand for commodity x_i is then calculated as $x_i = (Sh_i w)/p_i$.

B3 Repeat steps B1 and B2 many times, at least as often as with A1 – A4, say 20,000.

The final step is to compute the loss of power of the test for all possible AEI and VEI:

C1 Generate utility maximizing sets of observations with stochastic error, following procedure A. Then for each set of *n* budgets, compute the AEI and the VEI.

C2 Generate sets of observations following procedure B. Again, compute the AEI and the VEI for each set.

C3 Sort the sets from C1 by their AEI and VEI, respectively. For each set from C1, compute the percentage of sets from step C2 that have a higher AEI and VEI, respectively.

3.2 Results

3.2.1 Descriptive Statistics

Table 1 reports the Bronars' Power of the data sets generated by procedure B. Note that in the three sets with a power of 100% indeed not a single set of choices was consistent with GARP; of course, the real power has to be less than this approximation.

		ALGORITHM	
		FIRST	SECOND
20 BUDGETS	$\mathcal{P}_1$	96.57%	90.80%
	$\mathcal{P}_2$	99.62%	99.22%
40 BUDGETS	$\mathcal{P}_1$	100.00%	99.98%
	$\mathcal{P}_2$	100.00%	100.00%

Table 1: Bronars' Power of the used data sets.

Table 2 reports the fraction of choice sets generated by procedure A which are inconsistent with GARP. The choice of parameters does not seem to be important. What is perhaps a bit surprising is that the fraction of GARP-consistent sets is higher for the second utility function.

3.2.2 Loss of Power

The main result is that for all of the different data sets we generated, the loss of power is mostly smaller and never greater if the cutoff point is based on the VEI rather than the AEI. Perhaps a bit surprisingly, the result is robust with respect to the functional form. It suggests that the VEI is better suited than the AEI to capture small deviations from utility maximization and distinguish between a set of decisions that are close to utility maximizing on the one hand and purely random behavior on the other hand.

For the sets with $n_2 = 40$ budgets, the difference is less pronounced. This is not surprising because the power of the setup for these cases is already very close to 100%;

		COBB-DOUGLAS FUNCTION					
		FIXED			RANDOM		
		$\kappa = .05$	$\kappa = .1$	$\kappa = .2$	$\kappa = .05$	$\kappa = .1$	$\kappa = .2$
n_1	$\mathcal{P}_1$	3.43%	13.06%	31.45%	3.04%	12.19%	31.64%
	$\mathcal{P}_2$	1.08%	11.50%	39.51%	1.08%	10.49%	39.39%
n_2	$\mathcal{P}_1$	12.01%	41.03%	77.83%	12.37%	41.47%	77.99%
	$\mathcal{P}_2$	4.77%	38.46%	86.58%	4.25%	35.39%	85.95%
		VES FUNCTION					
		FIXED			RANDOM		
		$\kappa = .05$	$\kappa = .1$	$\kappa = .2$	$\kappa = .05$	$\kappa = .1$	$\kappa = .2$
n_1	$\mathcal{P}_1$	1.35%	7.00%	20.70%	1.89%	7.09%	20.63%
	$\mathcal{P}_2$	0.44%	4.82%	22.77%	0.39%	6.36%	31.88%
n_2	$\mathcal{P}_1$	5.35%	25.97%	60.63%	7.01%	34.65%	74.25%
	$\mathcal{P}_2$	1.49%	19.01%	65.84%	1.46%	23.27%	79.07%

Table 2: Fractions of GARP-inconsistent choice sets.

the procedure used here is mainly useful for applications in which the test power is less than perfect.

Figures 4 and 5 (on pages 20 and 21) report the proportion of utility maximizing observations with stochastic error that are accepted as “consistent enough” with GARP, depending on the desired power of the test (from the Cobb-Douglas function with fixed parameters, $\mathcal{P}_1$ and $n_1 = 20$, using the random choices generated by Bronars’ first and second algorithm, respectively). Figure 6 (on page 22) report the same results for the VES function (Bronars’ second algorithm). In all cases, we lose less test power when basing decisions on the AEI.

Tables 3 and 4 report the retained test power when all utility maximizing choices with stochastic error are accepted as utility maximizing, using Bronars’ first and second algorithm, respectively. The tables only report the results for fixed parameters; the results for random parameters are largely the same.

		Cobb-Douglas Function					
		AEI			VEI		
		$\kappa = .05$	$\kappa = .1$	$\kappa = .2$	$\kappa = .05$	$\kappa = .1$	$\kappa = .2$
n_1	$\mathcal{P}_1$	95.21%	91.88%	73.21%	96.56%	96.43%	94.83%
	$\mathcal{P}_2$	99.53%	98.87%	92.37%	99.62%	99.61%	99.41%
n_2	$\mathcal{P}_1$	99.98%	99.94%	98.41%	99.99%	99.99%	99.99%
	$\mathcal{P}_2$	100.00%	100.00%	99.92%	100.00%	100.00%	100.00%
		VES Function					
		AEI			VEI		
		$\kappa = .05$	$\kappa = .1$	$\kappa = .2$	$\kappa = .05$	$\kappa = .1$	$\kappa = .2$
n_1	$\mathcal{P}_1$	95.76%	93.04%	86.49%	96.57%	96.44%	95.88%
	$\mathcal{P}_2$	99.58%	99.33%	96.46%	99.62%	99.62%	99.58%
n_2	$\mathcal{P}_1$	99.99%	99.96%	99.82%	100.00%	100.00%	99.99%
	$\mathcal{P}_2$	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 3: The retained test power when all choices are accepted as utility maximizing, using Bronars' first algorithm.

4 Application to Experimental Data

Andreoni and Miller (2002) report results of an experimental dictator game. It was designed to test rationality of altruistic preferences. It is a generalized dictator game in which one subject (the dictator) allocates token endowments between himself and an anonymous other subject (the beneficiary) with different transfer rates. The payoffs of the dictator and the beneficiary are interpreted as two distinct goods, and the transfer rates as the price ratio of these two goods.

Figure 7 (on page 23) reports the proportion of subjects which are accepted as "consistent enough" with GARP, depending on the desired power of the test, using Bronars' first algorithm (results are similar for the second algorithm). Contrary to the results from Section 3, we lose somewhat less power if we base decisions on the AEI.

		COBB-DOUGLAS FUNCTION					
		AEI			VEI		
		$\kappa = .05$	$\kappa = .1$	$\kappa = .2$	$\kappa = .05$	$\kappa = .1$	$\kappa = .2$
n_1	$\mathcal{P}_1$	87.66%	78.66%	44.27%	90.75%	90.20%	82.98%
	$\mathcal{P}_2$	97.66%	94.16%	71.32%	98.22%	98.10%	96.30%
n_2	$\mathcal{P}_1$	99.92%	99.35%	83.40%	99.98%	99.97%	99.91%
	$\mathcal{P}_2$	99.99%	99.99%	97.50%	100.00%	100.00%	100.00%
		VES FUNCTION					
		AEI			VEI		
		$\kappa = .05$	$\kappa = .1$	$\kappa = .2$	$\kappa = .05$	$\kappa = .1$	$\kappa = .2$
n_1	$\mathcal{P}_1$	88.76%	81.82%	66.80%	90.79%	90.28%	88.04%
	$\mathcal{P}_2$	97.95%	96.63%	84.49%	98.22%	98.20%	97.73%
n_2	$\mathcal{P}_1$	99.97%	99.70%	97.89%	99.98%	99.97%	99.97%
	$\mathcal{P}_2$	99.99%	99.99%	99.89%	100.00%	100.00%	100.00%

Table 4: The retained test power when all choices are accepted as utility maximizing, using Bronars' second algorithm.

Interpretation of this result is not straightforward. First note that the data set is too small to draw strong conclusions, as there were 142 subjects of which only 13 violated GARP. Secondly, we do not know for sure whether these 13 subjects are actually utility maximizers who made only minor errors. Comparing the results with the results from the previous section suggests that the data from the subjects who violated GARP is indeed not so much a result of imperfect utility maximization but rather random behavior. Another possible explanation is that even the VES utility function used to generate data in the previous sections is still a bit too idealized.

5 Discussion and Conclusion

In this paper a new measure for the severity of a violation of utility maximization, the volume efficiency index, was suggested. The measure is based on the extent to which the upper bound of the indifference surface of a decision intersects the budget on which the decision was made. This measure has several advantages.

The measure is intuitively appealing as it can be easily illustrated with graphical tools covered in any intermediate course in microeconomic theory. In two dimensions the measure is easy to compute. It provides a convenient way of relating the extent of a violation to the maximum extent possible. It performs very well as a cut-off rule for determining whether or not observations on a single consumer can still be considered "close enough" to maximizing behavior.

A disadvantage is the computational effort needed to compute the measure in high dimensions.[5] However, note that the dimension of most data obtained by laboratory experiments is naturally bounded.

[5] From experimentation with simulated data it seems that even Monte Carlo experiments are still quite feasible in six dimensions and 40 observation points per consumer; in fact, it takes about as much time as computing the AEI.

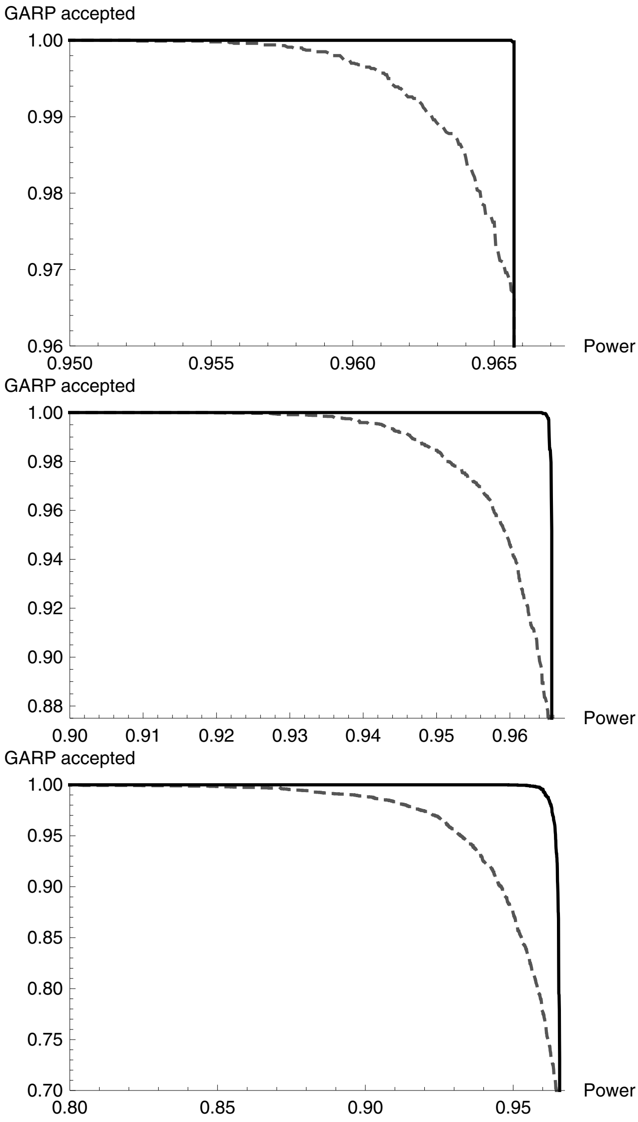

Figure 4: Results for the Cobb-Douglas function, using $\mathcal{P}_1$, n_1, fixed parameters, Bronars' first algorithm. Top: κ = .05, middle: κ = .1, bottom: κ = .2. The figure reports the proportion of utility maximizing observations that are accepted as consistent with GARP, depending on the desired power of the test. The dashed line gives the proportion of accepted observations according to the AEI, and the solid line gives the proportion according to the VEI.

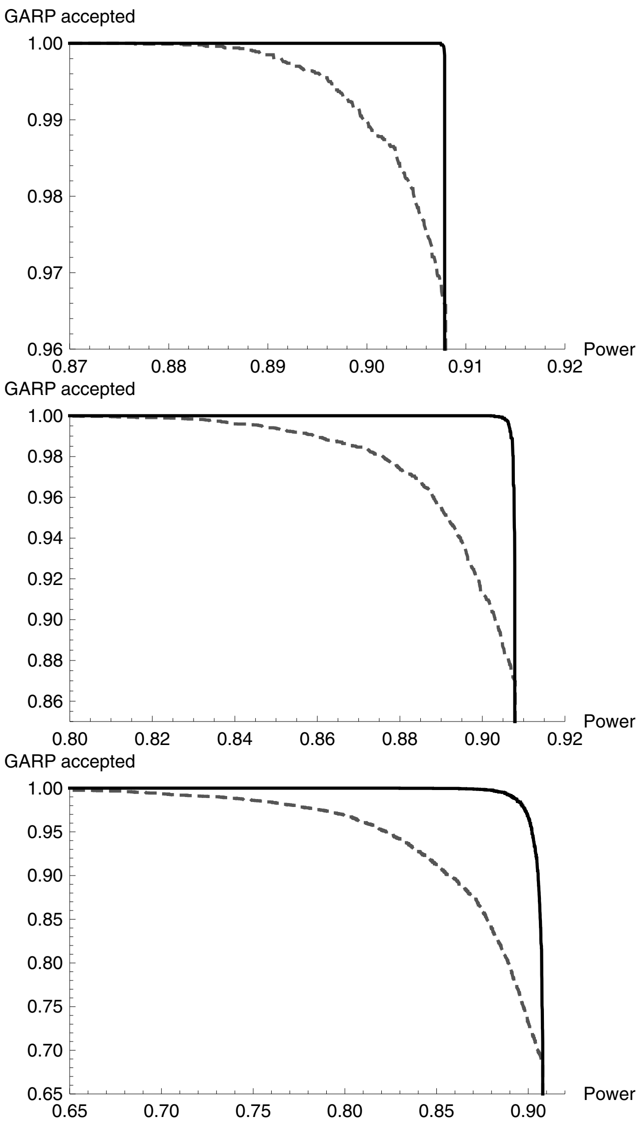

Figure 5: Same as for Figure 4, but using Bronars' second algorithm.

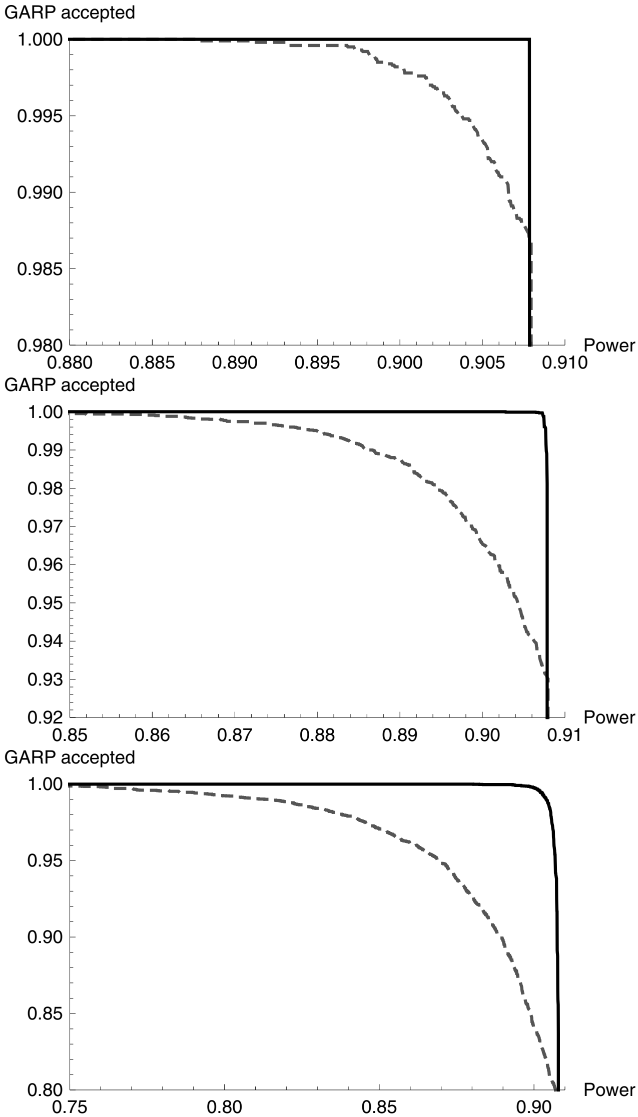

Figure 6: Same as for Figure 5, but for VES function.

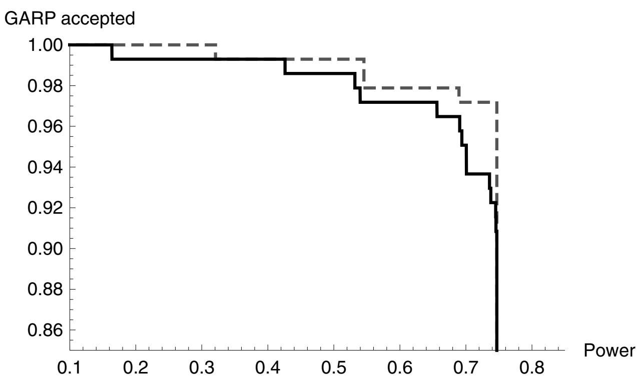

Figure 7: Results for data from Andreoni and Miller (2002) with Bronars' first algorithm. The figure reports the proportion of utility maximizing observations that are accepted as consistent with GARP, depending on the desired power of the test. The dashed line gives the proportion of accepted observations according to the AEI, and the solid line gives the proportion according to the VEI.

References

Afriat, S. N. (1967): The Construction of a Utility Function From Expenditure Data, *International Economic Review*, 8(1), 67–77.

——— (1972): Efficiency Estimation of Production Functions, *International Economic Review*, 13(3), 568–598.

Andreoni, J. and J. Miller (2002): Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism, *Econometrica*, 70(2), 737–753.

Barber, C. B., D. P. Dobkin, and H. Huhdanpaa (1996): The Quickhull Algorithm for Convex Hulls, *ACM Transactions on Mathematical Software*, 22(4), 469–483, http://www.qhull.org.

Becker, G. S. (1962): Irrational Behavior and Economic Theory, *Journal of Political Economy*, 70(1), 1–13.

Bronars, S. G. (1987): The Power of Nonparametric Tests of Preference Maximization, *Econometrica*, 55(3), 693–698.

Choi, S., R. Fisman, D. M. Gale, and S. Kariv (2007): Revealing Preferences Graphically: An Old Method Gets a New Tool Kit, *American Economic Review*, 97(2), 153–158.

Famulari, M. (1995): A Household-Based, Nonparametric Test of Demand Theory, *Review of Economics and Statistics*, 2(2), 285–382.

Février, P. and M. Visser (2004): A Study of Consumer Behavior Using Laboratory Data, *Experimental Economics*, 7, 93–114.

Fleissig, A. R. and G. A. Whitney (2003): A New PC-Based Test for Varian's Weak Separability conditions, *Journal of Business and Economic Statistics*, 21(1), 133–144.

——— (2005): Testing for the Significance of Violations of Afriat's Inequalities, *Journal of Business and Economic Statistics*, 23(3), 355–362.

Gross, J. (1995): Testing Data for Consistency with Revealed Preference, *Review of Economics and Statistics*, 77(4), 701–710.

Harbaugh, W. T., K. Krause, and T. R. Berry (2001): GARP for Kids: On the Development of Rational Choice Behavior, *American Economic Review*, 91(5), 1539–1545.

Knoblauch, V. (1992): A Tight Upper Bound on the Money Metric Utility Function, *American Economic Review*, 82(3), 660–663.

Mattei, A. (2000): Full-scale real tests of consumer behavior using experimental data, *Journal of Economic Behavior and Organization*, 43, 487–497.

McMillan, M. L. and J. Amoako-Tuffou (1988): An Examination of Preferences for Local Public Sector Outputs, *Review of Economics and Statistics*, 70(1), 45–54.

Sippel, R. (1997): An Experiment on the Pure Theory of Consumer's Behavior, *The Economic Journal*, 107, 1431–1444.

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(3), 458–464.

Varian, H. R. (1982): The Nonparametric Approach to Demand Analysis, *Econometrica*, 50, 945–972.

——— (1985): Non-Parametric Analysis of Optimizing Behavior With Measurement Error, *Journal of Econometrics*, 30, 445–458.

——— (1990): Goodness of Fit for Revealed Preference Tests, university of Michigan CREST Working Paper Number 13.

Weisstein, E. W. (2008): Content, from MathWorld – A Wolfram Web Resource. http://mathworld.wolfram.com/Content.html.