Goodness-of-Fit Measures for Revealed Preference Tests: Complexity Results and Algorithms

BART SMEULDERS, FRITS C. R. SPIEKSMA, and LAURENS CHERCHYE,
University of Leuven
BRAM DE ROCK, Université Libre de Bruxelles

We provide results on the computational complexity of goodness-of-fit measures (i.e., Afriat's efficiency index, Varian's efficiency vector-index, and the Houtman-Maks index) associated with several revealed preference axioms (i.e., WARP, SARP, GARP, and HARP). These results explain the computational difficulties that have been observed in literature when computing these indices. Our NP-hardness results are obtained by reductions from the independent set problem. We also show that this reduction can be used to prove that no approximation algorithm achieving a ratio of $O(n^{1-\delta})$, $\delta > 0$ exists for Varian's index, nor for Houtman-Maks' index (unless P = NP). Finally, we give an exact polynomial-time algorithm for finding Afriat's efficiency index.

Categories and Subject Descriptors: F.2.0 [**Analysis of Algorithms and Problem Complexity**]: General; G.2.2 [**Discrete Mathematics**]: Graph Theory—*Graph algorithms*

General Terms: Algorithms, Economics

Additional Key Words and Phrases: Computational complexity, goodness-of-fit measures, NP-complete, revealed preference, testing individual rationality

ACM Reference Format:
Smeulders, B., Spieksma, F. C. R., Cherchye, L., and De Rock, B. 2014. Goodness-of-fit measures for revealed preference tests: Complexity results and algorithms. *ACM Trans. Econ. Comp.* 2, 1, Article 3 (March 2014), 16 pages.
DOI: http://dx.doi.org/10.1145/2560793

1. INTRODUCTION

Utility maximization is a core hypothesis in neoclassical microeconomics, and testing the empirical validity of this assumption has attracted considerable attention in the literature. Such tests based on revealed preference theory have become increasingly popular. An attractive feature of these tests is that they are intrinsically nonparametric: they check consistency with the utility-maximization hypothesis without requiring a (typically nonverifiable) functional specification of the utility function, and so they

An extended abstract corresponding to an earlier version of this article appeared in *Proceedings of the 4th International Workshop on Computational Social Choice (COMSOC'12)*.
This research has been partially funded by FWO grant G.0447.10, and by the Inter-university Attraction Poles Programme initiated by the Belgian Science Policy Office. B. De Rock was supported by a Starting Grant from the European Research Council (ERC). L. Cherchye was supported by the Research Fund K.U.Leuven through grant STRT1/08/004.
Authors' addresses: B. Smeulders (corresponding author) and F. C. R. Spieksma, ORSTAT, Faculty of Economics and Business, University of Leuven; email: bart.smeulders@kuleuven.be; L. Cherchye, CES, Faculty of Economics and Business, University of Leuven; B. De Rock, ECARES-ECORE, Université Libre de Bruxelles.

DOI: http://dx.doi.org/10.1145/2560793

maximally avoid the risk of erroneous conclusions due to a misspecified functional form. The empirical requirements for utility maximization are summarized in terms of revealed preference axioms, which can be directly applied to consumption data (i.e., prices and quantities) without requiring auxiliary assumptions. For example, a key result of revealed-preference theory is that consumption can be represented as maximizing a (well-behaved) utility function if and only if it satisfies the Generalized Axiom of Revealed Preference (GARP) [Varian 1982]. Three other axioms that are most frequently considered in the literature are the Weak, Strong, and Homothetic Axioms of Revealed Preference (WARP, SARP, and HARP; see Section 2 for exact definitions).

However, a frequently cited weakness of the basic revealed-preference tests is that they are 'sharp' tests: they only tell us whether or not observed behavior is consistent with the revealed-preference axiom that is being tested. When consumption data do not pass the test, there is no indication concerning the severity or the number of violations. To deal with this, a number of measures have been proposed in the literature to express how close a dataset is to satisfying rationality. In what follows, we will call these measures *goodness-of-fit* measures; they tell us how well a revealed-preference axiom fits the data at hand. Probably the most popular goodness-of-fit measure in applied work is Afriat's efficiency index (AI) [1973]. Other frequently used measures are the ones of Houtman and Maks (HI) [1985] and Varian (VI) [1990]. Section 2 provides a precise description of these alternative goodness-of-fit measures.

The revealed-preference axioms and goodness-of-fit measures have been used intensively in the literature. The first tests of the axioms of revealed preference go back to the sixties and seventies. Aggregated household consumption data was used in tests of SARP by Koo [1963, 1971], Koo and Hasenkamp [1972], Mossin [1972], and Landsburg [1981]. Varian [1982] tested GARP using similar data. Only Koo tried to measure the severity of the rejections by focusing on the number of violations and using a measure similar to HI. Over the last decade, the goodness-of-fit measures have been used more and more often. Sippel [1997] tests relaxations of WARP, SARP, and GARP related to AI. AI and GARP are used in papers by Mattei [2000], Harbaugh et al. [2001], Andreoni and Miller [2002], Février and Visser [2004], Choi et al. [2007, 2011], Dean and Martin [2010], and Burghart et al. [2013]; the last four papers also use HI. VI and GARP appear in Cox [1997], Mattei [2000], Choi et al. [2007, 2011], and Dean and Martin [2010]. For WARP, all three indices appear in Choi et al. [2007]. To the best of our knowledge, there do not exist any studies that compute goodness-of-fit measures for HARP, although there exist papers in which HARP is tested (see, e.g., Manser and McDonald [1988]). Finally, we also notice continuing interest in goodness-of-fit measures, illustrated by the recent introduction of several new indices in the literature, specifically, the money pump index by Echenique et al. [2011] which calculates the monetary cost of irrational behavior and the minimal swaps and minimal loss index by Apesteguia and Ballester [2011].

This article is specifically concerned with the computational complexity of the goodness-of-fit measures used in revealed-preference analysis. As such, it fits into the recent literature that utilizes insights from computational complexity theory in the study of interesting economic subjects.[1]

In general, computational complexity becomes an important issue if one wants to consider large datasets. In this respect, we indicate that large consumption datasets

[1]See, among many others, Gilboa and Zemel [1989], Chu and Halpern [2001], Cechlarova and Hajdukova [2002], Fang et al. [2002], Woeginger [2003], Baron et al. [2004, 2008], Brandt and Fisher [2008], Conitzer and Sandholm [2008], Kalyanaraman and Umans [2008], Procaccia and Rosenschein [2008], Galambos [2009], Hudry [2009], Brandt et al. [2010], Deb [2010], Apestequia and Ballester [2010], Cherchye et al. [2011], and Nobibon et al. [2011].

Table I. Overview of Results

	WARP	SARP	GARP	HARP
AI (sec 6)	$n^2 \log n$	$n^2 \log n$	$n^\alpha \log n$	n^3
VI (sec 4)	Inapproximable	Inapproximable	Inapproximable	Inapproximable
HI (sec 5)	Inapproximable	Inapproximable	Inapproximable	Inapproximable

are increasingly available (see, e.g., the scanner consumption data that nowadays can be used), which directly motivates the research question we consider here. Indeed, while the computational complexity of methods for testing GARP and the other revealed-preference axioms is well understood by now, this is not always the case for computing the previously mentioned goodness-of-fit measures.

It is generally thought that calculating AI is easy. However, to our knowledge, no exact algorithm is described in the literature. Varian [1990] provides an approximation algorithm which comes within $\left(\frac{1}{2}\right)^m$ of the true index value in m GARP tests. As for the other two indices (HI and VI), it has been empirically recognized that computing them is computationally intensive. For instance, Varian [1990] writes as follows.

> "Computing the set of efficiency indices [VI] that are as close as possible to 1 in some norm is substantially more difficult ... This approach is significantly more difficult from a computational perspective."

Similarly, Choi et al. [2011] state the following.

> "All indices [VI and HI] are computationally intensive for even moderately large data sets."

In fact, because of the apparent difficulty of exactly calculating VI, some authors have focused on designing approximate heuristics (see, e.g., Varian [1993] and Alcantud et al. [2010]).

The goal of the current article is to give a theoretical foundation for these practical observations and to strengthen the existing results. As far as we are aware, explicit complexity results are known only for index HI. More specifically, Houtman and Maks establish a link between their index for SARP and a feedback vertex set on a digraph, which implies NP-Hardness. Next, Dean and Martin [2010] state that HI for GARP is also NP-hard.

We establish the computational complexity for every combination of the three goodness-of-fit measures (AI, VI, and HI) and the four revealed-preference axioms (GARP, SARP, WARP, and HARP) already mentioned. We will refer to these problems as $\{A,V,H\}$I-$\{G,S,W,H\}$ARP, where choosing a symbol from the set $\{A,V,H\}$ and a symbol from the set $\{G,S,W,H\}$ identifies a particular problem. For example, AI-GARP is the problem of computing the maximum index AI such that the dataset satisfies a relaxation of GARP. Our main results are summarized in Table I, where a column corresponds to a specific axiom and a row to a specific measure, and where n stands for the number of observations.[2]

The rest of this article unfolds as follows. The next section sets the stage by introducing the basic revealed-preference concepts that we will use throughout. Section 2.2 provides a statement of the computational problems we focus on. Sections 4 and 5 present our results on the computational complexity for indices VI and HI. Section 6 contains a polynomial-time algorithm for AI. Section 7 concludes.

[2]*Inapproximable* means that no polynomial-time algorithm can achieve a constant-factor approximation unless $P = NP$. α represents the constant needed for matrix-multiplication.

2. REVEALED-PREFERENCE CONCEPTS

We start by stating the four revealed-preference axioms that we will consider. Subsequently, we present the different goodness-of-fit measures.

2.1. Axioms of Revealed-Preference

Our analysis starts from a dataset $S = \{(p_i, q_i) | \, i = 1, \ldots, n\}$, where p_i (q_i) is an N-dimensional vector of prices (quantities) corresponding to observation $i = 1, \ldots, n$. Without loss of generality, we will assume that prices are normalized such that $p_i q_i = 1$ for every observation i.

To define the concept of revealed preferences, we consider two observations i and j. If $(p_i q_i =) \; 1 \geq p_i q_j$, we say that bundle q_i is *directly revealed preferred* to bundle q_j. This is expressed by writing $q_i R_0 q_j$, where R_0 captures the direct revealed-preference relation. The transitive closure of R_0 is denoted by R and is called the indirect revealed-preference relation. If $1 > p_i q_j$, we say that bundle q_i is *strictly directly revealed preferred* to bundle q_j, which is denoted by $q_i P_0 q_j$. Finally, P stands for the transitive closure of P_0.

We can now state the four revealed-preference axioms that we consider in this article.

Definition 2.1 *Axioms of Revealed Preference*. A dataset S satisfies the following.

— *WARP*. If and only if for each pair of distinct bundles, q_i, q_j ($i, j = 1, \ldots, n$ with $i \neq j$): if $q_i R_0 q_j$, then it is not the case that $q_j R_0 q_i$.
— *SARP*. If and only if for each pair of distinct bundles, q_i, q_j ($i, j = 1, \ldots, n$ with $i \neq j$): if $q_i R q_j$, then it is not the case that $q_j R_0 q_i$.
— *GARP*. If and only if for each pair of distinct bundles, q_i, q_j ($i, j = 1, \ldots, n$ with $i \neq j$): if $q_i R q_j$, then it is not the case that $q_j P_0 q_i$.
— *HARP*. If and only if for every sequence of observations, $i, j, k, \ldots, l (= 1, \ldots, n)$: $\log(p_i q_j) + \log(p_j q_k) + \ldots + \log(p_l q_i) \geq 0$.

In other words, the main differences between the alternative axioms can be summarized as follows (see Varian [2006] for a more extensive discussion on the meaning of the axioms). Data consistency with WARP implies that the direct revealed-preference relation R_0 is asymmetric. By construction, this direct revealed-preference relation is not complete, since $q_i R_0 q_j$ and $q_j R_0 q_k$ do not need to imply $q_i R_0 q_k$. This is no longer the case for the revealed-preference relation R: by taking the transitive closure, we obtain a complete relation. Data consistency with WARP is a necessary condition for data consistency with SARP and implies that the revealed-preference relation is asymmetric (and complete).

Next, data consistency with SARP means that consumption behavior can be described as maximizing a utility function that generates single-valued demand. Similarly, data consistency with GARP means that consumption behavior can be described as maximizing a utility function that generates multivalued demand. As such, GARP is a generalization of SARP, and data consistency with GARP still makes the revealed-preference relation R an asymmetric and complete relation. Finally, data consistency with HARP means that consumption behavior can be described as maximizing a utility function that is homothetic. This implies that GARP is a necessary condition for HARP and that the revealed preference relation is asymmetric and complete. Figure 1 illustrates the relations between the different axioms of revealed preference.

2.2. Goodness-of-Fit Measures

In practice, direct application of any of the preceding revealed-preference axioms to some given dataset effectively obtains a 'sharp' test: a dataset either satisfies the

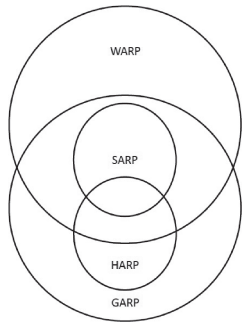

Fig. 1. Relations of the axioms of revealed preference.

axiom or it does not. In words, such a test allows us to conclude whether or not observed behavior is exactly consistent with the hypothesis of utility maximization (of a particular form, depending on whether we consider WARP, SARP, GARP, or HARP). However, a dataset that is not exactly consistent may actually be very close to consistency. For example, there may be only a limited number of observations that cause the observed violations of the axiom that is subject to testing. Or, the violations may be very insignificant in that small adjustments of the observations' expenditures (i.e., prices times quantities) may suffice to obtain consistency. Generally, it is interesting to quantify the degree to which a given dataset is close to consistency (see [Varian 1990] for extensive motivation).

To account for these considerations, a number of goodness of fit measures have been described in the literature. Three often used measures are Afriat's efficiency index (AI), Varian's efficiency vector index (VI) and the Houtman and Maks index (HI). Essentially, the indices AI and VI look for minimal expenditure perturbations to obtain consistency with the revealed preference axiom under evaluation: the AI index applies a common perturbation to all observations, while the VI index allows a different perturbation for each individual observation. Next, the index HI identifies the largest subset of observations that are consistent with the axiom. Essentially, this quantifies the degree of violation in terms of the number of observations that are involved in a violation of the revealed preference axiom that is tested. We refer to Varian [2006] for a more detailed discussion of the different goodness of fit measures we evaluate.

To formally introduce our goodness-of-fit measures, we make use of the vector $e = (e_1, e_2, \ldots, e_n)$, with $0 \leq e_i \leq 1$. This vector introduces an index e_i for each observation i, which relaxes the revealed-preference relations R_0 and P_0, as follows:

$$\text{if } e_i(= e_i p_i q_i) \geq p_i q_j, \text{then } q_i R_0(e) q_j,$$

$$\text{if } e_i(= e_i p_i q_i) > p_i q_j, \text{then } q_i P_0(e) q_j.$$

Analogously, $R(e)$ and $P(e)$ represent the transitive closures of $R_0(e)$ and $P_0(e)$. These newly defined relations $R_0(e)$, $P_0(e)$, $R(e)$, and $P(e)$ give rise to relaxed versions of the earlier axioms of revealed preference, which are defined for a given vector e. Clearly these axioms comply with the original versions of WARP, SARP, GARP, and HARP as soon as $e_i = 1$ for all i.

Definition 2.2. For a given $e = (e_1, e_2, \ldots, e_n)$, with $0 \leq e_i \leq 1$, a dataset S satisfies the following.

— *WARP(e)*. If and only if for each pair of distinct bundles, q_i, q_j $(i, j = 1, \ldots, n$ with $i \neq j)$: if $q_i R_0(e) q_j$, then it is not the case that $q_j R_0(e) q_i$.

— *SARP(e)*. If and only if for each pair of distinct bundles, q_i, q_j $(i, j = 1, \ldots, n$ with $i \neq j)$: if $q_i R(e) q_j$, then it is not the case that $q_j R_0(e) q_i$.
— *GARP(e)*. If and only if for each pair of distinct bundles, q_i, q_j $(i, j = 1, \ldots, n$ with $i \neq j)$: if $q_i R(e) q_j$, then it is not the case that $q_j P_0(e) q_i$.
— *HARP(e)*. If and only if for every sequence of observations, $i, j, k, \ldots, l (= 1, \ldots, n)$: $\log(p_i q_j) + \log(p_j q_k) + \ldots + \log(p_l q_i) \geq \log(e_i) + \log(e_j) + \ldots + \log(e_l)$.

To define the Afriat index (AI), we assume that $e_1 = \cdots = e_n$, which does indeed comply with a common perturbation for all observations. AI equals the supremum over all values for which the data is consistent with the tested revealed-preference axiom. More precisely, if $AI = 1$, then the data is consistent with the tested axiom. While if $AI < 1$, then this indicates that we need to pertubate the data to make it consistent with the revealed-preference axiom under study. The smaller the number AI is, the higher the perturbation or, alternatively, the more severe the rejection of the axiom. Finally, we notice that AI is well defined. If for a given e, the data is consistent with, for example, WARP(e), then the same holds for all $e' < e$. Indeed, by construction, we have that the revealed-preference relations in terms of e' are always a subset of the ones in terms of e (e.g., $R_0(e') \subseteq R_0(e)$).

The Varian index (VI) differs from AI by allowing for observation-specific perturbations. VI equals the vector e that is closest to one, for some given norm, such that the data satisfies the revealed-preference axiom under study. For example, if we use the quadratic norm, then VI should minimize $\sum_i (1 - e_i)^2$ such that, for example, WARP(e) is satisfied. Further, VI is subject to the same qualifications as AI.

Finally, the Houtman and Maks index (HI) equals the size of the largest subset of observations which satisfy the axioms of revealed preference. Formally, this complies with restricting the possible values of e_i so that $e_i \in \{0, 1\}$.

3. PROBLEM STATEMENT

In this section, we introduce the tools that we need to prove the results announced in Table I. In particular, in Section 3.1, we show how to reformulate the goodness-of-fit measures using graph theory, and in Section 3.2, we state the corresponding optimization problems.

3.1. Graph Representation

In order to verify whether a dataset actually satisfies some revealed-preference axiom, it is natural to construct a graph (see Koo [1971]). We now extend this procedure by taking into account a given vector $e = (e_1, \ldots, e_n)$. For some dataset S, we construct the associated graph $G_e(S)$. In this graph, there is a node for every observation. Next, for each pair of observations (i, j) $(i \neq j)$, there is an arc from node i to node j when $e_i \geq p_i q_j$. The length of this arc is equal to $p_i q_j - e_i$.

The graph $G_e(S)$ will be used to test WARP, SARP, and GARP. To test HARP, we make use of another graph $G'_e(S)$. The nodes and arcs of this alternative graph are defined in the same way as for graph $G_e(S)$, but now the length of the arc is given by $\log(p_i q_j) - \log(e_i)$.

The axioms of revealed preference can then be formulated as follows.

Definition 3.1. For a given $e = (e_1, e_2, \ldots, e_n)$, with $0 \leq e_i \leq 1$, the dataset S satisfies the following.

— *WARP(e)*. If and only if each cycle consisting of two arcs in graph $G_e(S)$ involves observations that have identical bundles.
— *SARP(e)*. If and only if each cycle in graph $G_e(S)$ contains only observations with identical bundles.

— *GARP(e)*. If and only if graph $G_e(S)$ does not contain any cycles of negative length.
— *HARP(e)*. If and only if graph $G'_e(S)$ does not contain any cycles of negative length.

3.2. Problem Descriptions

We are now in a position to define an optimization problem that measures how close a given dataset is to satisfying a particular axiom of revealed preference. This leads to twelve different problems. For example, for SARP(e), we obtain the problems AI-SARP, VI-SARP, and HI-SARP, each corresponding to a specific index. Straightforward adaptations define the problems AI-$\{S,G,H\}$ARP, VI-$\{S,G,H\}$ARP, and HI-$\{S,G,H\}$ARP. For compactness, we only state the optimization problems with respect to SARP; the optimization problems corresponding to $\{W,G,H\}$-ARP are defined analogously.

Problem 3.2 VI-SARP. Given a dataset S, for what values e_i, with $0 \leq e_i \leq 1$ for each i, is $\sum_{i=1}^{n} e_i$ maximized, while S satisfies SARP(e)?

Clearly, objective functions other than $\sum_{i=1}^{n} e_i$ are possible. We will come back to this issue in Section 4.

Problem 3.3 HI-SARP. Given a dataset S, what is the largest subset of observations $Q \subseteq \{1,\ldots,n\}$ such that Q satisfies SARP?

Results concerning this problem will be given in Section 5.

Problem 3.4 AI-SARP. Given a dataset S, for what value e_1, with $0 \leq e_1 \leq 1$, is e_1 maximized, while S satisfies SARP(e), with $e = (e_1,\ldots,e_1)$?

Our results for Afriat's index will be given in Section 6.

4. THE COMPLEXITY OF VARIAN'S INDEX

Clearly, when given a vector $e = (e_1,\ldots,e_n)$, there are different ways to specify an objective function measuring the quality of e. Obvious candidates are minimize $\sum_{i=1}^{n}(1-e_i)$, minimize $\sum_{i=1}^{n}(1-e_i)^2$, or minimize $\max_i (1-e_i)$. In fact, all these objective functions can be captured by considering minimize $\sum_{i=1}^{n}(1-e_i)^\rho$ for $\rho \geq 1$. Observe that, since *minimize* $\lim_{\rho\to\infty}\sum_{i=1}^{n}(1-e_i)^\rho$ is equivalent to *minimize* $\max_i (1-e_i)$, the Afriat index arises when $\rho \to \infty$. The results in this section are phrased for $\rho = 1$, that is, for the case where we minimize $\sum_{i=1}^{n}(1-e_i)$ or equivalently maximize $\sum_{i=1}^{n} e_i$. At the end of the section, we point out that the reduction remains valid for every fixed $\rho \geq 1$. Notice that for every fixed ρ, the problem is hard, while for $\rho \to \infty$, the problem becomes easy.

Let us now consider the following decision problem associated with VI-SARP (VI-SARP$_d$).

Input. A dataset $S = \{p_i, q_i \,|\, i = 1,\ldots,n\}$ and a number Z.
Question. Do there exist n numbers e_i, with $0 \leq e_i \leq 1$, such that

(1) the dataset S satisfies SARP(e), and
(2) $\sum_{i=1}^{n} e_i \geq Z$.

Obviously, being able to solve Problem 1 from Section 3.2, that is, being able to solve VI-SARP in polynomial time implies that VI-SARP$_d$ can be solved in polynomial time. That, however, is unlikely, as witnessed by our next result.

THEOREM 4.1. *VI-SARP$_d$ is NP-complete.*

PROOF. First, we show that the VI-SARP$_d$ is in NP. A certificate consists of a vector $(e_1, e_2, \ldots, e_n)$. Since calculating $\sum_{i=1}^{n} e_i$, building the graph $G_e(S)$ and testing $G_e(S)$ for acyclicity (see Definition 3.1) can be done in polynomial time, the problem is in NP.

Next, we prove that VI-SARP$_d$ is NP-hard by a reduction from the well-known NP-hard independent set problem [Karp 1972], which is formulated as follows.

Input. A graph $G = (V, E)$ and a number k.
Question. Does there exist a subset $V' \subseteq V$ of at least k vertices such that for every pair of vertices $i, j \in V'$, the edge (i, j) is not in E?

Given an instance of IS, we now construct the following instance of VI-SARP$_d$. For every node $i \in V$, there is an observation in VI-SARP$_d$: $n := |V|$. The vectors $p_i = (p_i^1, \ldots, p_i^N), q_i = (q_i^1, \ldots, q_i^N)$ are created as follows. We set, for $i = 1, \ldots, n$, $q_i^i := 1$, all remaining $q_i^j := 0$. Further, we set $p_i^i := 1$, for $i = 1, \ldots, n$. If there is an edge between node i and node j in G, that is, if $\{i, j\} \in E$, then $p_i^j := \epsilon$ (for some $0 < \epsilon < \frac{1}{n}$), otherwise $p_i^j := 2$. Finally, we set $Z := k$. This completes the description of the instance of VI-SARP$_d$. Notice that this construction implies that if an edge exists between i and j in G, then $p_i q_j = p_j q_i = \epsilon$, else $p_i q_j = p_j q_i = 2$.

We now argue the equivalence between IS and VI-SARP$_d$. Suppose the instance of independent set is a yes-instance, that is, an independent set of size at least k exists. For every vertex in that independent set, set $e_i = 1$, and for every other vertex, set $e_i = 0$. It is clear that $\sum e_i \geq Z$. Consider the graph $G_e(S)$, and recall that an arc is present from i to j if and only if $p_i q_j \leq e_i$. We claim that graph $G_e(S)$ is acyclic. Indeed, notice that vertices outside the independent set will not have any outgoing arcs in $G_e(S)$, since for each such vertex i: $p_i q_j - e_i = p_i q_j > 0$. Also notice that no arc connects two observations corresponding to nodes in the independent set, since for a pair of such observations i, j we have $p_i q_j - e_i = p_j q_i - e_j = 2 - 1 > 0$. Thus, arcs in $G_e(S)$ only exist from vertices in the independent set to vertices outside the independent set. It follows that the graph is acyclic.

Now, suppose that the instance of VI-SARP$_d$ is a yes-instance, so $\sum e_i \geq Z = k$. Then for at least k observations $e_i > \epsilon$; if not, at most $k - 1$ e_i-values exceed ϵ; since $e_i \leq 1$, $\sum e_i$ is then bounded by $k - 1 + (n - k - 1)\epsilon < k - 1 + 1 = k$, which contradicts with the requirements for a yes-instance. We will call such an e_i value *large*. We claim that the vertices with large e_i-values constitute an independent set in G. Indeed, consider two vertices i and j with a large e_i value. If i and j are connected in G, then $p_i q_j = p_j q_i = \epsilon$, implying that there is an arc in the graph $G_e(S)$ from i to j and from j to i, which is a cycle. Therefore i and j are not connected in G. Thus the set of vertices with large e_i is an independent set of size at least k. □

We now proceed with VI-GARP; its corresponding decision problem is VI-GARP$_d$.

Input. A dataset $S = \{p_i, q_i \,|\, i = 1, \ldots, n\}$ and a number Z.
Question. Do there exist n numbers e_i, with $0 \leq e_i \leq 1$, such that

(1) the dataset S satisfies GARP(e), and
(2) $\sum_{i=1}^{n} e_i \geq Z$.

THEOREM 4.2. *VI-GARP$_d$ is NP-complete.*

PROOF. This proof is an adaptation of the proof of Theorem 4.1. Observe that we are now interested in the question whether there exists a cycle in graph $G_e(S)$ that has negative length (see Definition 3.1).

Again, VI-GARP$_d$ is easily seen to be in NP. The certificate is the same, calculating $\sum_{i=1}^{n} e_i$, building graph $G_e(S)$ and testing Definition 3.1 can all be done in polynomial time.

The instance that we build using independent set is exactly the same as in the proof of Theorem 4.1. We now show the equivalence. From the proof of Theorem 4.1, we know that if an independent set of size k exists, we can find a vector e for which $\sum e_i \geq Z = k$ and SARP(e) is satisfied. As GARP(e) is a relaxation of SARP(e), GARP(e) holds as well.

Vice versa, we now argue that a yes-instance of VI-GARP$_d$ corresponds with an independent set of size at least k. Consider two nodes in G, i and j, and assume that both e_i and e_j are large. If i and j are connected in G, then an arc from i to j and an arc from j to i, both with negative length, are present in $G_e(S)$. This, however, is impossible, since the instance of VI-GARP$_d$ is a yes-instance. Thus since there are at least k observations with a large e-value, an independent set of size at least k exists in G. □

Next we consider the problem VI-WARP and its corresponding decision problem VI-WARP$_d$.

Input. A dataset $S = \{p_i, q_i \,|i = 1, \ldots, n\}$ and a number Z.
Question. Do there exist n numbers e_i, with $0 \leq e_i \leq 1$, such that

(1) the dataset S satisfies WARP(e), and
(2) $\sum_{i=1}^{n} e_i \geq Z$.

THEOREM 4.3. *VI-WARP$_d$ is NP-complete.*

PROOF. For VI-WARP$_d$, the proof of Theorem 4.1 is also easily adapted. Again, no changes are made to the graph construction, the only difference compared to VI-SARP$_d$ is that cycles are now allowed, as long as they involve more than two vertices (see Definition 3.1). Clearly, VI-WARP$_d$ is in NP.

From the proof of Theorem 4.1, we know that if an independent set of size k exists, we can find a vector e for which $\sum e_i \geq Z = k$ and graph $G_e(S)$ is acyclic. As it is acyclic, clearly no cycles involving only two vertices exist, and WARP(e) is satisfied.

Finally, if the instance of VI-WARP$_d$ is a yes-instance, we claim that an independent set of size k exists. As shown before, the observations for which the e_i value is large are not connected in the graph G, so they form an independent set. □

We end this section with the problem VI-HARP; here is the corresponding decision version.

Input. A dataset $S = \{p_i, q_i \,|i = 1, \ldots, n\}$ and a number Z.
Question. Do there exist n numbers e_i, with $0 \leq e_i \leq 1$, such that the following hold?

(1) The dataset S satisfies HARP(e).
(2) $\sum_{i=1}^{n} e_i \geq Z$.

THEOREM 4.4. *VI-HARP$_d$ is NP-complete.*

PROOF. As with the other axioms, VI-HARP$_d$ is in NP, since the certificate consists of the e_i values. Testing the graph for Definition 3.1 can be done using a minimum cycle mean algorithm. This algorithm identifies the cycle for which the mean weight of its arcs is the minimum mean weight of all cycles in the graph. If the mean weight of the arcs of this cycle is positive, HARP(e) is satisfied.

An instance of VI-HARP$_d$ is built in the same way as the instances in the proofs of the previous theorems. Consider now the equivalence. If an independent set of size k

exists in G, we choose $e_i = 1$ for observations corresponding to nodes in the independent set and $e_i = \epsilon^{n+1}$ for the other nodes. Observe that the resulting graph $G'_e(S)$ has the following properties.

(1) An arc in $G'_e(S)$ emanating from an observation corresponding to a node not in the independent set has a length of either $\log\epsilon - \log\epsilon^{n+1}$ or $\log 2 - \log\epsilon^{n+1}$.
(2) Every other arc in $G'_e(S)$ has length $\log\epsilon - \log 1$ or $\log 2 - \log 1$.

Any cycle that contains only observations corresponding to nodes within the independent set has positive length, since the length of each arc equals $\log 2$. Further, a cycle in $G'_e(S)$ going through an observation corresponding to a node not in the independent set contains an arc with length at least $\log\epsilon - \log\epsilon^{n+1}$. Hence, the length of this cycle is at least $\log\epsilon - \log\epsilon^{n+1} + n\log\epsilon = (n+1)\log\epsilon - \log\epsilon^{n+1} = 0$. Thus each cycle has nonnegative length, and the instance is a yes-instance of VI-HARP$_d$.

Consider now a yes-instance of VI-HARP$_d$. Clearly, there will be at least k observations with a large e_i value. Consider two nodes in G, each corresponding to an observation with a large e-value. If these two nodes are connected in G, arcs of length $\log\epsilon - \log e_i$ and $\log\epsilon - \log e_j$ are present, yielding a negative cycle. Thus, two observations with large e-values can not correspond to nodes that are connected in G. □

From the NP-completeness of the decisions problems, it follows that the optimization problems VI-$\{W,S,G,H\}$ARP are NP-hard. Next, we show that not only is computing Varian's index NP-hard, but also that approximating it in polynomial time is difficult, unless P = NP.

THEOREM 4.5. *For each fixed $\delta > 0$, the existence of a polynomial-time approximation algorithm VI $-$ $\{W,S,G,H\}$ achieving a ratio of $O(n^{1-\delta})$ implies P = NP.*

PROOF. Consider an instance of independent set and the corresponding instance of VI-$\{W,S,G,H\}$ARP, as constructed in Theorem 4.1. Clearly, if the optimum value of VI-$\{W,S,G,H\}$ARP instance equals z, then the optimum value for the IS instance equals $\lfloor z \rfloor$ (if not, then there exists an independent set of size $\lfloor z \rfloor + 1$, and by the previous reduction, we can find e so that $\sum_{i=1}^{n} e_i \geq \lfloor z \rfloor + 1$). Now, let z represent the optimum value of the instance of VI-$\{W,S,G,H\}$ARP, and let us assume that we have a polynomial-time approximation algorithm for VI-$\{W,S,G,H\}$ARP achieving a ratio $O(n^{1-\delta})$. Thus, more concretely, assume that we can compute in polynomial time a vector index e such that $\sum_{i=1}^{n} e_i \geq \frac{2z}{n^{1-\delta}}$. Given this vector index, we can find an independent set of size $max(1, \lfloor \frac{2z}{n^{1-\delta}} \rfloor)$. Since $n^{1-\delta} \times max(1, \lfloor \frac{2z}{n^{1-\delta}} \rfloor) \geq n^{1-\delta} \times \lfloor \frac{2z}{n^{1-\delta}} \rfloor \geq n^{1-\delta} \times \frac{\lfloor 2z \rfloor}{n^{1-\delta}+1} \geq n^{1-\delta} \times \frac{\lfloor 2z \rfloor}{2n^{1-\delta}} \geq \frac{\lfloor 2z \rfloor}{2} \geq \lfloor z \rfloor$, we conclude that

$$\max(1, \lfloor \frac{2z}{n^{1-\delta}} \rfloor) \geq \frac{\lfloor z \rfloor}{n^{1-\delta}}.$$

It follows that we have a polynomial-time algorithm achieving a ratio $n^{1-\delta}$ of for IS. Håstad [1999] and Zuckerman [2006] have shown that an approximation algorithm for IS running in polynomial time can not guarantee a ratio of $n^{1-\delta}$, unless P = NP. Our result follows. □

Let us now return to the general objective function $\sum_{i=1}^{n}(1-e_i)^{\rho}$ (with $\rho \geq 1$) given at the start of this section. We now consider the following problem.

Input. A dataset $S = \{p_i, q_i \,|i = 1, \ldots, n\}$ and a number Z.
Question. Do there exist n numbers e_i, with $0 \leq e_i \leq 1$, such that

(1) the dataset S satisfies SARP(e), and
(2) $\sum_{i=1}^{n}(1-e_i)^{\rho} \leq Z$.

COROLLARY 4.6. *Computing Varian's index is NP-hard for objective functions of the form* $\sum_{i=1}^{n}(1-e_i)^{\rho}$*, for any fixed* $\rho \geq 1$.

PROOF. Given an instance of independent set, create an instance of VI-$\{W,S,G,H\}$ARP, as in the proof of their respective theorems, with the following differences. Set $Z := n-k$ and let $0 < \epsilon < 1 - (\frac{n-k}{n-k+1})^{(1/\rho)}$. It can be easily checked that the equivalence holds. □

5. THE INDEX HI

In this section, we consider the problems HI-$\{W,S,G,H\}$ARP. We give the problem HI-SARP, as all other problems are analogous, differing only in the axiom of revealed preference to be satisfied. Notice that in their original paper, Houtman and Maks already showed a relation between HI and a feedback vertex set (see also [Dean and Martin 2010]).

Input. A dataset $S = \{p_i, q_i \,|i = 1, \ldots, n\}$ and a number Z.
Question. Do there exist n numbers e_i, with $e_i \in \{0, 1\}$, such that

(1) the dataset S satisfies SARP(e), and
(2) $\sum_{i=1}^{n} e_i \geq Z$.

THEOREM 5.1. *HI-*$\{W,S,G,H\}$*ARP*$_d$ *is NP-complete.*

PROOF. The proofs of NP-completeness for the problems VI-$\{W,S,G,H\}$ARP$_d$ are easily extended to HI. As the choice of e_i is now limited to either zero or one, it is clear that every large $e_i = 1$ and every other $e_j = 0$. □

The NP-hardness of the optimization problems follows from the NP-completeness of the decision problems.

THEOREM 5.2. *For each fixed* $\delta > 0$*, the existence of a polynomial-time approximation algorithm HI* $-\{W,S,G,H\}$ *achieving a ratio of* $n^{1-\delta}$ *implies P = NP.*

PROOF. Consider an instance of independent set and the corresponding instance of HI-SARP as constructed in Theorem 4.1. Clearly, if the optimum value of the HI-SARP instance equals z, then the optimum value for the IS instance also equals z (if not, then there exists an independent set of size $z+1$, and by the previous reduction, we can find e so that $\sum_{i=1}^{n} e_i \geq z+1$). Now assume that there exists a polynomial-time approximation algorithm achieving a ratio of $n^{1-\delta}$ for HI-SARP, then we can find a vector index so that $\sum_{i=1}^{n} e_i \geq \frac{z}{n^{1-\delta}}$. Given this vector index, we can find an independent set of size $\lceil \frac{z}{n^{1-\delta}} \rceil$ as follows: for every i for which $e_i = 1$, add vertex i to the independent set. This would give us an $n^{1-\delta}$-approximation for IS in polynomial time. Relying on the same results as in Theorem 4.5, this provides the bound of Houtman and Maks' index. □

6. AFRIAT'S INDEX (AI)

6.1. Introductory Observations

As with the previous indices, it is our goal to find the maximum value of e ($e_1 = e_2 = \ldots = e_T = e$), such that a given dataset still passes $\{W, S, G, H\}$ARP. However, such a maximum-value frequently does not exist. For example, consider the following matrix of the values $p_i q_j$ (for two observations),

$$\begin{pmatrix} 1 & 0.50 \\ 0.60 & 1 \end{pmatrix}.$$

As long as $e \in [0; 0.6[$, all axioms of revealed preference will be satisfied, but for $e \geq 0.6$, a cycle of negative length between the two vertices exists in both $G_e(S)$ and $G'_e(S)$, and thus the axioms are violated. Since there is no maximum feasible value for e, we look for the value e^* that is the supremum of the values of e for which the axioms of revealed preference are satisfied. Varian [1990] describes an approximation algorithm which approximates e^* to within $(1/2)^t$ by testing the axiom under e t times. In an overview paper, Varian [2006] mentions that it is also easy to calculate e^* exactly, and exact values are calculated for AI-GARP in a number of papers (see, for instance, Choi et al. [2007]). However, to the best of our knowledge, no exact polynomial algorithm has been described in the literature. In the next section, we provide such a polynomial-time exact algorithm for AI-$\{W, S, G\}$ARP and a separate algorithm for AI-HARP.

6.2. Complexity Results

Consider the following algorithm for AI-WARP.

ALGORITHM 1: AI-WARP, (input: p_i, q_i; output e^*)

1: Initialization: Construct an array A of all values $p_i q_j \leq 1, i \neq j$. Sort these values in ascending order, and let x be the median value in A.
2: Test WARP(x). If WARP(x) is satisfied, remove all values lower than or equal to x from A, otherwise remove all higher values.
3: If more than one element remains in the array, repeat step 2; otherwise x is the remaining value in A. Set $e^* = x$

THEOREM 6.1. *Algorithm 1 solves AI-WARP in $O(n^2 \log(n))$ time.*

PROOF. We first argue that Algorithm 1 is correct. Clearly, if the dataset satisfies WARP(e), then it satisfies WARP(e') for all $e' \leq e$. Moreover, the dataset satisfies WARP(0). Thus, for an increasing e, WARP(e) becomes infeasible at some value e^*. This can only happen when an arc, completing a cycle consisting of two arcs, is added to the graph $G_e(S)$, that is, at some value $p_i q_j$. It follows that Algorithm 1 is correct.

Next, we analyze the complexity of this algorithm. To construct A, $p_i q_j$ must be calculated for all pairs of observations, which takes $O(n^2)$ time. In the worst case, this array is of size $O(n^2)$, so sorting is done in $O(n^2 \log(n))$. In the second step of the algorithm, WARP(e) is tested for different values of e. As the array is halved in each iteration, at most $O(\log(n^2))$ such tests are needed, and each such test can be done in $O(n^2)$ by checking each pair of nodes for violations of WARP(e). This gives a total time complexity for the second step of $O(n^2 \log(n))$. The total time complexity is thus determined by the sorting of the array, and the second step and is $O(n^2 \log(n))$. □

THEOREM 6.2. *AI-SARP can be solved in $O(n^2 \log(n))$.*

PROOF. For AI-SARP, we consider Algorithm 1, with the adjustment that SARP(e) is tested instead of WARP(e). This will give us the lowest value for which SARP(e) is violated and thus a solution to AI-SARP.

The time complexity of testing SARP(e) is also $O(n^2)$ time. This can be achieved by testing for acyclicity using, for example, the topological ordering algorithm [Ahuja et al. 1993]. The time bound follows. □

ALGORITHM 2: AI-GARP, (input: p_i, q_i; output e^*)

1: Initialization: Construct an array A of all values $p_i q_j \leq 1, i \neq j$. Sort these values in ascending order, and let x be the median value in A.
2: Test GARP(x). If GARP(x) is satisfied, remove all values lower than x from A, otherwise remove all higher values.
3: If more than two element remains in array A, repeat step 2.
4: Let the lowest of the two remaining values be x_1 and the highest x_2. Test GARP(e) for $e = x_1$, and for $e = x_2$. Then
   i. If GARP(x_1) and GARP(x_2) are satisfied, test GARP($x_2 + \epsilon$), (where $\epsilon > 0$ is an arbitrarily small number). If GARP($x_2 + \epsilon$) is satisfied, then $e^* = 1$ otherwise $e^* = x_2$.
   ii. If both are not satisfied, $e^* = x_1$.
   iii. If GARP(x_1) is satisfied and GARP(x_2) is not, test GARP($x_1 + \epsilon$). If GARP($x_1 + \epsilon$) is satisfied, $e^* = x_2$, otherwise $e^* = x_1$.

In the case of AI-SARP and AI-WARP, an e^* value is computed that corresponds to some arc appearing in the graph, at which point a cycle appears. In GARP(e), however, a cycle does not necessarily indicate a violation, since the length of the cycle may be 0. Therefore, we need a subtle change, as can be seen in in Algorithm 2.

THEOREM 6.3. *Algorithm 2 solves AI-GARP in* $O(n^\alpha \log(n))$ *time.*

PROOF. We first notice that value e^* can be feasible for GARP(e), if for that value, a cycle of length 0 exists in the graph $G(S)$. Therefore, we consider Algorithm 2, which does not discard the highest-known feasible value of e.

The time complexity of this algorithm is similar to that for AI-SARP. Throughout most of the algorithm, the only difference is the testing of GARP(e) instead of SARP(e). In the final step, GARP(e) is tested twice, which has no impact on the overall bound. However, testing GARP(e) is done by computing the transitive closure of the graph. By way of matrix multiplication, this takes $O(n^\alpha)$ (see Coppersmith and Winograd [1990]). The overall complexity is thus $O(n^\alpha \log(n))$. □

Finally, we provide a polynomial-time algorithm for AI-HARP. In this algorithm, we need to compute the minimum cycle mean, that is, the cycle with the shortest average arc length.

ALGORITHM 3: AI-HARP, (input: p_i, q_i; output e^*)

1: Initialization: Construct the graph $G'_1(S)$.
2: Calculate the minimum cycle mean (MCM), which is the shortest average length of the arcs in any cycle in the graph $G'_1(S)$.
3: Calculate e^* as follows: $e^* = \exp(MCM)$.

THEOREM 6.4. *Algorithm 3 solves AI-HARP in* $O(n^3)$ *time.*

PROOF. We will show that computing the minimum cycle mean (MCM) of $G'_1(S)$ is sufficient for finding e^*. HARP(e) is satisfied if there are no cycles of negative length in $G'_e(S)$. Thus, if such a cycle exists, we need to remove it by lowering e. A decrease in e will lengthen every arc in the graph by the same amount, as the length of an arc is $\log(p_i q_j) - \log(e)$. It is clear that if we set the value of e^* so that the cycle with the shortest average arc length has a length of zero, the average arc length of every other cycle will be nonnegative, and no cycles of negative length will remain. Indeed, by setting $e^* := \exp(\text{MCM})$, the length of each arc becomes $\log(p_i q_j) - \log(\exp(\text{MCM})) = \log(p_i q_j) - \text{MCM}$.

The time complexity of this algorithm is polynomial, as there exist algorithms for finding the MCM in $O(nm)$ time [Karp 1978], with m being the number of arcs in the graph. In $G'_1(S)$, there will be n^2 arcs, as there exist arcs from every vertex to every other vertex in the graph. The building of the graph in the initialization step takes $O(n^2)$ time. The overall time bound of the algorithm is thus $O(n^2 \log(n) + n^3) = O(n^3)$ time. □

7. CONCLUSION

Motivated by the increasing availability of large-scale consumption datasets and by the observed empirical difficulty of computing goodness-of-fit indices, we have investigated the computational complexity of testing the utility-maximization hypothesis in revealed preference terms. In particular, we have focused on three goodness-of-fit measures for four different revealed-preference axioms (i.e., WARP, SARP, GARP, and HARP). We have demonstrated that, for all four axioms, both Varian's and Houtman and Maks' index are inapproximable. Next, we have shown that these conclusions do not apply to Afriat's index, and we have presented exact polynomial algorithms for computing this index (for every revealed-preference axiom that we considered).

There are different avenues for further research. Clearly, when insisting on optimal solutions for one of the indices considered here (except Afriat), one needs to accept long running times. It is therefore interesting to develop bounds: not only should heuristics be defined and tested, but also the design of good upper bounds for the indices is an interesting (and largely unexplored) issue. Another direction is to consider datasets or utility-maximization models that are in some way restricted. Essentially, this may ensure that the graphs built in our analysis simplify such that computation of the respective indices becomes easier.

ACKNOWLEDGMENTS

We gratefully acknowledge helpful discussions with Fabrice Talla Nobibon and Gerhard Woeginger, as well as the comments of the referees.

REFERENCES

S. N. Afriat. 1973. On a system of inequalities in demand analysis: An extension of the classical method. *Int. Econ. Rev. 14*, 2, 460–472.

R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. 1993. *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall.

J. C. R. Alcantud, D. L. Matos, and C. R. Palmero. 2010. Goodness-of-fit in optimizing a consumer model. *Math. Comput. Model. 52*, 7–8, 1088–1094.

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

J. Apesteguia and M. A. Ballester. 2010. The computational complexity of rationalizing behavior. *J. Math. Econ. 46*, 356–363.

J. Apesteguia and M. Ballester. 2011. A measure of rationality and welfare. Working Paper, Universitat Pompeu Fabra, Departamento de Economía y Empresa, No. 1220.

R. Baron, J. Durieu, H. Haller, and P. Solal. 2004. Finding a Nash equilibrium in spatial games is an NP-complete problem. *Econ. Theory 23*, 445–454.

R. Baron, J. Durieu, H. Haller, R. Savani, and P. Solal. 2008. Good neighbors are hard to find: Computational complexity of network formation. *Rev. Econ. Des. 12*, 1–19.

F. Brandt and F. Fisher. 2008. Computing the minimal covering set. *Math. Social Sci. 56*, 254–268.

F. Brandt, F. Fisher, P. Harrenstein, and M. Mair. 2010. A computational analysis of the tournament equilibrium set. *Social Choice Welfare 34*, 597–609.

D. R. Burghart, P. W. Glimcher, and S. C. Lazzaro. 2013. An expected utility maximizer walks into a bar. *J. Risk Uncertain. 46*, 3, 215–246.

K. Cechlarova and J. Hajdukova. 2002. Computational complexity of stable partitions with B-preferences. *Int. J. Game Theory 31*, 353–364.

L. Cherchye, T. Demuynck, and B. De Rock. 2011. Testable implications of general equilibrium models: An integer programming approach. *J. Math. Econ. 47*, 564–575.

S. Choi, R. Fisman, D. Gale, and S. Kariv. 2007. Consistency and heterogeneity of individual behavior under uncertainty. *Am. Econ. Rev. 97*, 5, 1921–1938.

S. Choi, S. Kariv, W. Müller, and D. Silverman. 2011. Who is (More) rational? Tech. rep., National Bureau of Economic Research.

F. Chu and J. Halpern. 2001. On the NP-completeness of finding an optimal strategy in games with common payoffs. *Int. J. Game Theory 30*, 99–106.

V. Conitzer and T. Sandholm. 2008. New complexity results about Nash equilibria. *Games Econ. Behav. 63*, 621–641.

D. Coppersmith and S. Winograd. 1990. Matrix multiplication via arithmetic progressions. *J. Symbol. Comput. 9*, 3, 251–280.

J. C. Cox. 1997. On testing the utility hypothesis. *Econ. J. 107*, 443, 1054–1078.

M. Dean and D. Martin. 2010. How rational are your choice data? In *Proceedings of the Conference on Revealed Preferences and Partial Identification*.

R. Deb. 2010. An efficient nonparametric test of the collective household model. Working Paper, University of Toronto.

F. Echenique, S. Lee, and M. Shum. 2011. The money pump as a measure of revealed preference violations. *J. Political Econ. 119*, 6, 1201–1223.

Q. Fang, S. Zhu, M. Cai, and X. Deng. 2002. On computational complexity of membership test in flow games and linear production games. *Int. J. Game Theory 31*, 39–45.

P. Février and M. Visser. 2004. A study of consumer behavior using laboratory data. *Exp. Econ. 7*, 1, 93–114.

A. Galambos. 2009. The complexity of Nash rationalizability. Tech. rep., Lawrence University.

I. Gilboa and E. Zemel. 1989. Nash and correlated equilibria: Some complexity considerations. *Games Econ. Behav. 1*, 80–93.

W. T. Harbaugh, K. Krause, and T. R. Berry. 2001. Garp for kids: On the development of rational choice behavior. *Am. Econ. Rev. 91*, 5, 1539–1545.

J. Håstad. 1999. Clique is hard to approximate withinn 1- ε. *Acta Mathematica 182*, 1, 105–142.

M. Houtman and J. Maks. 1985. Determining all maximal data subsets consistent with revealed preference. *Kwantitatieve methoden 19*, 89–104.

O. Hudry. 2009. A survey on the complexity of tournament solutions. *Math. Social Sci. 57*, 292–303.

S. Kalyanaraman and C. Umans. 2008. The complexity of rationalizing matchings. In *Proceedings of the 19th International Symposium on Algorithms and Computation*. Lecture Notes in Computer Science, Springer, Berlin, 171–182.

R. M. Karp. 1972. Reducibility among combinatorial problems. *Complex. Comput. Comput. 40*, 4, 85–103.

R. M. Karp. 1978. A characterization of the minimum cycle mean in a digraph. *Discrete Math. 23*, 3, 309–311.

A. Y. C. Koo. 1963. An empirical test of revealed preference theory. *Econometrica 31*, 4, 646–664.

A. Y. C. Koo. 1971. Revealed preference–A structural analysis. *Econometrica 39*, 1, 89–97.

A. Y. C. Koo and G. Hasenkamp. 1972. Structure of revealed preference: Some preliminary evidence. *J. Political Econ. 80*, 4, 724–744.

S. E. Landsburg. 1981. Taste change in the united kingdom, 1900–1955. *J. Political Econ. 89*, 1, 92–104.

M. E. Manser and R. J. McDonald. 1988. An analysis of substitution bias in measuring inflation, 1959–85. *Econometrica 56*, 4, 909–930.

A. Mattei. 2000. Full-scale real tests of consumer behavior using experimental data. *J. Econ. Behav. Org. 43*, 4, 487–497.

A. Mossin. 1972. A mean demand function and individual demand functions confronted with the weak and the strong axioms of revealed preference: An empirical test. *Econometrica 40*, 1, 177–192.

F. T. Nobibon, L. Cherchye, B. De Rock, J. Sabbe, and F. C. R. Spieksma. 2011. Heuristics for deciding collectively rational consumption behavior. *Comput. Econ. 38*, 173–204.

A. D. Procaccia and J. S. Rosenschein. 2008. On the complexity of achieving proportional representation. *Social Choice Welfare 30*, 353–362.

R. Sippel. 1997. An experiment on the pure theory of consumer's behaviour. *Econ. J. 107*, 1431–1444.

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

H. R. Varian. 1990. Goodness-of-fit in optimizing models. *J. Econometrics 46*, 1–2, 125–140.

H. R. Varian. 1993. Goodness-of-fit for revealed preference tests. Unpublished.

H. R. Varian. 2006. Revealed preference. In *Samuelsonian Economics and the Twenty-First Century*, 99–116.

G. J. Woeginger. 2003. Banks winners in tournaments are difficult to recognize. *Social Choice Welfare 20*, 523–528.

D. Zuckerman. 2006. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*. ACM, 681–690.

Received September 2012; revised January 2013; accepted March 2013