Goodness of fit measures for revealed preference tests: Complexity results and algorithms

Laurens Cherchye, Bram De Rock, Bart Smeulders
and Frits C.R. Spieksma

Abstract

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). 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 constant factor approximations algorithm exists for the Houtman-Maks index (unless P = NP). Finally, we give an exact polynomial time algorithm for finding Afriat's efficiency index.

Keywords: Revealed preference, Complexity, Nonparametric rationality tests

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 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 (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) [26]. Three other axioms that are most frequently considered in the applied 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 exactly 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 amount of violations. To deal with this, a number of measures have been proposed in the literature to express how close a data set 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) [1]. Other frequently used measures are the ones of Houtman and Maks (HI) [15] and Varian (VI) [27]. 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 applied 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 [18, 19], Koo and Hasenkamp [20], Mossin [24] and Landsburg [21]. Varian [26] tested GARP using similar data. Only Koo tried to measure the severity of the rejections by focusing on the amount 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 [25] tests relaxations of WARP, SARP and GARP related to AI. AI and GARP are used in papers by Mattei [23], Harbaugh et al. [14] Andreoni and Miller [4], Février and Visser [13], Choi et al. [7, 8], Dean and Martin [11] and Burghart [6]; the last four papers also use HI. VI and GARP appears in Cox [10], Mattei [23], Choi et al. [7, 8] and Dean and Martin [11]. For WARP, all three indices appear in Choi et al. [7]. 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 for example Manser and McDonald [22]). Finally, we also note 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. [12] which calculates the monetary cost of irrational behaviour and the minimal swaps and minimal loss index by Apesteguia and Ballester [5].

This paper is specifically concerned with the computational complexity of the goodness of fit measures used in revealed preference analysis. In general, computational complexity becomes an important issue if one wants to consider large data sets. In this respect, we indicate that large consumption data sets 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 above mentioned goodness of fit measures.

It is generally thought that calculating AI is easy. However, to our knowledge, no exact algorithm exists in the literature. Varian [27] 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.[1] For instance, Varian [27] writes:

> "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. [8] state:

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

The goal of the current paper 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 feedback vertex set on a digraph, which implies NP-Hardness. Next, Dean and Martin [11] state that HI for GARP is also NP-HARD.

We define 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) mentioned above. 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 data set satisfies a relaxation of GARP.

---

[1]Because of the difficulty to exactly calculate VI, some authors have focused on designing approximate heuristics. See, for example, Varian [28] and Alcantud et al. [3].

Our main results are summarized in Table 1, 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]

	WARP	SARP	GARP	HARP
AI (sec 6)	$n^2 \log n$	$n^2 \log n$	$n^{2.376} \log n$	n^3
VI (sec 4)	NP-HARD	NP-HARD	NP-HARD	NP-HARD
HI (sec 5)	Inapproximable	Inapproximable	Inapproximable	Inapproximable

Table 1: Overview of Results

The rest of this paper unfolds as follows. The next section sets the stage by introducing the basic revealed preference concepts that we will use throughout. Section 3 provides a statement of the computational problems we focus on. Sections 4 and 5 then presents our results on computational complexity for the indices VI and HI. Section 6 does the same for the index AI. Here, we also give exact polynomial time algorithms for computing this index in practical applications. Section 7 concludes.

2 Revealed preference concepts

We start by stating the 4 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 data set $S = \{(p_i, q_i) | i = 1, \ldots, n\}$, 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 then state the four revealed preference axioms that we consider in this paper.

Definition 1. (WARP) *A data set S satisfies WARP if for each pair of bundles, q_i, q_j $(i, j = 1, \ldots, n$ with $i \neq j)$, the following holds: if $q_i R_0 q_j$ then it is not the case that $q_j R_0 q_i$.*

Definition 2. (SARP) *A data set S satisfies SARP if for each pair of bundles, q_i, q_j, $(i, j = 1, \ldots, n$ with $i \neq j)$, the following holds: if $q_i R q_j$ then it is not the case that $q_j R_0 q_i$.*

Definition 3. (GARP): *A data set S satisfies GARP if for each pair of bundles, q_i, q_j, $(i, j = 1, \ldots, n$ with $i \neq j)$, the following holds: if $q_i R q_j$ then it is not the case that $q_j P_0 q_i$.*

Definition 4. (HARP): *A data set S satisfies HARP if for every sequence of observations, $i, j, k, \ldots, l (= 1, \ldots, n)$, the following holds: $\log(p_i q_j) + \log(p_j q_k) + \ldots + \log(p_l q_i) \geq 0$.*

In words, the main differences between the alternative axioms can be summarized as follows (see Varian [29] for a more extensive discussion on the meaning of the axioms). Data consistency with WARP is a necessary condition for data consistency with SARP;

---

[2] *'inapproximable'* stands for: no polynomial-time algorithm can achieve a constant-factor approximation unless $P = NP$.

the essential difference is that WARP (in contrast to SARP) does not require transitivity of preferences. 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 multi-valued demand. As such, GARP is a generalization of SARP. 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.

2.2 Goodness of fit measures

In practice, direct application of any of the above revealed preference axioms to some given data set effectively obtains a 'sharp' test: a data set either satisfies the 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 data set 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 data set is close to consistency (see [27] 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 [29] 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 $0 \le e_i \le 1$. This vector introduces an index e_i for each observation i, which relaxes the revealed preference relations R_0 and P_0 as follows:

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

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

Analogous to before, $R(e)$ and $P(e)$ represent the transtive 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 5. (WARP(e)) *A data set S satisfies WARP(e) if for each pair of bundles, $(i, j = 1, \ldots, n$ with $i \ne j)$, the following holds: if $q_i R_0(e) q_j$ then it is not the case that $q_j R_0(e) q_i$.*

Definition 6. (SARP(e)) *A data set S satisfies SARP(e) if for each pair of bundles, q_i, q_j, $(i, j = 1, \ldots, n$ with $i \ne j)$, the following holds: if $q_i R(e) q_j$ then it is not the case that $q_j R_0(e) q_i$.*

Definition 7. (GARP(e)) *A data set S satisfies GARP(e) if for each pair of bundles, q_i, q_j, $(i, j = 1, \ldots, n$ with $i \neq j)$, the following holds: if $q_i R(e) q_j$ then it is not the case that $q_j P_0(e) q_i$.*

Definition 8. (HARP(e)) *A data set S satisfies HARP(e) if for every sequence of observations $i, j, k, \ldots, l (= 1, \ldots, n)$,, the following holds: $\log(p_i q_j) + \log(p_j q_k) + \ldots + \log(p_l q_i) \geq \log(e_i) + \log(e_j) + \ldots + \log(e_l)$.*

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. The index AI equals the highest value 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 note 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 the index AI by allowing for observation specific perturbations. The index 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, the index VI is subject to the same qualifications as the index 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 1. 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 data set actually satisfies some revealed preference axiom, it is natural to construct a graph (see Koo [19]). We now extend this procedure by taking into account a given vector $e = (e_1, \ldots, e_n)$. For some data set 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 the 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 9. (WARP(e)) *The data set S satisfies WARP(e) if and only if the graph $G_e(S)$ does not contain any cycle consisting of two arcs.*

Definition 10. (SARP(e)) *The data set S satisfies SARP(e) if and only if the graph $G_e(S)$ is acyclic.*

Definition 11. *(GARP(e))* *The data set S satisfies GARP(e) if and only if the graph $G_e(S)$ does not contain any cycles of negative length.*

Definition 12. *(HARP(e))* *The data set S satisfies HARP(e) if and only if the 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 data set 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 1. *(VI-SARP) Given a data set 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, other objective functions are possible, We will give results and come back to this issue in Section 4.

Problem 2. *(HI-SARP) Given a data set 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. *(AI-SARP) Given a data set S, for what value e_1, with $0 \leq e_1 \leq 1$, is e_1 maximized while S satisfies SARP(e), with $e = (e_1, \ldots, e_1)$?.*

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 $(\sum_{i=1}^{n}(1 - e_i)^\rho)^{1/\rho}$ for $\rho \geq 1$. Observe that, since $\lim_{\rho \to \infty}(\sum_{i=1}^{n}(1 - e_i)^\rho)^{1/\rho} = max_i (1 - e_i)$, the Afriat index arises when $\rho \to \infty$. The results in this section are phrased for $\rho = 1$, i.e., 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 $\rho \geq 1$.

Let us now consider the following decision problem associated with VI-SARP:

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

(*i*) The data set S satisfies SARP(e), and

(*ii*) $\sum_{i=1}^{n} e_i \geq Z$?

Theorem 1. *VI-SARP is NP-Hard.*

Proof. We prove that VI-SARP is NP-Hard by a reduction from the well-known NP-Hard independent set problem [16], 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. For every node $i \in V$, there is an observation in VI-SARP: $n := |V|$. 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, i.e., 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. 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. Suppose the instance of independent set is a yes-instance, i.e., 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 the 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 note 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 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. $\qquad\square$

We now proceed with the closely related problems VI-GARP, VI-WARP and VI-HARP:

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

(i) The data set S satisfies the appropriate axiom, GARP(e), WARP(e) or HARP(e), and

(ii) $\sum_{i=1}^{n} e_i \geq Z$?

Theorem 2. *VI-GARP is NP-Hard.*

Theorem 3. *VI-WARP is NP-Hard.*

Theorem 4. *VI-HARP is NP-Hard.*

Instances of these problems are built as in the proof of Theorem 1. The proofs of equivalence are relatively straightforward.

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

   (i) The data set S satisfies SARP(e), and

 (ii) $\sum_{i=1}^{n} (1 - e_i)^\rho \leq Z$?

Corollary 1. *Varian's Index is NP-HARD for objective functions of the form minimize* $(\sum_{i=1}^{n} (1 - e_i)^\rho)^{1/\rho}$, *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 equivalences hold. $\qquad \square$

5   The index HI

In this section, we consider the problems HI-$\{W, S, G, H\}$ARP. We give the problem HI-SARP, 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 feedback vertex set, see also [11]

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

   (i) The data set S satisfies SARP(e), and

 (ii) $\sum_{i=1}^{n} e_i \geq Z$?

Theorem 5. *HI-$\{W, S, G, H\}$ARP is NP-Hard.*

Proof. The proof of NP-hardness for maximizing the sum of the elements of VI is 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 $e_j = 0$. $\qquad \square$

Theorem 6. *No polynomial time ρ-approximation algorithm exists for HI-$\{W, S, G, H\}$ARP, unless $P = NP$.*

Proof. Consider an instance of Independent Set, and the corresponding instance of HI-SARP as constructed in Theorem 1. Now consider that the optimum of the HI-SARP instance is z, then the optimum for IS is also z. If not, then for we could find an independent set of size $z + 1$ and by the previous reduction we could find e so that $\sum e_i \geq z + 1$.

Now consider we have a ρ-approximation for HI-SARP, then we could find a vector-index so that $\sum e_i \geq z \times \rho$ in polynomial time. Given this vector-index we could find an independent set of size $z \times \rho$ as follows, for every i for which $e_i = 1$ add the vertex i to the independent set. This would give us a ρ-approximation for IS in polynomial time. This implies that $P = NP$. $\qquad \square$

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 data set 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 [27] 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 [29] 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. [7]. However, to the best of our knowledge, no exact polynomial algorithm has been published 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

Theorem 7. *AI-WARP can be solved in $O(n^2 \log(n))$.*

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 $G_e(S)$, i.e., at some value $p_i q_j$. It follows that Algorithm 1 is correct.

Next we analyse 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 $O(n^2 \log(n))$. □

Algorithm 1 AI-WARP

1: Initialization: Construct an array A of all values $p_i q_j \leq 1, i \neq j$. Sort these values in ascending order.
2: Let x be the median value in A. Test WARP(x), if WARP(e) 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 let x be the remaining value in A, then $e^* = x$

Theorem 8. *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). SARP(e) can also be tested in $O(n^2)$, for example by a topological ordering algorithm [2], leading to the same time complexity. □

Theorem 9. *AI-GARP can be solved in $O(n^{2.376} \log(n))$.*

Proof. We first note that the 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 a variant of the algorithm, which does not discard the highest known feasible value of e.

The time complexity of this algorithm is similar to that for AI-WARP and AI-SARP, but differs in that testing GARP(e) takes $O(n^{2.376})$. This test is done by finding the transitive closure by way of matrix multiplication [9]. □

Finally, we provide a polynomial time algorithm for AI-HARP.

Algorithm 2 AI-HARP

1: Input: A set of observations $p_t = (p_t^1, \ldots, p_t^N), q_t = (q_t^1, \ldots, q_t^N)$ for $t = 1, \ldots, T$
2: Initialization: Construct the graph $G_1'(S)$
3: Calculate the minimum cycle mean (MCM), which is the shortest average length of the arcs in any cycle in the graph $G_1'(S)$.
4: Calculate e^* as follows: $e^* = \exp(MCM)$.

Theorem 10. *AI-HARP can be solved in time proportional to $O(n^3)$.*

Proof. We will show that computing the minimum cycle mean (MCM) of $G_1'(S)$ is sufficient to find 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 the 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 non-negative 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 [17], with m being the number of arcs in the graph. In $G_1'(S)$ there will be n^2 arcs. The building of the graph 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 data sets, 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, the Houtman and Maks index is inapproximable and that computing Varian's index is NP-Hard. 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).

References

[1] S.N. Afriat. On a system of inequalities in demand analysis: an extension of the classical method. *International Economic Review*, 14(2):460–472, 1973.

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

[3] J.C.R. Alcantud, D.L. Matos, and C.R. Palmero. Goodness-of-fit in optimizing a consumer model. *Mathematical and Computer Modelling*, 52(7-8):1088–1094, 2010.

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

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

[6] D.R. Burghart, P.W. Glimcher, and S.C. Lazzaro. An expected utility maximizer walks into a bar... Working Paper, 2011.

[7] S. Choi, R. Fisman, D. Gale, and S. Kariv. Consistency and heterogeneity of individual behavior under uncertainty. *The American economic review*, 97(5):1921–1938, 2007.

[8] S. Choi, S. Kariv, W. Müller, and D. Silverman. Who is (more) rational? Technical report, National Bureau of Economic Research, 2011.

[9] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *Journal of symbolic computation*, 9(3):251–280, 1990.

[10] J.C. Cox. On testing the utility hypothesis. *The Economic Journal*, 107(443):1054–1078, 1997.

[11] M. Dean and D. Martin. How rational are your choice data? In *Conference on Revealed Preferences and Partial Identification*, 2010.

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

[13] P. Février and M. Visser. A study of consumer behavior using laboratory data. *Experimental economics*, 7(1):93–114, 2004.

[14] W.T. Harbaugh, K. Krause, and T.R. Berry. Garp for kids: On the development of rational choice behavior. *The American economic review*, 91(5):1539–1545, 2001.

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

[16] R.M. Karp. Reducibility among combinatorial problems. *Complexity of Computer Computations*, 40(4):85–103, 1972.

[17] R.M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete mathematics*, 23(3):309–311, 1978.

[18] A.Y.C. Koo. An empirical test of revealed preference theory. *Econometrica*, 31(4):646–664, 1963.

[19] A.Y.C. Koo. Revealed preference–a structural analysis. *Econometrica*, 39(1):89–97, 1971.

[20] A.Y.C. Koo and G. Hasenkamp. Structure of revealed preference: some preliminary evidence. *The Journal of Political Economy*, 80(4):724–744, 1972.

[21] S.E. Landsburg. Taste change in the united kingdom, 1900-1955. *The Journal of Political Economy*, 89(1):92–104, 1981.

[22] M.E. Manser and R.J. McDonald. An analysis of substitution bias in measuring inflation, 1959-85. *Econometrica*, 56(4):909–930, 1988.

[23] A. Mattei. Full-scale real tests of consumer behavior using experimental data. *Journal of Economic Behavior & Organization*, 43(4):487–497, 2000.

[24] A. Mossin. 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, 1972.

[25] R. Sippel. An experiment on the pure theory of consumer's behaviour. *The Economic Journal*, 107:1431–1444, 1997.

[26] H.R. Varian. The nonparametric approach to demand analysis. *Econometrica*, 50(4):945–973, 1982.

[27] H.R. Varian. Goodness-of-fit in optimizing models. *Journal of Econometrics*, 46(1-2):125–140, 1990.

[28] H.R. Varian. Goodness-of-fit for revealed preference tests. Unpublished, 1993.

[29] H.R. Varian. Revealed preference. *Samuelsonian economics and the twenty-first century*, pages 99–116, 2006.

Laurens Cherchye
Faculty of Business and Economics - Center for Economic Studies
KU Leuven
Naamsestraat 69
8500 Kortrijk, Belgium
Email: `laurens.cherchye@econ.kuleuven.be`

Bram De Rock
ECARES
Université libre de Bruxelles
Av. F.D. Roosevelt 50
1050 Brussels, Belgium
Email: `bderock@ulb.ac.be`

Bart Smeulders
Faculty of Business and Economics - ORSTAT
KU Leuven
Naamsestraat 69
3000 Leuven, Belgium
Email: `bart.smeulders@econ.kuleuven.be`

Frits C.R. Spieksma
Faculty of Business and Economics - ORSTAT
KU Leuven
Naamsestraat 69
3000 Leuven, Belgium
Email: `frits.spieksma@econ.kuleuven.be`