ARTICLE IN PRESS
JID: EOR [m5G;May 4, 2018;11:9]

European Journal of Operational Research 000 (2018) 1–13

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Invited Review

# Revealed preference theory: An algorithmic outlook

Bart Smeulders[a], Yves Crama[a,*], Frits C. R. Spieksma[b]

[a] HEC Management School, University of Liège, Liège 4000, Belgium
[b] Department of Mathematics and Computer Science, TU Eindhoven, Eindhoven 5600 MB, The Netherlands

ARTICLE INFO

*Article history:*
Received 31 August 2017
Accepted 11 April 2018
Available online xxx

*Keywords:*
Preference learning
Revealed preference
Rationality axioms
Utility theory
Computational complexity

ABSTRACT

Revealed preference theory is a domain within economics that studies rationalizability of behavior by (certain types of) utility functions. Given observed behavior in the form of choice data, testing whether certain conditions are satisfied gives rise to a variety of computational problems that can be analyzed using operations research techniques. In this survey, we provide an overview of these problems, their theoretical complexity, and available algorithms for tackling them. We focus on consumer choice settings, in particular individual choice, collective choice and stochastic choice settings.



## 1. Introduction

### 1.1. Motivation

Our world is full of choices. Before we step outside the door in the morning, we have already chosen what to eat for breakfast and which clothes to wear. For the morning commute, we decide how to travel, by what route, and whether we will pick up coffee along the way. Dozens of small choices are made before it is even time for lunch, and then there are the less frequent, but more important decisions like buying a car, moving to a new home, or setting up retirement savings. Neoclassical economists hypothesize that such consumption choices are made so as to maximize utility. Given this hypothesis, it follows that each choice tells us something about the decision maker. In other words, choices *reveal preferences*, and thereby provide information about an underlying utility function. As we observe the choices of a decision maker over time, we can piece together more and more information. Given this information about choices made, a number of questions naturally arise:

i) Does there exist a utility function which is consistent with the observed choices?
ii) When a consistent utility function exists, does there exist one in a prespecified class?
iii) When no consistent utility function exists, how close are the observed choices to being consistent?

These questions belong to the domain of *revealed preference theory*, pioneered by Samuelson (1938, 1948). In this theory, it is usual to formulate a minimum set of prior assumptions, also known as *axioms*, which are based on a theory of choice behavior. Thus, revealed preference characterizations are defined as conditions on the observed choices of decision makers. This approach allows for direct tests of the decision models, without running the risk that excessively strong functional (mis)specifications lead to rejections of the model.

Testing the axioms of revealed preference theory is a topic at the interface of economics and operations research. We focus on the algorithmic aspects of solving the corresponding optimization/decision problems, and we highlight some of the issues of interest from the operations research viewpoint. In particular, we examine algorithms that can be used to test whether observed consumer choices satisfy certain revealed preference conditions. We also look at the tractability, that is, the computational complexity of algorithms for answering these questions. Following the classical framework of computational complexity (see, for instance, Garey & Johnson, 1979 or Cormen, Leiserson, Rivest, & Stein, 2001), we focus on worst-case time-bounds of algorithms. We are especially interested in whether a particular question is easy (that is, solvable in polynomial time) or difficult (NP-HARD), and what the best-known method is for answering the question.

Let us first motivate this computational point of view. In a very general way, it is clear that computational issues have become increasingly important in all aspects of science, and economics is no exception. This is reflected, in particular, in the economic literature on revealed preference, where computational challenges are frequently and explicitly mentioned. We illustrate this claim with three quotes from recent papers.

* Corresponding author.
*E-mail addresses:* bart.smeulders@uliege.be (B. Smeulders), Yves.Crama@uliege.be (Y. Crama), f.c.r.spieksma@tue.nl (F.C.R. Spieksma).

https://doi.org/10.1016/j.ejor.2018.04.026
0377-2217/

Please cite this article as: B. Smeulders et al., Revealed preference theory: An algorithmic outlook, European Journal of Operational Research (2018), https://doi.org/10.1016/j.ejor.2018.04.026

Echenique, Lee, and Shum (2011) write:

> "Given [that calculating money pump costs can be a huge computational task], we check only for violations of GARP that involve cycles of limited length: lengths 2, 3, and 4."

Choi, Kariv, Müller, and Silverman (2014) write (in the online appendix):

> "Since the algorithm is computationally very intensive, for a small number of subjects we report upper bounds on the consistent set."

Kitamura and Stoye (2014) write:

> "It is computationally prohibitive to test stochastic rationality on 25 periods at once. We work with all possible sets of eight consecutive periods, a problem size that can be very comfortably computed."

These quotes signify the need for fast algorithms that can test rationality of choices made by an individual (or a group of individuals), or at least to better understand the tractability of these underlying questions.

Another trend that emphasizes the relevance of efficient computations in the domain of revealed preference is the ever-increasing size of datasets. As in many other fields of social and exact sciences, and as underlined by the pervasiveness of buzzwords such as "big data" and "data science", more and more information is available about actual choices of decision makers. As a striking example, it is now commonplace for brands or large retailers to track the purchases of individual consumers or households. This activity yields numerous datasets with sizes far beyond those provided by laboratory experiments. This only reinforces the need for efficient methods, in order to be able to tackle and to draw meaningful conclusions from huge datasets. For example, Cherchye et al. (2017a) use revealed preference models to study food choices. The sample they analyze contains records of all grocery purchases of 3645 individuals over a period of 24 months. It is extracted from the Kantar Worldpanel, which records the purchases of 25,000 households. Long-running longitudinal studies actually provide large datasets of household consumption and other economic indicators. Cherchye, Demuynck, De Rock, and Vermeulen (2017b) identify intrahousehold decision structures using such large datasets.

In view of these considerations, there is a quickly growing body of work on computation and economics. As mentioned above, our objective is to give an overview of algorithmic problems arising in revealed preference theory. Due to the wide range of choice situations to which revealed preference has been applied, providing a comprehensive overview is not a realistic goal. In this paper, we focus on algorithmic results concerning tests of rational behavior in consumer choice settings. For different discussions of the topic, we refer the reader to the recent monograph on the theory of revealed preference by Chambers and Echenique (2016), and to a survey by Crawford and De Rock (2014) on empirical revealed preference; an earlier overview can be found in Houtman (1995). Finally, we should note that certain aspects of revealed preference theory, as a way of explaining choice behavior, have also been criticized; see, e.g., the works of Hausman (2000) and Wong (2006).

### 1.2. Preference modeling and utility theory

Before we close this introductory section, we find it useful to formulate a few comments on the relations between the stream of literature that we cover in this paper, and the literature on preference modeling and utility-based decision making, as they have classically been handled in operations research (OR) and, more recently, in artificial intelligence (AI). Our goal is obviously not to survey these huge and active fields of research. Rather, we simply intend to clarify some of the similarities and differences that exist between the "economic" setting of revealed preference theory, and an "operations research" or "artificial intelligence" perspective which may be more familiar to readers of this journal.

Many of the results surveyed in this paper express conditions for the existence of a utility function which represents the preferences revealed through the choices made by consumers. Most of these results have been published by economists. On the other hand, in operations research and in decision theory, there is a long tradition of building utility functions (sometimes called "value functions" in the deterministic setting) based on information provided by one or several decision makers. Classical references are, for instance, Fishburn (1970), Keeney and Raiffa (1976). Typically, in such settings, the preferences of the decision maker are expressed by a limited number of pairwise comparisons of alternatives, or by rankings of the alternatives on several criteria. The objective is then to build a utility function which is coherent with the expressed preferences, and which can be used, for instance, in order to evaluate each and every alternative on a numerical or ordinal scale, or to evaluate alternatives that have not yet been seen. The utility functions under consideration may be as simple as a (weighted) sum of criteria, or may be selected within a parameterized class of functions whose parameters are to be determined. This type of approach has been extensively investigated, in particular, by researchers interested in multiple criteria problems with discrete alternatives (MCDA) (see, e.g., Greco, Ehrgott, & Figueira, 2016, and in particular Bouyssou & Pirlot, 2016; Dyer, 2016; Moretti, Öztürk, & Tsoukiàs, 2016; Siskos, Grigoroudis, & Matsatsinis, 2016 for recent surveys of closely related topics; see also Corrente, Greco, Matarazzo, & Słowiński, 2016 for extensions), or in conjoint analysis (see, e.g., Giesen, Mueller, Taneva, & Zolliker, 2010; Gustafsson, Herrmann, & Huber, 2007; Rao, 2014). More recently, similar questions have also been investigated in preference learning, a subfield of artificial intelligence (see, e.g., Corrente, Greco, Kadziński, & Słowiński, 2013; Fürnkranz & Hüllermeier, 2010).

Not surprisingly, all of these fields share a common theoretical basis, as well as many methodological concepts: preference relations, transitivity, pairwise comparisons, to name but a few. Nevertheless, they also all have their own specific purposes, assumptions, and applications, which lead to a variety of research questions and results. The objective of this survey is not to carry out a systematic comparison of these various settings. However, in order to avoid any confusion in the mind of the reader, we find it useful to briefly outline some of the most striking differences between revealed preference theory and other utility-based frameworks.

- The approaches proposed in OR and in AI are mostly *prescriptive* or *operational* in nature. Their main objective is to help an individual, or a group of individuals, to express and to structure their preferences, *so as to allow them to make informed decisions*. This is the case in MCDA, in conjoint analysis, and in preference learning. In contrast, the revealed preference literature is mostly *normative* (to the extent that it posits axioms of rational choice behavior) and *descriptive* (to the extent that it attempts *to test whether actual consumer choices are consistent* with the stated axioms), but *it is not meant to support any decision making process*. This is definitely a major distinguishing feature of revealed preference theory.
- As a corollary of the previous item, an objective frequently pursued in OR and in AI is *to explicitly build* ("assess", "elicit") a utility function which is compatible with the data; this is the case in multiattribute utility theory or in conjoint analysis, most noticeably. (Of course, some classical approaches to multicriteria decision making do not explicitly attempt to

build the utility function of the decision maker; this is the case, for instance, of the interactive methods developed by Zionts and Wallenius (1976, 1983), and of outranking methods such as described by Roy (1991).) On the other hand, in the economic literature, a main objective is *to check the coherence* of consumer choices with rationality axioms proposed in the theory. Hence, *building* a compatible utility function (sometimes called the “recovery” issue in economics) is usually not viewed as the primary outcome of the process. It should be noted, however, that the existence proofs provided for instance by Afriat (1967b) or Varian (1982) (see Section 3 hereunder) are constructive and provide an analytical expression of the utility function, when it exists. Predicting, or bounding the demand bundles associated with future prices is also a topic in interest in economics; see, e.g., (Blundell, 2005; Varian, 1982).

- In utility theory and in MCDA, the alternatives are often considered as “abstract”, “unspecified” entities: most papers in this stream start with the assumption that the decision maker is facing “a set $A$ of alternatives”, or potential actions, but *the nature of these alternatives is not directly relevant* for the development of the theoretical framework (although, of course, the alternatives must be fully determined in any specific application of the theory); see (Dyer, 2016; Fishburn, 1970; Keeney & Raiffa, 1976). In conjoint analysis or in preference learning, the alternatives are represented as multidimensional vectors associated with product attributes or other measurable features. In revealed preference theory, on the other hand, the observations consist of *bundles of goods* and their *associated prices*: this assumption is crucial for the definition of the preference relation, as we explain next.
- In OR or AI, preferences among alternatives can be formulated in a variety of ways (e.g., through pairwise comparisons of alternatives), but are solely based on declarations of the decision maker. In revealed preference settings, on the contrary, the preferences between bundles are explicitly derived by the analyst from pairwise comparisons of the *prices of the bundles* purchased by the decision maker. As a consequence, goods and their prices play a central role and provide another distinguishing feature of the theory. In particular, many of the theorems regarding the existence of utility functions can be stated in terms of prices and quantities of goods.
- In MCDA, in conjoint analysis, or in preference learning, the procedure used to elicit the utility function often rests on the formulation of questions that can be submitted to the decision maker, possibly in an interactive, dynamic process; so, the *design of the most appropriate experiments* is an important issue to be tackled by the analyst, as it influences the relevance of the collected data and the efficiency of the elicitation process (see, e.g., Gustafsson et al., 2007; Rao, 2014; Riabacke, Danielson, & Ekenberg, 2012 for a discussion of such design issues). In revealed preference settings, the analyst usually faces the results of *uncontrolled experiments*, in the form of a database of observations which have been typically collected for other purposes (although the issue of experimental design is also discussed, for instance, in Blundell, 2005).
- As a consequence of the previous point, the datasets considered in the OR literature on preference modeling are often quite small, and computational complexity, or even algorithmic considerations have not been a main focus of attention in this area. (This is true, at least, for multiple criteria problems with discrete alternatives, as opposed to multiple criteria optimization problems which may feature an infinite set of feasible alternatives, such as a polyhedron described by linear inequalities, and which call for more efficient algorithmic approaches; see, e.g., (Wallenius et al., 2008) for a discussion of the growing importance of algorithmic issues in multicriteria decision making.) On the other hand, the databases to be handled in revealed preference studies are potentially huge, so that complexity issues naturally arise and have been considered, more or less explicitly, by various researchers. They provide the main theme to be covered in this paper.

As previously mentioned, in spite of the inherent differences outlined above, and in spite of the fact that the streams of research on utility-based decision making and on revealed preference have evolved in almost total separation, there remain some obvious commonalities between these topics. The objective of our survey, however, is not to establish a comparative study, but rather to provide the reader with an overview of fundamental results and of recent developments in the field of algorithmic revealed preference theory. We hope that this may lay the ground for future cross-fertilization between operations research and revealed preference theory.

### *1.3. Outline of the survey*

We begin this survey by introducing key concepts in revealed preference theory, such as utility functions and preference relations, in Section 2. Next, in Section 3, we state the fundamental theorems that characterize rationalizability in revealed preference theory. We explicitly connect rationalizability with properties of certain graphs, and we state the worst-case complexity of algorithms that establish whether a given dataset satisfies a particular “axiom” of revealed preference. In Section 4, we look at various kinds of utility functions that have been considered in the literature, and we provide corresponding rationalizability theorems. Section 5 deals with goodness-of-fit and power measures, which respectively quantify the severity of violations and give a measure of how stringent the tests are. In Section 6, we explore collective settings, where the observed choices are the result of joint decisions by several individuals. Finally, in Section 7, we look at stochastic preference settings where the decision maker still attempts to maximize her utility, but her preferences are not necessarily constant over time. Instead, the decision maker has a number of different utility functions, and the function that she maximizes at any given time is probabilistically determined. We conclude in Section 8.

## 2. Preliminaries

In this section, we lay the groundwork for the remainder of this paper: Section 2.1 introduces utility functions and their properties, Section 2.2 states the different axioms of revealed preference, and Section 2.3 shows how graphs can be built from a given set of observations.

### *2.1. Basic properties of utility functions*

Let us first introduce the basic ideas of revealed preference, by considering purchasing decisions and utility maximization. Specifically, consider a world with $m$ different goods. The decision maker selects a *bundle* of goods, denoted by the $(m \times 1)$ vector $q \in \mathbb{R}^m_+$. Throughout this paper, except where noted otherwise, we assume this choice is constrained by a *linear* budget constraint. The $(1 \times m)$ vector $p \in \mathbb{R}^m_{++}$ denotes the *prices* of the goods, and $b$ the available budget. Under the classical hypothesis of utility maximization, the choice of the decision maker is guided by a utility function $u(q): \mathbb{R}^m_+ \to \mathbb{R}_+$. Thus, the decision maker selects (consciously or not) an optimal bundle $q$ by solving the following problem, for any given price vector $p$ and budget $b$.

$$\text{Maximize } u(q) \tag{1}$$

$$\text{subject to } pq \leq b. \tag{2}$$

Following standard economic theory, we assume the utility function to be concave, continuous and strictly monotone, a set of properties we capture in the following definition:

**Definition 1 Well-behaved utility function.** A utility function $u(q): \mathbb{R}^m_+ \to \mathbb{R}_+$ is *well-behaved* if and only if $u$ is concave, continuous, and strictly monotone.

Notice that in this survey, we restrict ourselves exclusively to the deterministic setting where the utility function does not depend on unobservable, random elements beyond the bundle $q$.

Another relevant property of a utility function is the potential uniqueness of its optima. This is formulated as follows:

**Definition 2 Single-valued utility function.** A utility function $u(q): \mathbb{R}^m_+ \to \mathbb{R}_+$ is *single-valued* if and only if, for each $p$, $b$, the problem $\{\text{Maximize}\, u(q)\, \text{subject to}\, pq \le b\}$ has a unique optimal solution $q$.

Of course, there are many other properties that one may want to require from a utility function; we come back to this issue in Section 4.

### 2.2. Preference relations and axioms of revealed preference

In the remainder of the paper, we assume that data is collected by observing, at $n$ different points in time, the prices and quantities of all goods that are bought. This yields a dataset $S = \{(p_i, q_i)|\ i \in N\}$, where $p_i \in \mathbb{R}^m_{++}$ is the vector of prices at time $i$, $q_i \in \mathbb{R}^m_+$ is the bundle purchased at time $i$, and $N = \{1, 2, \dots, n\}$. We use the word *observation* to denote a pair $(p_i, q_i)$, $i \in N$.

Samuelson (1938) introduced the definition of the *direct revealed preference relation* over the set of bundles.

**Definition 3 Direct revealed preference relation.** For any pair of observations $i, j \in N$, if $p_i q_i \ge p_i q_j$, we say that $q_i$ is *directly revealed preferred over* $q_j$, and we write $q_i\, R_0\, q_j$.

The interpretation of Definition 3 is quite intuitive: indeed, note that $p_i q_i$ and $p_i q_j$ respectively express the total price of bundle $q_i$ and bundle $q_j$ at time $i$, that is, when the prices $p_i$ apply. If the inequality $p_i q_i \ge p_i q_j$ holds, we thus observe that bundle $q_i$ was purchased at time $i$ *in spite of the fact that* $q_i$ *was at least as expensive as* $q_j$ at time $i$. The natural conclusion is that the decision maker prefers bundle $q_i$ over $q_j$ (otherwise, she would have bought $q_j$), and this is the meaning of the relation $R_0$.

Assume now that we wish to test the hypothesis of utility maximization. In the empirical setting, the budget available to the decision maker at time $i \in N$ is generally unobservable, but it is natural to assume that it is equal to $p_i q_i$. (As a matter of fact, if the decision maker maximizes her utility and if the utility function is monotonic, then the bundle picked at each period must exhaust the available budget, which is therefore equal to $p_i q_i$ at time $i$.)

We now wish to test whether the given dataset is consistent with the theory of utility maximization. For the data to be consistent with that theory, there must exist a utility function such that all purchasing decisions maximize utility under the budget constraints. We say that a utility function satisfying this requirement *rationalizes* the data, and we call it a *rationalizing utility function*.

**Definition 4. Rationalizability**

A dataset $S = \{(p_i, q_i)|\ i \in N\}$ is *rationalizable by a well-behaved (single-valued) utility function* if and only if there exists a well-behaved (single-valued) utility function $u$ such that for every observation $i \in N$,

$$u(q_i) \ge u(q_j) \text{ for all } j \in N \text{ with } p_i q_i \ge p_i q_j.$$

This rationalizability concept is key in revealed preference theory, and goes back to the work of Antonelli (1886). In words, Definition 4 expresses that, at each time $i \in N$, the choice of the decision maker was rational in the sense that she picked the bundle which maximizes her utility among all (observed) bundles $q_j$, $j \in N$, whose total price $p_i q_j$ (at time $i$) was within the budget $p_i q_i$. Restricting the attention to the finite set of bundles $\{q_j | j \in N\}$ that actually have been observed in the dataset, rather than considering the infinite universe $\mathbb{R}^m_+$ of all bundles that could potentially be bought by the decision maker, will allow us to test Definition 4 in an empirical setting, as we will find out in the next sections.

In terms of the direct revealed preference relation, the utility function $u(q)$ rationalizes the data if and only if $u(q_i) \ge u(q_j)$ for all $i, j \in N$ such that $q_i\, R_0\, q_j$: in the terminology of Fishburn (1970), this means that $u(q)$ is *order-preserving* for $R_0$; see also Bouyssou and Pirlot (2016). Therefore, it is natural to investigate conditions on $R_0$ which ensure that a data set is rationalizable. This observation led Samuelson (1938) to formulate the *Weak Axiom of Revealed Preference*.

**Definition 5 Weak Axiom of Revealed Preference (WARP).** A dataset $S$ satisfies WARP if and only if, for each pair of distinct bundles $q_i$, $q_j$, $i, j \in N$ with $q_i\, R_0\, q_j$, it is not the case that $q_j\, R_0\, q_i$.

WARP is the first rationalizability condition proposed in the literature. It requires the revealed preference relation to be asymmetric. The intuition behind it is simple: if the decision maker shows through her decision that she prefers bundle $q_i$ over $q_j$ at time $i$, then she cannot at another time show that she prefers $q_j$ over $q_i$ (assuming she behaves as a utility maximizer). In other words, WARP is a necessary condition for rationalizability by a single-valued utility function (see Section 3). On the other hand, we notice that WARP does not require the direct revealed preference relation to be transitive, so that WARP is not sufficient for rationalizability.

The work of Samuelson was further developed by Houthakker (1950), who noted that by using transitivity, the direct revealed preference relation could be extended to an indirect relation.

**Definition 6 Revealed preference relation.** For any sequence of observations $i_1, i_2, \dots, i_k \in N$, if $q_{i_1}\, R_0\, q_{i_2}\, R_0 \dots R_0\, q_{i_k}$, we say that $q_{i_1}$ is *revealed preferred over* $q_{i_k}$, and we write $q_{i_1}\, R\, q_{i_k}$.

Using these revealed preference relations, Houthakker formulated the *Strong Axiom of Revealed Preference*.

**Definition 7 Strong Axiom of Revealed Preference (. SARP )**

A dataset $S$ satisfies SARP if and only if for each pair of distinct bundles $q_i$, $q_j$, $i, j \in N$ with $q_i\, R\, q_j$, it is not the case that $q_j\, R_0\, q_i$.

In order to allow for indifference between bundles, Varian (1982) introduced the strict direct revealed preference relation, and using this relation, defined the *generalized axiom of revealed preference*, GARP.

**Definition 8 Strict direct revealed preference relation.** For any pair of observations $i, j \in N$, if $p_i q_i > p_i q_j$, we say that $q_i$ is *strictly revealed preferred over* $q_j$, and we write $q_i\, P_0\, q_j$.

**Definition 9 Generalized Axiom of Revealed Preference (. GARP)**

A dataset $S$ satisfies GARP if and only if for each pair of distinct bundles, $q_i$, $q_j$, $i, j \in N$, such that $q_i\, R\, q_j$, it is not the case that $q_j\, P_0\, q_i$.

**Example 1.** Consider the following small dataset consisting of four observations.

$$p_1 = (2, 2, 2) \qquad q_1 = (2, 2, 2)$$
$$p_2 = (1, 2, 4) \qquad q_2 = (4, 0, 2)$$
$$p_3 = (2, 1, 3) \qquad q_3 = (4, 4, 0)$$
$$p_4 = (4, 2, 1) \qquad q_4 = (0, 1, 4)$$

Table 1 contains the values $p_i q_j$ for $i, j = 1, \dots, 4$.

**Table 1**
$p_i q_j$ for $i, j = 1, \ldots, 4$.

| | $q_1$ | $q_2$ | $q_3$ | $q_4$ |
|---|---|---|---|---|
| $p_1$ | **12** | 12 | 16 | 10 |
| $p_2$ | 13 | **12** | 12 | 18 |
| $p_3$ | 12 | 14 | **12** | 13 |
| $p_4$ | 14 | 18 | 24 | **8** |

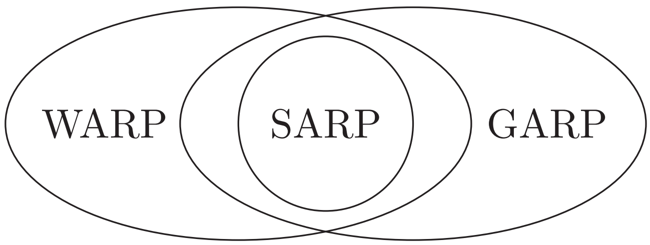


**Fig. 1.** Relations of the axioms of revealed preference.

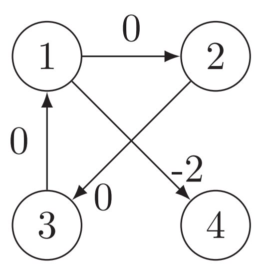


**Fig. 2.** A revealed preference graph.

Clearly, there are direct revealed preference relations $q_1 R_0 q_2$, $q_2 R_0 q_3$, $q_3 R_0 q_1$ and a strict direct revealed preference relation $q_1 P_0 q_4$. This dataset satisfies both WARP and GARP, but not SARP since $q_1 R q_3$ and $q_3 R_0 q_1$.

Fig. 1 illustrates the relations between the different core axioms of revealed preference theory (WARP, SARP, and GARP). Indeed, any dataset satisfying SARP satisfies both WARP and GARP, and there exist datasets not satisfying SARP that satisfy both WARP and GARP (see Example 1).

### 2.3. Graphs representing a dataset

We now describe how to build a directed graph that can be used to represent a dataset; this construction originates from Koo (1971). As we wil see in Section 3, such graphs are very useful tools in deciding rationalizability. Given a datset $S = \{(p_i, q_i) | \; i \in N\}$, we build a directed weighted graph $G_S = (V_S, A_S)$ as follows. For each observation $i \in N$, there is a node in $V_S$, i.e., $V_S := N$. Further, there is an arc from node $i$ to node $j$ in $A_S$ exactly when $p_i q_i \geq p_i q_j$ and $q_i \neq q_j$ (or equivalently, when $q_i R_0 q_j$ and $q_i \neq q_j$). Observe that in $G_S$ there is no arc between distinct observations that feature an identical bundle. Finally, the length of an arc $(i, j) \in A_S$ equals $p_i(q_j - q_i)$. Notice that this length is always nonpositive.

**Example 1 Continued.** The revealed preference graph corresponding to the dataset is given in Fig. 2. Notice that the direct, but not strict, revealed preference relations correspond to an arc of length 0, while the strict revealed preference relations correspond to arcs of strictly negative length.

An alternative version of this construction was proposed by Talla Nobibon et al. (2016). These authors defined a directed graph $G_{R_0}$ which is simply the graph of the direct preference relation $R_0$: the node set of $G_{R_0}$ is again $N$, and there is an arc from node $i$ to node $j$ if and only if $q_i R_0 q_j$ (including when $q_i = q_j$). For the dataset in Example 1, $G_{R_0} = G_S$ since no bundle appears twice.

## 3. Fundamental results

In this section, we connect the fundamentals given in Section 2, and we formulate the theorems that characterize rationalizability. Clearly, a main goal within revealed preference theory is to test whether there exists a (particular) utility function rationalizing a given dataset $S$.

### 3.1. Testing GARP

Necessary and sufficient conditions for rationalizability of a given dataset by a well-behaved utility function are given in Theorem 1.

**Theorem 1.** *(GARP)*
*The following statements are equivalent:*

1. *The dataset $S = \{(p_i, q_i) | \; i \in N\}$ is rationalizable by a well-behaved utility function $u(q)$.*
2. *There exist strictly positive numbers $U_i$, $\lambda_i$ for $i \in N$ satisfying the system of linear inequalities*
$$U_i \leq U_j + \lambda_j p_j (q_i - q_j) \quad \forall i, j \in N. \tag{3}$$
3. *$S$ satisfies GARP.*
4. *Each arc contained in a cycle of the graph $G_S$ has length 0.*

The inequalities comprising system (3) are called the *Afriat Inequalities*. It is not difficult to see that system (3) can be reformulated as a linear program. Indeed, notice that multiplying a given feasible solution $(U_i, \lambda_i: i \in N)$ by any positive constant gives again a feasible solution; thus, one can require each of the variables to be at least equal to 1, and not just strictly positive. The equivalence of statements 1 and 2 in Theorem 1 was established by Afriat (1967b), and their equivalence with statement 3 is due to Varian (1982). Statement 4 is easily derived from the definition of GARP. Thus, Afriat (1967b) provided a linear program, formed by the Afriat Inequalities, that characterizes rationalizability by a well-behaved utility function. This allows us to conclude that GARP can be tested in polynomial time (although no polynomial time algorithms for solving linear programming problems were known at the time when Afriat published his work).

Rationalizability tests for consistency of datasets with GARP have gone through a number of stages. Diewert (1973) states another linear programming formulation. Varian's formulation of GARP (Varian, 1982) provides another algorithm for testing rationalizability. This formulation shows that rationalizability can be tested by computing the transitive closure of the direct revealed preference relation. This transitive closure yields all revealed preference relations, direct and indirect. Given the transitive closure, GARP can be tested by checking, for each pair of bundles $q_i$, $q_j$, $i$, $j \in N$, whether both $q_i R q_j$ and $q_j P_0 q_i$ simultaneously hold. The bottleneck in this procedure is the computation of the transitive closure. Varian suggests to use Warshall's algorithm (Warshall, 1962), which has a worst-case time complexity of $O(n^3)$; he also notes the existence of faster algorithms based on matrix multiplication, which at the time achieved $O(n^{2.74})$ complexity (Munro, 1971). By now, these algorithms have improved, the best known algorithms for general matrices having $O(n^{2.373})$ time complexity (Coppersmith & Winograd, 1990; Le Gall, 2014; Williams, 2012).

Recently, Talla Nobibon, Smeulders, and Spieksma (2015) described an algorithm with a worst-case bound of $O(n^2)$ for GARP, based on the computation of strongly connected components of the graph $G_S$. An alternative, simple statement of the $O(n^2)$ test is derived in Talla Nobibon et al. (2016) from the observation that a dataset $S$ satisfies GARP if and only if $p_i q_i = p_i q_j$ for each arc ($i$, $j$) contained in a strongly connected component of $G_{R_0}$ (see Condition 4 of Theorem 1). Shiozawa (2016) describes yet another way to test GARP in $O(n^2)$ time, using shortest path algorithms. Talla Nobibon et al. (2015) prove a lower bound on testing GARP, showing that no algorithm can exist with time complexity smaller than $O(n \log n)$.

### 3.2. Testing SARP

Analogously to Theorem 1, we now give a theorem that provides necessary and sufficient conditions relating to SARP.

**Theorem 2.** *(SARP)*

*The following statements are equivalent:*

1. *The dataset $S = \{(p_i, q_i) | \ i \in N\}$ is rationalizable by a well-behaved, single-valued utility function $u(q)$.*
2. *There exist strictly positive numbers $U_i$, $\lambda_i$ for $i \in N$ satisfying the system of linear inequalities*

$$U_i < U_j + \lambda_j p_j (q_i - q_j) \quad \forall i, j \in N. \tag{4}$$

3. *$S$ satisfies SARP.*
4. *The graph $G_S$ is acyclic.*

Houthakker (1950), extending the work of Samuelson, introduced the formulation of SARP and proved the equivalence of statements 1 and 3. Statement 2 is an extension of Theorem 1.

Again, observe that system (4) can be cast into a linear optimization format. Using a matrix representation of the direct revealed preference relations, Koo (1963) describes a sufficient condition for consistency with SARP. Dobell (1965) is the first to describe conditions which are both necessary and sufficient. Dobell's test is based on the matrix representation of direct revealed preference relations. He proposes checking whether every square submatrix of the direct revealed preference matrix contains at least one row and one column consisting completely of elements equal to 0. Since there is an exponential number of such submatrices, this test runs in exponential time. Koo (1971) later publishes another paper where he observes that testing SARP amounts to checking whether $G_S$ is acyclic: this can be done in $O(n^2)$ time, and is to-date the most efficient available method for testing consistency with SARP. An alternative version of this test is provided by Talla Nobibon et al. (2016). These authors observe that $S$ satisfies SARP if and only if, within each strongly connected component of $G_{R_0}$, all bundles are identical. This condition can again be checked in $O(n^2)$ time by relying on Tarjan's algorithm to compute all strong components of $G_{R_0}$ (Tarjan, 1972).

### 3.3. Testing WARP

For the sake of completeness, let us now state an easy result which is in fact nothing but a restatement of the definition of WARP.

**Theorem 3.** *(WARP)*

*The following statements are equivalent:*

1. *The dataset $S = \{(p_i, q_i) | \ i \in N\}$ satisfies WARP.*
2. *The graph $G_S$ does not contain any cycle consisting of two arcs.*

As mentioned before, satisfying WARP is only a necessary condition for rationalizability by a single-valued utility function. However, in the special case where the dataset involves only two goods (i.e., $m = 2$), WARP is both a necessary and sufficient condition for rationalizability by a single-valued utility function (Little, 1949; Samuelson, 1948).

Testing WARP can be done in $O(n^2)$ time, since it is sufficient to test each pair of observations for a violation. More explicitly, after having computed the quantities $p_i q_i$ and $p_i q_j$ for all distinct $i, j \in N$, WARP can be rejected if and only if there exists a pair of distinct $i$, $j \in N$ such that $p_i q_i \geq p_i q_j$ and $p_j q_j \geq p_j q_i$.

Finally, let us point out that the graph characterization of GARP, SARP and WARP allows us to easily conclude (using Fig. 2) that the dataset given in Example 1 satisfies WARP (as there are no 2-cycles in $G_S$), satisfies GARP (as the cycle 1-2-3 has length 0), and does not satisfy SARP (as $G_S$ is not acyclic).

Rationalizability questions are not limited to general utility functions. In the next sections, we are interested in the question whether datasets can be rationalized by utility functions of a specific form (Section 4), by collective choice processes (Section 6), or by stochastic choice processes (Section 7).

## 4. Other classes of utility functions and their rationalizability

Besides the basic tests discussed in the previous paragraphs, conditions and tests have been derived for testing rationalizability by various specific forms of utility functions. In this section we consider two additional classes of utility functions: utility functions that are *separable* (Section 4.1), and utility functions that are *homothetic* (Section 4.2). In addition, we assume from now on that the utility functions are *non-satiated*. This is a concept used to model the property that for every bundle $q$ there is another bundle $q'$ in the neighborhood of $q$ that is preferred over $q$. Formally (Jehle & Reny, 2011):

**Definition 10 Non-satiated utility functions.** A utility function $u(\cdot)$ is non-satiated if, for each $q \in \mathbb{R}^m$ and for each $\epsilon > 0$, there exists $q' \in \mathbb{R}^m$ with $||q' - q|| \leq \epsilon$ such that $u(q') > u(q)$.

The property of non-satiatedness expresses that, in the absence of a budget constraint, no particular bundle is preferred to all other bundles. It also imposes some form of continuity to the preferences over bundles.

### 4.1. Separable utility functions

Separability of a utility function refers to the property that different goods in a bundle may have no joint effect on the utility of the bundle; then, goods can be regarded as independent of each other. More generally, it is often assumed that there exists a partition of the goods into $R$ subsets such that goods from different sets do not interact. Hence, separability of a utility function is defined with respect to a given partition of the goods. More concretely, given a partition of the goods into $R$ disjoint sets, we denote by $m_j$ the number of goods in set $j$, $1 \leq j \leq R$. Any bundle of goods can then be written as $q = (q^1, \ldots, q^R)$, with $q^j \in \mathbb{R}_+^{m_j}$ denoting the vector of quantities for the goods in set $j$, $1 \leq j \leq R$.

There are two versions of separability: strong and weak. We first provide the definition of a strongly separable (also known as *additive*) utility function.

**Definition 11 Strongly separable utility functions.** A utility function $u(q)$ is strongly separable with respect to a given partition of the set of goods $\{1, 2, \ldots, m\}$ if and only if there exist well-behaved functions $f_j(q^j) : \mathbb{R}_+^{m_j} \to \mathbb{R}_+$ for each $j \in \{1, \ldots, R\}$ such that

$$u(q) = f_1(q^1) + f_2(q^2) + \cdots + f_R(q^R).$$

The case where we partition the set of goods into two subsets, i.e., the case $R = 2$, allows the following theorem due to Varian (1983):

**Theorem 4.** *The following statements are equivalent:*

1. *There exists a strongly separable, well-behaved, non-satiated utility function $u(f(q^1), q^2)$ rationalizing the dataset $S = \{(p_i, q_i) | \ i \in N\}$.*

2. *There exist strictly positive numbers $U_i$, $V_i$, $\lambda_i$ with $i \in N$ satisfying the system of linear inequalities*

$$U_i \leq U_j + \lambda_j p_j^1 (q_i^1 - q_j^1) \ \forall i, j \in N, \tag{5}$$

$$V_i \leq V_j + \lambda_j p_j^2 (q_i^2 - q_j^2) \ \forall i, j \in N. \tag{6}$$

Varian (1983) also gives a linear programming formulation for arbitrary $R$, allowing for a polynomial-time test of rationalizability by a strongly separable utility function.

A weaker version of separability occurs when the utilities of the different sub-bundles are not necessarily summed to obtain the total utility; weak separability rather assumes that there exists a function, denoted $u'$, that takes as input the utilities of the individual groups of goods, and translates these into a total utility.

**Definition 12 Weakly separable utility functions.** A utility function $u(q)$ is weakly separable with respect to $q^1, \ldots, q^{R-1}$ if and only if there exist functions $f_j(q^j) : \mathbb{R}_+^{m_j} \to \mathbb{R}_+$ for each $j \in \{1, \ldots, R-1\}$ and a function $u'(x_1, \ldots, x_{R-1}, q^R)$ such that $u(q) = u'(f_1(q^1), \ldots, f_{R-1}(q^{R-1}), q^R)$.

Following his paper on general utility functions, Afriat also wrote an unpublished work on separable utility functions (Afriat, 1967a). Varian (1983) built further on this, giving a non-linear system of inequalities, reproduced below in Theorem 5, for which the existence of a solution is a necessary and sufficient condition for rationalizability by a well-behaved, weakly separable utility function with $R = 2$ sets of goods.

**Theorem 5.** *The following statements are equivalent.*

1. *There exists a weakly separable, well-behaved, non-satiated utility function $u(f(q^1), q^2)$ rationalizing the dataset $S = \{(p_i, q_i) | \ i \in N\}$.*
2. *There exist strictly positive numbers $U_i$, $V_i$, $\lambda_i$, $\mu_i$ for $i \in N$ satisfying the system of non-linear inequalities*

$$U_i \leq U_j + \lambda_j p_j^2 (q_i^2 - q_j^2) + (\lambda_j / \mu_j)(V_i - V_j) \ \forall i, j \in N, \tag{7}$$

$$V_i \leq V_j + \mu_j p_j^1 (q_i^1 - q_j^1) \ \forall i, j \in N. \tag{8}$$

Diewert and Parkan (1985) extend this result to multiple separable subsets. Cherchye, Demuynck, De Rock, and Hjertstrand (2015) prove that testing rationalizability by a weakly separable utility function is NP-HARD even for $R = 2$. They also provide an integer programming formulation which is equivalent to (7) and (8). Several heuristic approaches have been formulated for testing weak separability. Varian attempts to overcome the computational difficulties by finding a solution to the linear part of the system of inequalities and then fixing variables based on this solution, which linearizes the remainder of the inequalities. This implementation can be too restrictive, as the variables are usually fixed with values making the system infeasible, even if a solution exists, as shown by Barnett and Choi (1989). Fleissig and Whitney (2003) take a similar approach, but improve on it by fixing variables with values that are more likely to allow solutions to the rest of the system of equalities. Exact tests of (adaptations of) Varian's inequalities are described in Swofford and Whitney (1994) and Fleissig and Whitney (2008). Both use non-linear programming packages to find solutions and are limited in the size of datasets they can handle. Computational results in Cherchye et al. (2015) suggest that the integer programming approach is feasible for moderately sized datasets. Hjertstrand, Swofford, and Whitney (2016) use this approach in an application testing separability of consumption, leisure and money. When dropping the concavity assumption, the rationalizability problem remains NP-HARD, even if the dataset is limited to 9 goods (Echenique, 2014). Quah (2014) provides an algorithm for testing separable utility functions without the concavity assumption. Swofford and Whitney (1994) modify (7) and (8) to account for consumers needing time to adjust their spending.

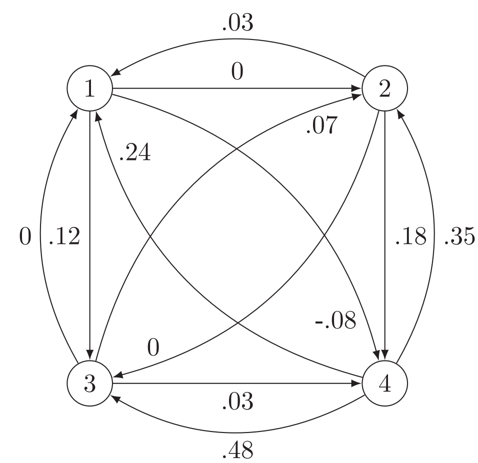


**Fig. 3.** A revealed preference graph for testing homotheticity.

*4.2. Homothetic utility functions*

Another class of utility functions of interest are the *homothetic utility functions*. Their definition is based on the concept of a homogenous function.

**Definition 13 Homogenous functions.** A function $f(\cdot)$ is homogenous when $f(\lambda q) = \lambda f(q)$, for each $q \in \mathbb{R}^m$ and for each $\lambda \in \mathbb{R}$.

**Definition 14 Homothetic utility functions.** A utility function $u(\cdot)$ is homothetic when there exist a homogenous function $f$ and a monotonic function $\ell$ such that $u(q) = \ell(f(q))$ for each $q \in \mathbb{R}^m$.

In effect, if $u$ is homothetic and if $u(q_i) \geq u(q_j)$ for two bundles $q_i$, $q_j$, then for any constant $\alpha > 0$, $u(\alpha q_i) \geq u(\alpha q_j)$. Theorem 6 gives necessary and sufficient conditions for rationalizability of a dataset by a homothetic utility function. Notice that for tests of homothetic utility functions described in the theorem, we assume the price vectors are normalized so that $p_i q_i = 1$ for all $i \in N$. One of these conditions is based on the following graph $H = (V_S, A_S)$ (whose construction is in the spirit of the construction described in Section 2.3). For each observation $i \in N$, there is a node in $V_S$, i.e., $V_S := N$. Further, for each ordered pair of observations $(i, j)$, there is an arc of length $\log(p_i q_j)$ between the corresponding nodes. Fig. 3 shows a graph to test homotheticity for the dataset given in Example 1.

**Theorem 6.** *The following statements are equivalent:*

1. *There exists a non-satiated homothetic utility function $u(\cdot)$ rationalizing the dataset $S = \{(p_i, q_i) | \ p_i q_i = 1, \ \forall i \in N\}$.*
2. *There exist strictly positive numbers $U_i$ for $i \in N$ satisfying the inequalities*

$$U_i \leq U_j p_j q_i \ \forall i, j \in N. \tag{9}$$

3. *For all distinct choices of observations $(i_1, i_2, \ldots, i_k)$, we have*

$$(p_{i_1} q_{i_2})(p_{i_2} q_{i_3}) \ldots (p_{i_k} q_{i_1}) \geq 1. \tag{10}$$

4. *The graph $H_S$ does not contain a cycle of negative length.*

The equivalence of statements 1, 2 and 3 was proven by Afriat (1972, 1981). Based on statement 4, Varian (1983) proposes a combinatorial test which can be implemented in $O(n^3)$ time.

**Table 2**
Complexity results for testing rationalizability by utility functions of specific forms.

| Type of utility function | Type of test | Time complexity |
|---|---|---|
| General | Graph test | $O(n^2)$ |
| Single-valued | Graph test | $O(n^2)$ |
| Strongly separable | System of linear ineq. | Polynomial |
| Weakly separable | System of non-linear ineq. | NP-HARD |
| Homothetic | Graph test | $O(n^3)$ |
| Homothetic and separable | System of non-linear ineq. | Open |

Varian (1983) also provides a test for homothetic, separable utility functions, which is again a difficult-to-solve system of non-linear inequalities. Finally, utility maximization in case of rationing (i.e., when there are additional linear constraints on the bundles which can be bought, on top of the budget constraint) is also handled by Varian. He provides a linear system of inequalities whose feasibility is a necessary and sufficient condition for rationalizability.

In summary, various forms of utility functions are usually associated with a system of inequalities, for which the existence of a solution is a necessary and sufficient condition for rationalizability by such a utility function. The difficulty of these rationalizability tests crucially depends on whether the systems are linear or non-linear. General, single-valued and strongly separable utility functions are easy to rationalize, as their associated systems of inequalities are linear. The same holds true for utility maximization by a general utility function under rationing constraints. For general and single-valued utility functions, more straightforward tests have been developed. A polynomial test also exists for rationalizability by a homothetic utility function. On the other hand, for those utility functions associated with non-linear systems of inequalities, that is, weakly separable and homothetic separable functions, no efficient tests are known. For weakly separable utility, formal NP-HARDNESS results exist. For homothetic separable functions, the complexity question remains open. Varian (1982, 1983) provides a way to construct consistent utility functions for all of these settings. Table 2 summarizes these results.

To complete our overview on rationalizability by general utility functions, we mention some recent work on indivisible goods and non-linear budget sets. More precisely, these are settings where the optimization problems (1) and (2) are further constrained by the conditions that (i) some components of $q$ are integral, and (ii) the budget constraint is non-linear (e.g., in the presence of quantity discounts), and/or there are multiple budget constraints. Forges and Minelli (2009) give a revealed preference characterization for non-linear budgets, for which GARP is a sufficient and necessary condition for rationalizability by an increasing and continuous utility function. Cherchye, Demuynck, and De Rock (2014) give conditions for rationalizability by an increasing, concave and continuous utility function for the setting with non-linear budgets. They note that, together with the results by Forges and Minelli, this allows for tests of the concavity of utility functions which are not possible in the setting with linear budgets. Computationally there is no obvious easy way to test the conditions laid out by Cherchye et al. in general. However, they show that if the budgets can be represented by a finite union of polyhedral convex sets, a system of linear inequalities provides conditions for rationalizability. Fujishige and Yang (2012) and Polisson and Quah (2013) extend the revealed preference results to the case with indivisible goods. They find that GARP is a necessary and sufficient test for rationalizability, given a suitable adaptation of the revealed preference relations for their setting. Cosaert and Demuynck (2015) look at choice sets which are non-linear and have a finite number of choice alternatives. They provide revealed preference characterizations for weakly monotone, strongly monotone, weakly monotone and concave, and strongly monotone and concave utility functions, all of which are easy to test, either by some variant of GARP or a system of linear inequalities.

## 5. Goodness-of-fit and power measures

An often cited limitation of rationalizability tests is that they are binary tests: either the dataset is rationalizable or it is not. Thus, when violations of rationalizability conditions are found, there is no indication of their severity. Likewise, when the rationalizability conditions are satisfied, this could be because the choices faced by the decision maker make it unlikely that violations would occur. To refine this yes/no verdict inherent to rationalizability, so-called goodness-of-fit measures and power measures have been proposed in the literature. Goodness-of-fit measures (Section 5.1) quantify the severity of violations, while power measures (Section 5.2) indicate how far the choices are from violating rationalizability conditions.

### 5.1. Goodness-of-fit measures

A first class of goodness-of-fit measures is based on the systems of inequalities which are used to establish rationalizability of many different forms of utility functions (see Section 3). Slack variables are added to these systems, so as to relax the constraints on the data. An optimization problem can then be defined, for which the objective function is the minimization of some appropriate function of the slack variables, such as their sum, under the constraint that the system of equalities is satisfied. The goodness-of-fit measure is then equal to the value of the optimal solution of this optimization problem. Such an approach was first described by Diewert (1973) and has since been used in a number of different papers for various forms of utility functions (see Diewert & Parkan, 1985; Fleissig & Whitney, 2005; Fleissig & Whitney, 2008 for weak separability, Fleissig and Whitney (2007) for additive separability). Computing the goodness-of-fit measure is easy if the system of inequalities is linear, which is the case for general utility functions and additive separable utility functions. In the case of non-linear systems of inequalities, minimizing the sum of the slack variables is at least as hard as finding a solution to the system without slack variables. Since this is already NP-HARD for weakly separable utility functions, the hardness result remains valid for these goodness-of-fit measures.

A second class of goodness-of-fit measures is due to Afriat (1973), and is based on strengthening the revealed preference relations. In this case, revealed preference relations are assumed to hold if the difference in price between the chosen bundle and another affordable bundle is big enough. This is done by introducing *efficiency indices* $0 \leq e_i \leq 1$ for each observation $i \in N$, and defining the revealed preference relation $R_0(e_1, \ldots, e_n)$ as follows:

$$\text{for all } i, j \in N, \text{ if } e_i p_i q_i \geq p_i q_j, \text{ then } q_i R_0(e_1, \ldots, e_n)\, q_j. \quad (11)$$

Obviously, when $e_i = 1$, conditions (11) are the same revealed preference relations as in Definition 3; when $e_i < 1$, condition (11) can be interpreted as defining a revealed preference relation between two bundles for which the price difference exceeds a certain fraction of the budget. As a result, there will be fewer revealed preference relations, and axioms such as WARP, SARP and GARP will be easier to satisfy. A goodness-of-fit measure is then the maximum value of the sum of the $e_i$ values, under the constraint that a given axiom of revealed preference is satisfied by $R_0(e_1, \ldots, e_n)$. Three different goodness-of-fit indices based on this idea have been respectively described by Afriat (1973), Varian (1990) and Houtman and Maks (1985). Of these three, Afriat's index is the simplest, as it constrains the $e_i$ values to be equal for every observation ($e_1 = e_2 = \cdots = e_n$). Afriat's index can be computed in polynomial time (see Smeulders, Spieksma, Cherchye, & De Rock, 2014), although for a long time the only published algorithm was an approximation algorithm due to Varian (1990). Varian's index, in contrast, allows the $e_i$ values to differ between observations. This makes computation less straightforward and the computation of this index was thus perceived to be hard (as confirmed by Smeulders et al. (2014) who showed that computing Varian's index is NP-HARD). This led to work on heuristic algorithms for computing Varian's index by Varian (1990), Tsur (1989), and more recently by Alcantud, Matos, and Palmero (2010). Finally, Houtman and Maks (1985) proposed to constrain the $e_i$ values to be either 0 or 1. In effect, maximizing the sum of the $e_i$'s then amounts to removing the minimum number of observations so that the remaining dataset is rationalizable. Houtman and Maks established a link between the feedback vertex set problem (known to be NP-HARD) and their index, thus informally showing its difficulty; see Hjertstrand and Heufer (2015) for two methods computing the Houtman–Maks index. The complexity of computing all three of the above indices is addressed by Smeulders et al. (2014), who provide polynomial time algorithms for Afriat's index for various axioms of revealed preference, and establish NP-hardness of Varian's index, and of the Houtman–Maks index. Even stronger, it is shown that no constant-factor approximation algorithms running in polynomial time exist for these indices unless P = NP. Boodaghians and Vetta (2015) strengthen these hardness results, by showing that computing the Houtman–Maks index is already NP-HARD for datasets with only 3 goods.

A third approach to the definition of goodness-of-fit measures was introduced by Varian (1985). When a dataset fails to satisfy the rationalizability conditions, the goal is here to find a dataset which does satisfy the conditions and is only minimally different from the observed dataset. The problem of finding these minimally different rationalizable datasets can be formulated as a non-linear optimization problem, which, in general, is hard to solve. To avoid solving large scale non-linear problems, De Peretti (2005) approaches this problem with an iterative procedure. Working on GARP, his algorithm tackles violations one at a time, also perturbing only one observation at a time. If a preference cycle exists between two bundles of goods $q_i$ and $q_j$, $i, j \in N$, he computes the minimal perturbation necessary to remove the violation both for the case in which $q_i R_0 q_j$ (in which case $q_i$ is perturbed) and for the case in which $q_j R_0 q_i$ (in which case $q_j$ is perturbed). The smallest of the two perturbations is then used to update the dataset, and the new dataset is checked again for GARP violations. While this algorithm does not guarantee an optimal solution, it allows handling large datasets, especially if the number of violations is small.

A number of recent papers introduce new goodness-of-fit measures, thus showing continued interest in this topic. Echenique et al. (2011) define the mean and median money pump indices. In their paper, the severity of violations of rationality is measured by the amount of money which an arbitrageur could extract from the decision maker by exploiting her irrational choices. This is reflected by a money pump index for every violation of rationality. Echenique et al. propose to calculate the money pump index of the mean and median violation as measures of the irrationality of the decision maker. Computing these measures is NP-HARD, as shown in Smeulders, Cherchye, De Rock, and Spieksma (2013). In the latter paper, it is also shown that computing the money pump index for the most and least severe violations can be done in polynomial time. Furthermore, Apesteguia and Ballester (2015) introduce the minimal swaps index. Informally, the swaps index of a given preference ordering over the alternatives is calculated by counting how many better alternatives (according to the preference order) were not chosen over all choice situations. The minimal swaps index is then the swaps index of the preference order for which this index is minimal. Apesteguia and Ballester show that computing the minimal swaps index is equivalent to the NP-HARD linear ordering problem. Finally, Dean and Martin (2016) define the minimum cost index. This index is the minimum cost of removing revealed preference relations, such that the remaining relations induce no violations. The cost of removing violations is weighted by the price difference of the considered bundles. Dean and Martin show that computing this index is NP-HARD by a reduction from the set covering problem.

### 5.2. Power measures

Power measures were first introduced by Bronars (1987), with the following motivation. Consider a test that allows us to determine whether the observations in a dataset are coherent with the choices of a utility-maximizing decision maker. If the outcome of the test is positive for most datasets, *including those where choices were not made so as to maximize a utility function,* then obviously the test is not good at discriminating between utility maximizing behavior and alternative behaviors. Power measures are numerical values indicating to what extent a test is able to discriminate between samples coming from a rational or from an irrational decision maker.

Bronars (1987) proposes to use random choices as an alternative model of behavior. The likelihood of this alternative model satisfying the rationalizability conditions (that is, passing the test) is determined by Monte Carlo simulation. The higher this likelihood, the lower the power of the test. Andreoni and Miller (2002) use a similar approach: they generate synthetic datasets by bootstrapping from observed choices, and use these alternative datasets to establish the power of their test.

Bronars's Monte Carlo approach has also been applied to goodness-of-fit measures. The value of a goodness-of-fit measure is hard to interpret without context. There is no natural level which, if crossed, indicates a large deviation from rational behavior. Furthermore, the values of goodness-of-fit indices which point to large deviations may vary from dataset to dataset, as the choices faced by a decision maker may or may not allow large violations of rationalizability. One way to establish what values are significant, is to generate random datasets by a Monte Carlo approach and to calculate their goodness-of-fit measures. This yields a distribution of the values of goodness-of-fit measures for datasets of random choices. It can then be checked whether the goodness-of-fit measures computed for the actual decision makers are significantly different. Examples of this approach are found in Choi, Fisman, Gale, and Kariv (2007) and Heufer (2012). As this framework requires a large number of computations of the goodness-of-fit measures, there is a strong incentive to use efficient algorithms and to favor measures which are easy to calculate.

Beatty and Crawford (2011) propose to evaluate the power of a test by calculating the proportion of possible choices which would pass the test. Andreoni, Gillen, and Harbaugh (2013) give an overview of power measures and introduce a number of new power measures themselves. The measures they introduce are adaptations of goodness-of-fit measures. For example, they introduce a jittering index, which is the minimum perturbation of the data such that the rationalizability conditions are no longer satisfied, in line with the work of Varian (1985). They also introduce an Afriat Power Index, which is the converse of Afriat's goodness-of-fit measure; that is, instead of considering the maximum value of $e \leq 1$ in (11) such that the dataset satisfies the considered axiom of revealed preference, they propose to determine the minimum value of $e \geq 1$ such that the dataset does not satisfy the conditions.

## 6. Collective choices

In the preceding sections, datasets are analyzed as if a single person buys or chooses goods, so as to maximize her own utility function. However, in many cases purchasing decisions are observed at the household level that consists of multiple decision makers. The choices that result from collective decision making may appear irrational, even if all individual decision makers have rational preferences. For example, Arrow's impossibility theorem (Arrow, 1950) shows that for non-dictatorial, unanimous preference aggregation functions, independence of irrelevant alternatives cannot be guaranteed. As a result, the group can exhibit choice reversals if more choice alternatives are added. Moreover, a group can use different choice mechanisms at different times, giving more or less power to different group members, also leading to choices that appear irrational. Analyzing datasets resulting from collective choices thus calls for collective models, which account for individually rational household members, and in addition, some decision process for splitting up the budget. Example 2 shows how the joint purchases of two rational decision makers can appear irrational when they are analyzed as if there was a unique decision maker.

**Example 2.** Consider the following dataset with 2 periods and 3 goods.

$$p_1 = (3, 2, 1) \qquad q_1 = (5, 4, 7) \tag{12}$$

$$p_2 = (2, 3, 1) \qquad q_2 = (3, 5, 9) \tag{13}$$

Then, bundle 1 would be strictly revealed preferred over bundle 2, since $p_1q_1 = 30 > 28 = p_1q_2$. Likewise, bundle 2 would be strictly revealed preferred over bundle 1, since $p_2q_2 = 30 > 29 = p_2q_1$. The dataset thus does not satisfy GARP. However, consider the following datasets.

$$p_1 = (3, 2, 1) \qquad q_1^1 = (5, 0, 0)$$
$$p_2 = (2, 3, 1) \qquad q_2^1 = (3, 0, 0)$$

and

$$p_1 = (3, 2, 1) \qquad q_1^2 = (0, 4, 7)$$
$$p_2 = (2, 3, 1) \qquad q_2^2 = (0, 5, 9)$$

It is clear that both of these satisfy GARP, since for the first dataset $q_1^1 > q_2^1$, and for the second dataset $q_2^2 > q_1^2$. Furthermore, notice that $q_1 = q_1^1 + q_1^2$ and $q_2 = q_2^1 + q_2^2$. The datasets (12) and (13) thus represent the joint purchases of two rational decision makers.

The initial contributions in revealed preference theory dealing with collective choice are published by Chiappori (1988), for the so-called *labor supply* setting. This setting corresponds to a situation in which there are two goods, namely leisure time and aggregated consumption, which are observed for each member in the household. Also, we assume that the household consists of two decision makers. The behavior of this household is then rationalizable if the consumption can be split up so that the resulting individual datasets of leisure and consumption are rationalizable for all individual household members. Chiappori provides conditions for rationalizability, both for the cases with and without externalities of private consumption. To model the labor supply setting in the collective choice model, we use a dataset of the form $S = \{(w_i^1, w_i^2, L_i^1, L_i^2, C_i) | \; i \in N\}$, with $w_i^1$ and $w_i^2$ corresponding to the wages of household members 1 and 2, with $L_i^1$ and $L_i^2$ corresponding to their respective leisure time, and with $C_i$ denoting the level of (collective) consumption in the household ($i \in N$). Notice that, since wages can be seen as the price of leisure time, and there is a unit price for aggregated consumption, we can write $p_i = (w_i^1, \mathbf{1})$ and $q_i = (L_i^1, fC_i)$ (for some fraction $0 \le f \le 1$). Hence, the dataset $S$ can still be seen as a set of observations consisting of price vectors and bundles.

**Theorem 7.** *(Chiappori's Theorem for collective rationalization by egoistical agents)*

*The following statements are equivalent.*

1. *There exists a pair of concave, monotonic, continuous non-satiated utility functions which provide a collective rationalization by egoistical agents.*
2. *There exist numbers $Z_i$ with $0 \le Z_i \le C_i$ such that the following (equivalent) conditions are satisfied.*
   (a) *The datasets $\{(w_i^1, \mathbf{1}), (L_i^1, Z_i) | \; i \in N\}$ and $\{(w_i^2, \mathbf{1}), (L_i^2, C_i - Z_i) | \; i \in N\}$ both satisfy SARP.*
   (b) *There exist strictly positive numbers $U_i^1, U_i^2, \lambda_i, \mu_i$ for $i \in N$ satisfying the non-linear inequalities*
   $$U_i^1 \le U_j^1 + \lambda_j w_j^1 (L_i^1 - L_j^1) + \lambda_j (Z_i - Z_j) \qquad \forall i, j \in N,$$
   $$U_i^2 \le U_j^2 + \mu_j w_j^2 (L_i^2 - L_j^2) + \mu_j (C_i - Z_i - C_j - Z_j) \qquad \forall i, j \in N,$$
   *with equality holding in the first (respectively, the second) inequality only if $L_i^1 = L_j^1$ and $Z_i = Z_j$ (respectively, $L_i^2 = L_j^2$ and $Z_i = Z_j$).*

Theorem 7 states Chiappori's result for collective rationalization by egoistical agents. (The agents are egoistical in the sense that they each spend their own personal wages, so that the observed consumption is just the sum of the individual ones.) No straightforward method is included in the paper to test the first condition; the second condition requires solving a system of non-linear inequalities. Similar conditions hold for the case with externalities. Snyder (2000) provides a reformulation of Chiappori's conditions for two periods and uses it in empirical tests. Thanks to the limit on the number of periods, this test is very easy: it requires solving four small linear systems of inequalities. Cherchye, De Rock, and Vermeulen (2011) depart from the labor supply setting by formulating a collective model with an arbitrary number of goods. In their model, each specific good is known to be either publicly or privately consumed. Given this information, rationalizability is tested by checking whether there exists a split of prices (for public goods) or quantities (for private goods), such that the dataset of personalized prices and quantities for each household member satisfies GARP. Cherchye et al. (2011) provide an integer programming formulation to test their model. Talla Nobibon et al. (2016) provide a large number of practical and theoretical computational results for this problem. First, they prove it is NP-HARD. Furthermore, they describe a more compact integer programming formulation, and provide a simulated annealing based metaheuristic. They compare the computational results with these different integer programming formulations and heuristics; they observe that the heuristic approach is capable of tackling larger datasets and seldom fails to find a feasible split when one exists. Smeulders, Cherchye, De Rock, Spieksma, and Talla Nobibon (2015) give further hardness results for a collective version of WARP: they find that the problem remains NP-HARD when testing for transitivity is dropped. All hardness results for these problems assume that the number of goods is not fixed a priori. It remains an open question whether the problems become easy for a small, fixed number of goods. In particular, the labor supply setting only requires one good to be partitioned over members of the household.

The work by Chiappori is generalized by Cherchye, De Rock, and Vermeulen (2007). Leaving the labor supply setting, they provide conditions for an arbitrary number of goods and without any prior allocation of goods, as was the case with leisure time in Chiappori's work. Cherchye et al. (2007) derive separate necessary and sufficient conditions for collective rationalizability by concave utility functions. In a later paper, Cherchye, De Rock, and Vermeulen

(2010) show that the necessary condition given in their earlier work is both necessary and sufficient, when dropping the assumption of concave utility functions. However, testing this condition is NP-HARD, as shown by Talla Nobibon and Spieksma (2010). Due to the hardness of rationalizability in collective settings, a number of papers have appeared on how to test this problem. An integer programming formulation is given by Cherchye, De Rock, Sabbe, and Vermeulen (2008) and an enumerative approach is provided by Cherchye, De Rock, and Vermeulen (2009). Talla Nobibon, Cherchye, De Rock, Sabbe, and Spieksma (2011) take a different approach and propose a heuristic algorithm. The goal of this algorithm is to quickly test whether the rationalizability conditions are satisfied. If this heuristic cannot prove that the conditions are satisfied, then an exact test is used. Using this heuristic pre-test, many computationally demanding exact tests can be avoided. Deb (2010) strengthens the hardness results by proving that a special case of this problem, the situation dependent dictatorship setting, is also NP-HARD. In this setting, the household decision process is such that each purchasing decision is made by a single household member, called the dictator. At different points in time, different household members can assume the role of the dictator; the goal is thus to partition the observations into datasets, so that each dataset is consistent with (unitary) GARP. Crawford and Pendakur (2013) also consider this problem in the context of preference heterogeneity, and provide algorithms for computing upper and lower bounds on the number of 'dictators'. Cosaert (2017) links this to the problem of computing the chromatic number of a graph. Furthermore, Cosaert formulates an integer program to partition the observations into sets, so that the observed characteristics within each set are as homogenous as possible. Smeulders et al. (2015) give further hardness results for a collective version of WARP: they find that dropping transitivity makes the test easy for households of two members, but the problem remains open for three or more members.

## 7. Revealed stochastic preference

In the previous sections, we have looked at methods that decide whether a set of observations can be rationalized by one or more decision makers, using different forms of utility functions, or different ways in which the choice process can be split over several decision makers. However, we assumed that utility functions and preferences are fully deterministic. As a result, if a choice situation repeats itself, we expect that the decision maker always chooses the same alternative. However, it is commonly observed in experiments on choice behavior that if a person is given the same choice situation multiple times, her decision may change. One possible way of explaining this behavior is by stochastic preferences, as pioneered by Block and Marschak (1960). Theories of stochastic preferences posit that, while at any point in time a decision maker has a preference ordering over all alternatives, these preferences are not constant over time and may fluctuate randomly. An observed behavior is rationalizable by stochastic preferences if and only if there exists a set of utility functions and a probability distribution over these utility functions, such that the frequency with which an alternative is chosen in any given choice situation is equal to the probability that this alternative has the highest utility in that situation. We note that many results on stochastic preferences are established for the case of finite choice sets, as opposed to the consumption setting, where there exists an infinite number of bundles that can be bought for a given expenditure level and prices. For an overview, we refer to McFadden (2005).

A very general result was established by McFadden and Richter (1990), namely, the *axiom of revealed stochastic preference* (ARSP), which states a necessary and sufficient condition for rationalizability of choice probabilities by stochastic preferences. The generality of this axiom allows it to be used for any form of choice situation, and all classes of decision rules. Besides the axiom, McFadden and Richter also provided a system of linear inequalities whose feasibility is a necessary and sufficient condition for rationalizability. Neither of these characterizations can be easily operationalized, since ARSP places a condition on every possible subset of observations, so that the resulting number of conditions is exponential in the number of observations. Furthermore, each condition requires finding a decision rule among all allowed decision rules which maximizes some function, and this can in itself be an NP-HARD problem (for example when the class of decision rules being tested are based on linear preference orders, this means solving an NP-HARD linear ordering problem; Karp (1972)). The linear system of inequalities, on the other hand, contains one variable for every possible decision rule within a class of decision rules, a number which is often exponential in the number of choice alternatives.

For the setting of consumer purchases (and thus infinite choice sets), Bandyopadhyay, Dasgupta, and Pattanaik (1999) formulate the *weak axiom of stochastic revealed preference* (WARSP). This axiom provides a necessary condition for rationalizability by stochastic preferences. Analogously to WARP, WARSP compares pairs of choice situations. Since the condition placed on these pairs is easy to test, WARSP allows for a polynomial time test. Heufer (2011) and Kawaguchi (2016) build further on this work. Heufer provides a sufficient condition for rationalizability in terms of stochastic preferences. Kawaguchi (2016) proposes the *strong axiom of revealed stochastic preference* (SARSP), a necessary condition for rationalizability by stochastic preferences. Both of these conditions seem difficult to test, requiring in the case of Heufer a feasible solution to a linear program with an exponential number of constraints and variables. Kawaguchi's SARSP likewise requires checking an exponential number of inequalities. Despite these challenges, Kitamura and Stoye (2014) develop a test which can be used to test rationalizability by stochastic preferences on consumption data, though for relatively small datasets. A key element in their approach is discretizing the dataset, so as to return to a setting with a finite number of choice options.

## 8. Conclusion

In this final section, let us summarize our discussion, and outline perspectives regarding possible future developments in the field. It is indisputable that revealed preference theory has established itself as an important tool in economics. On the other hand, testing revealed preference axioms on large datasets gives rise to numerous algorithmic challenges that should appeal to the operations researcher community. While a thorough understanding of individual rational choice, as it relates to revealed preference, has been achieved, we see (at least) three research directions emerging:

1. Economists are increasingly extending the revealed preference setting to more complex theories of choice behavior, such as collective decision making, or non-deterministic choices. The testing problems emerging in these cases are likewise more complex. Much work, both theoretical and algorithmically remains to be done in this area.
2. Many complexity hardness results have been established under the assumption that the number of goods can be arbitrarily large, as opposed to assuming that this number is limited and fixed (e.g., $m = 2$ or $m = 3$). We have mentioned in this survey a few results that hold when the number of goods is fixed, but many questions remain open in this direction. Beyond its theoretical interest, this setting has practical relevance, since in many empirical studies the number of goods is quite small, or

goods are aggregated into a limited number of classes. Tests that are difficult in general may turn out to be polynomially computable in these cases.

3. The relevance of efficient revealed preference tests for large datasets (see Section 1.1) continues to increase due to the ever growing size of available datasets. Better algorithms, both heuristic and exact, are required in order to be able to cope with this phenomenon. Thus, we need to further increase our understanding of the achievable running times for different versions of the rationalizability question.

Answering these questions will not only reveal the inherent difficulty of testing rationalizability of a given dataset by a utility function from a particular class, it will also shed light on the incentives and properties of human behavior.

## Acknowledgments

The authors thank Roman Słowiński and three anonymous reviewers for helpful suggestions on a preliminary version of the paper. This research was supported by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office (grant P7/36). Bart Smeulders is a postdoctoral fellow of the Belgian F.R.S.-FNRS.

## References

Afriat, S. (1967a). *The construction of separable utility functions from expenditure data*. Purdue University. Unpublished

Afriat, S. (1967b). The construction of utility functions from expenditure data. *International Economic Review, 8*(1), 67–77.

Afriat, S. (1972). The theory of international comparisons of real income and prices. In *Proceedings of the international comparisons of prices and output* (pp. 11–84). NBER.

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

Afriat, S. (1981). On the constructability of consistent price indices between several periods simultaneously. In A. Deaton (Ed.), *Essays in applied demand analysis*. Cambridge University Press.

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

Andreoni, J., Gillen, B., & Harbaugh, W. (2013). The power of revealed preference tests: Ex-post evaluation of experimental design. Available at http://people.hss.caltech.edu/~bgillen/papers/GARPPower.pdf.

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

Antonelli, G. (1886). *Sulla teoria matematica della economia politica*. Pisa: Tipografia Del Folchetto. [English translation in Preference, Utility and Demand, edited by J.S. Chipman et al. New York, Chapter 16 (1971)]

Apesteguia, J., & Ballester, M. (2015). A measure of rationality and welfare. *Journal of Political Economy, 123*(6), 1278–1310.

Arrow, K. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy, 58*(4), 328–346.

Bandyopadhyay, T., Dasgupta, I., & Pattanaik, P. (1999). Stochastic revealed preference and the theory of demand. *Journal of Economic Theory, 84*(1), 95–110.

Barnett, W., & Choi, S. (1989). A Monte Carlo study of tests of blockwise weak separability. *Journal of Business & Economic Statistics, 7*(3), 363–377.

Beatty, T., & Crawford, I. (2011). How demanding is the revealed preference approach to demand? *The American Economic Review, 101*(6), 2782–2795.

Block, H., & Marschak, J. (1960). Random orderings and stochastic theories of responses. *Contributions to Probability and Statistics, 2*, 97–132.

Blundell, R. (2005). How revealing is revealed preference? *Journal of the European Economic Association, 3*(2–3), 211–235.

Boodaghians, S., & Vetta, A. (2015). Testing consumer rationality using perfect graphs and oriented discs. *Proceedings of the web and internet economics* (pp. 187–200). Springer.

Bouyssou, D., & Pirlot, M. (2016). Conjoint measurement tools for MCDM: A brief introduction. In S. Greco, M. Ehrgott, & J. Figueira (Eds.), *Multiple criteria decision analysis: State of the art surveys* (pp. 97–151). Springer.

Bronars, S. (1987). The power of nonparametric tests of preference maximization. *Econometrica: Journal of the Econometric Society, 55*, 693–698.

Chambers, C., & Echenique, F. (2016). *Revealed preference theory*. Cambridge University Press.

Cherchye, L., De Rock, B., Griffith, R., O'Connell, M., Smith, K., & Vermeulen, F. (2017a). A new year, a new you? Heterogeneity and self-control in food purchases. *Technical Report*. Department of Economics, KU Leuven. Discussion Paper Series 17.19

Cherchye, L., De Rock, B., Sabbe, J., & Vermeulen, F. (2008). Nonparametric tests of collectively rational consumption behavior: An integer programming procedure. *Journal of Econometrics, 147*(2), 258–265.

Cherchye, L., De Rock, B., & Vermeulen, F. (2007). The collective model of household consumption: A nonparametric characterization. *Econometrica, 75*(2), 553–574.

Cherchye, L., De Rock, B., & Vermeulen, F. (2009). Opening the black box of intrahousehold decision making: Theory and nonparametric empirical tests of general collective consumption models. *Journal of Political Economy, 117*(6), 1074–1104.

Cherchye, L., De Rock, B., & Vermeulen, F. (2010). An Afriat theorem for the collective model of household consumption. *Journal of Economic Theory, 145*(3), 1142–1163.

Cherchye, L., De Rock, B., & Vermeulen, F. (2011). The revealed preference approach to collective consumption behaviour: Testing and sharing rule recovery. *The Review of Economic Studies, 78*(1), 176–198.

Cherchye, L., Demuynck, T., & De Rock, B. (2014). Revealed preference analysis for convex rationalizations on nonlinear budget sets. *Journal of Economic Theory, 152*, 224–236.

Cherchye, L., Demuynck, T., De Rock, B., & Hjertstrand, P. (2015). Revealed preference tests for weak separability: An integer programming approach. *Journal of Econometrics, 186*(1), 129–141.

Cherchye, L., Demuynck, T., De Rock, B., & Vermeulen, F. (2017b). Household consumption when the marriage is stable. *American Economic Review, 107*(6), 1507–1534.

Chiappori, P. (1988). Rational household labor supply. *Econometrica, 56*(1), 63–90.

Choi, S., Fisman, R., Gale, D., & Kariv, S. (2007). Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review, 97*(5), 1921–1938.

Choi, S., Kariv, S., Müller, W., & Silverman, D. (2014). Who is (more) rational? *American Economic Review, 104*(6), 1518–1550.

Coppersmith, D., & Winograd, S. (1990). Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation, 9*(3), 251–280.

Cormen, T., Leiserson, C., Rivest, R., & Stein, C. (2001). *Introduction to algorithms* (3rd). Cambridge, London: MIT Press.

Corrente, S., Greco, S., Kadziński, M., & Słowiński, R. (2013). Robust ordinal regression in preference learning and ranking. *Machine Learning, 93*, 381–422.

Corrente, S., Greco, S., Matarazzo, B., & Słowiński, R. (2016). Robust ordinal regression for decision under risk and uncertainty. *Journal of Business Economics, 86*(1), 55–83.

Cosaert, S. (2017). What types are there? Computational Economics.

Cosaert, S., & Demuynck, T. (2015). Revealed preference theory for finite choice sets. *Economic Theory, 59*, 169–200.

Crawford, I., & De Rock, B. (2014). Empirical revealed preference. *Annual Review of Economics, 6*, 503–524.

Crawford, I., & Pendakur, K. (2013). How many types are there? *The Economic Journal, 123*(567), 77–95.

De Peretti, P. (2005). Testing the significance of the departures from utility maximization. *Macroeconomic Dynamics, 9*(03), 372–397.

Dean, M., & Martin, D. (2016). Measuring rationality with the minimum cost of revealed preference violations. *Review of Economics and Statistics, 98*(3), 524–534.

Deb, R. (2010). An efficient nonparametric test of the collective household model. University of Toronto, Unpublished.

Diewert, E. (1973). Afriat and revealed preference theory. *The Review of Economic Studies, 40*(3), 419–425.

Diewert, E., & Parkan, C. (1985). Tests for the consistency of consumer data. *Journal of Econometrics, 30*(1), 127–147.

Dobell, A. (1965). A comment on AYC Koo's "An empirical test of revealed preference theory". *Econometrica: Journal of the Econometric Society, 33*(2), 451–455.

Dyer, J. (2016). Multiattribute utility theory (MAUT). In S. Greco, M. Ehrgott, & J. Figueira (Eds.), *Multiple criteria decision analysis: State of the art surveys* (pp. 285–314). Springer.

Echenique, F. (2014). Testing for separability is hard. Working paper, California Institute of Technology.

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

Fishburn, P. (1970). *Utility theory for decision-making*. Wiley.

Fleissig, A., & Whitney, G. (2003). A new PC-based test for Varian's weak separability conditions. *Journal of Business & Economic Statistics, 21*(1), 133–144.

Fleissig, A., & Whitney, G. (2005). Testing for the significance of violations of Afriat's inequalities. *Journal of Business & Economic Statistics, 23*(3), 355–362.

Fleissig, A., & Whitney, G. (2007). Testing additive separability. *Economics Letters, 96*(2), 215–220.

Fleissig, A., & Whitney, G. (2008). A nonparametric test of weak separability and consumer preferences. *Journal of Econometrics, 147*(2), 275–281.

Forges, F., & Minelli, E. (2009). Afriat's theorem for general budget sets. *Journal of Economic Theory, 144*(1), 135–145.

Fujishige, S., & Yang, Z. (2012). On revealed preference and indivisibilities. *Modern Economy, 3*(6), 752–758.

Fürnkranz, J., & Hüllermeier, E. (Eds.). (2010). *Preference learning*. Springer-Verlag.

Garey, M., & Johnson, D. (1979). *Computers and intractability*. San Francisco, CA: Freeman.

Giesen, J., Mueller, K., Taneva, B., & Zolliker, P. (2010). Choice-based conjoint analysis: Classification vs. discrete choice models. In J. Fürnkranz, & E. Hüllermeier (Eds.), *Preference learning* (pp. 297–315). Springer-Verlag.

Greco, S., Ehrgott, M., & Figueira, J. (Eds.). (2016). *Multiple criteria decision analysis: State of the art surveys* (2nd). Springer.

Gustafsson, A., Herrmann, A., & Huber, F. (Eds.). (2007). *Conjoint measurement: Methods and applications*. Springer-Verlag.

Hausman, D. (2000). Revealed preference, belief, and game theory. *Economics and Philosophy, 16*(1), 99–115.
Heufer, J. (2011). Stochastic revealed preference and rationalizability. *Theory and Decision, 71*(4), 575–592.
Heufer, J. (2012). Testing for utility maximization with error and the loss of power. *German Economic Review, 13*(2), 161–173.
Hjertstrand, P., & Heufer, J. (2015). Consistent subsets: Computationally feasible methods to compute the Houtman–Maks-index. *Economics Letters, 128*, 87–89.
Hjertstrand, P., Swofford, J., & Whitney, G. (2016). Mixed integer programming revealed preference tests of utility maximization and weak separability of consumption, leisure, and money. *Journal of Money, Credit and Banking, 48*(7), 1547–1561.
Houthakker, H. (1950). Revealed preference and the utility function. *Economica, 17*(66), 159–174.
Houtman, M. (1995). *Nonparametric consumer and producer analysis*. Maastricht University.
Houtman, M., & Maks, J. (1985). Determining all maximal data subsets consistent with revealed preference. *Kwantitatieve Methoden, 19*, 89–104.
Jehle, G., & Reny, P. (2011). *Advanced microeconomic theory*. Pearson.
Karp, R. (1972). Reducibility among combinatorial problems. *Complexity of Computer Computations, 40*(4), 85–103.
Kawaguchi, K. (2016). Testing rationality without restricting heterogeneity. *Journal of Econometrics, 197*(1), 153–171.
Keeney, R., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. Wiley.
Kitamura, Y., & Stoye, J. (2014). Nonparametric analysis of random utility models: Testing. Working paper, Cornell University.
Koo, A. (1963). An empirical test of revealed preference theory. *Econometrica, 31*(4), 646–664.
Koo, A. (1971). Revealed preference–a structural analysis. *Econometrica, 39*(1), 89–97.
Le Gall, F. (2014). Powers of tensors and fast matrix multiplication. In *Proceedings of the thirty-ninth international symposium on symbolic and algebraic computation* (pp. 296–303).
Little, I. (1949). A reformulation of the theory of consumer's behaviour. *Oxford Economic Papers, 1*(1), 90–99.
McFadden, D. (2005). Revealed stochastic preference: A synthesis. *Economic Theory, 26*(2), 245–264.
McFadden, D., & Richter, M. (1990). Stochastic rationality and revealed stochastic preference. In *Preferences, uncertainty, and optimality: Essays in honor of Leo Hurwicz* (pp. 161–186). Westview Press.
Moretti, S., Öztürk, M., & Tsoukiàs, A. (2016). Preference modelling. In S. Greco, M. Ehrgott, & J. Figueira (Eds.), *Multiple criteria decision analysis: State of the art surveys* (pp. 43–96). Springer.
Munro, I. (1971). Efficient determination of the transitive closure of a directed graph. *Information Processing Letters, 1*(2), 56–58.
Polisson, M., & Quah, J. (2013). Revealed preference in a discrete consumption space. *American Economic Journal: Microeconomics, 5*(1), 28–34.
Quah, J. (2014). A revealed preference test for weakly separable preferences. Discussion Paper Series, No. 708, University of Oxford.
Rao, V. (2014). *Applied conjoint analysis*. Springer-Verlag.
Riabacke, M., Danielson, M., & Ekenberg, L. (2012). State-of-the-art prescriptive criteria weight elicitation. *Advances in Decision Sciences, 2012*.
Roy, B. (1991). The outranking approach and the foundations of ELECTRE methods. *Theory and Decision, 31*, 49–73.
Samuelson, P. (1938). A note on the pure theory of consumer's behaviour. *Economica, 5*(17), 61–71.
Samuelson, P. (1948). Consumption theory in terms of revealed preference. *Economica, 15*, 243–253.
Shiozawa, K. (2016). Revealed preference test and shortest path problem; graph theoretic structure of the rationalizability test. *Journal of Mathematical Economics, 67*, 38–48.
Siskos, Y., Grigoroudis, E., & Matsatsinis, N. (2016). UTA methods. In S. Greco, M. Ehrgott, & J. Figueira (Eds.), *Multiple criteria decision analysis: State of the art surveys* (pp. 315–362). Springer.
Smeulders, B., Cherchye, L., De Rock, B., & Spieksma, F. (2013). The money pump as a measure of revealed preference violations: A comment. *Journal of Political Economy, 121*(6), 1248–1258.
Smeulders, B., Cherchye, L., De Rock, B., Spieksma, F., & Talla Nobibon, F. (2015). Complexity results for the weak axiom of revealed preference for collective consumption models. *Journal of Mathematical Economics, 58*, 82–91.
Smeulders, B., Spieksma, F., Cherchye, L., & De Rock, B. (2014). Goodness of fit measures for revealed preference tests: Complexity results and algorithms. *ACM Transactions on Economics and Computation, 2*(1), 3.
Snyder, S. (2000). Nonparametric testable restrictions of household behavior. *Southern Economic Journal, 67*(1), 171–185.
Swofford, J., & Whitney, G. (1994). A revealed preference test for weakly separable utility maximization with incomplete adjustment. *Journal of Econometrics, 60*(1), 235–249.
Talla Nobibon, F., Cherchye, L., Crama, Y., Demuynck, T., De Rock, B., & Spieksma, F. (2016). Revealed preference tests of collectively rational consumption behavior: Formulations and algorithms. *Operations Research, 64*(6), 1197–1216.
Talla Nobibon, F., Cherchye, L., De Rock, B., Sabbe, J., & Spieksma, F. (2011). Heuristics for deciding collectively rational consumption behavior. *Computational Economics, 38*(2), 173–204.
Talla Nobibon, F., Smeulders, B., & Spieksma, F. (2015). A note on testing axioms of revealed preference. *Journal of Optimization Theory and Applications, 166*(3), 1063–1070.
Talla Nobibon, F., & Spieksma, F. (2010). On the complexity of testing the collective axiom of revealed preference. *Mathematical Social Sciences, 60*(2), 123–136.
Tarjan, R. (1972). Depth-first search and linear graph algorithms. *SIAM Journal on Computing, 1*, 146–160.
Tsur, Y. (1989). On testing for revealed preference conditions. *Economics Letters, 31*(4), 359–362.
Varian, H. (1982). The nonparametric approach to demand analysis. *Econometrica, 50*(4), 945–973.
Varian, H. (1983). Non-parametric tests of consumer behaviour. *Review of Economic Studies, 50*(1), 99–110.
Varian, H. (1985). Non-parametric analysis of optimizing behavior with measurement error. *Journal of Econometrics, 30*(1), 445–458.
Varian, H. (1990). Goodness-of-fit in optimizing models. *Journal of Econometrics, 46*(1–2), 125–140.
Wallenius, J., Dyer, J., Fishburn, P., Steuer, R., Zionts, S., & Deb, K. (2008). Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. *Management Science, 54*, 1336–1349.
Warshall, S. (1962). A theorem on Boolean matrices. *Journal of the ACM, 9*(1), 11–12.
Williams, V. (2012). Multiplying matrices faster than Coppersmith–Winograd. In *Proceedings of the forty-fourth annual ACM symposium on theory of computing* (pp. 887–898). ACM.
Wong, S. (2006). *The foundations of Paul Samuelsons revealed preference theory* (Revised). Routledge, Taylor & Francis Group.
Zionts, S., & Wallenius, J. (1976). An interactive programming method for solving the multiple criteria problem. *Management Science, 22*, 652–663.
Zionts, S., & Wallenius, J. (1983). An interactive multiple objective linear programming method for a class of underlying nonlinear utility functions. *Management Science, 29*, 519–529.

Please cite this article as: B. Smeulders et al., Revealed preference theory: An algorithmic outlook, European Journal of Operational Research (2018), https://doi.org/10.1016/j.ejor.2018.04.026