Computing Afriat's Critical Efficiency Index via Graph Decomposition: Based on the Structure of Strongly Connected Components

Takeshi Naitoh*

November 7, 2025

Abstract This paper proposes a graph-based algorithm for computing Afriat's Critical Efficiency Index (ACEI), which measures the minimum relaxation required to rationalize a dataset under the Generalized Axiom of Revealed Preference (GARP). The method iteratively removes critical arcs within strongly connected components (SCCs), ordered by their revealed efficiency ratios. The final threshold produced by this deletion process is equal to the ACEI, providing a structure-aware alternative that avoids global enumeration of efficiency thresholds and yields an interpretable decomposition of revealed-preference violations. Furthermore, by accumulating the violation lengths of the removed arcs, the algorithm provides a tractable approximation to Dean and Martin's Minimum Cost Index (MCI), capturing both the threshold level of inefficiency and the cumulative cost of necessary corrections. Compared with existing threshold-based and global-optimization approaches, the proposed algorithm integrates computation and interpretation, offering a practical tool for evaluating and visualizing inconsistency in empirical consumption data.

Keywords: Revealed preference, Goodness-of-fit measures, Strongly connected components

1 Introduction

Afriat's theorem [1] provides necessary and sufficient conditions under which observed consumption behavior can be rationalized by a utility function. However, in the binary framework assumed by this theorem, even a slight violation of the Generalized Axiom of Revealed Preference (GARP) leads to the entire dataset being classified as irrational. To overcome such a strict dichotomy, researchers have proposed various *goodness-of-fit* indices that quantitatively assess the degree of deviation from rationality. Among them, Afriat's Critical Efficiency Index (ACEI), introduced by Afriat [2], measures the minimal uniform reduction in revealed expenditures required for the dataset to satisfy GARP. In addition, the Minimum Cost Index (MCI) proposed by Dean and Martin [3] quantifies GARP violations from a cost-removal perspective: it measures the minimal total expenditure that must be removed to eliminate all revealed preference violations.

In recent years, both the theoretical foundations and computational properties of such indices have been extensively studied. Smeulders, Crama, and Spieksma [14] survey the computational

*Faculty of Economics, Shiga University, 1-1-1 Banba, Hikone, Shiga 522-8522, Japan

development of ACEI, ranging from Varian's early approximation methods to the first exact polynomial-time algorithm. Dziewulski, Lanier, and Quah [4] further clarify that rationalizability conditions such as GARP are fundamentally characterized by the absence of cycles in the revealed preference graph. Moreover, Lanier and Quah [8] propose a conceptual framework based on three dimensions—*continuity*, *accuracy*, and *concreteness*—and demonstrate that no index can satisfy all three simultaneously. Against this background, and building on the interpretation of the CCEI as approximate cost-rationalizability [10], we set the stage for a structural, graph-based approach to measuring near-rationality.

Alongside these developments, recent studies have drawn attention to the structural role of *strongly connected components* (SCCs) in revealed preference graphs. Fujishige and Yang [7] use SCC decomposition to simplify GARP testing in the context of discrete rationalizability. Building on this, Shiozawa [12] proposes a unified framework that reformulates various rationalizability conditions—including GARP, homothetic, and quasi-linear preferences—as instances of the shortest path problem (SPP) and its generalization, the shortest path problem with weight adjustment (SPPWA). He also introduces the *Strongly Connected Component Index* (SCCI), a structurally motivated index based on internal violations within SCCs. More recently, Naitoh [9] proposes indices based on the lengths of violating arcs within SCCs, contributing to a more structural understanding of rationality violations. These structural insights motivate our contribution: we develop an SCC-based, graph-theoretic algorithm that identifies ACEI without threshold search, visualizes the structure of revealed preference violations, and yields an efficiently computable approximation to MCI.

The remainder of this paper is structured as follows. Section 2 introduces the revealed preference graph and reformulates GARP in graph-theoretic terms. Section 3 builds on the concept of e-GARP to define Afriat's Critical Efficiency Index (ACEI) as the minimal efficiency level required to eliminate all cycles in the graph. Section 4 presents a graph-based algorithm for computing ACEI and proves its correctness. Section 5 introduces a new index based on the accumulated cost of violations, and compares it with the Minimum Cost Index (MCI) proposed in previous studies.

2 Data, notation, and graph-theoretic preliminaries

We introduce some notation used throughout the paper. There are n different types of goods in the market. The consumer has a budget b for consumption and a utility function $U : \mathbf{R}^n_+ \to \mathbf{R}$. We consider a market analyst observing a finite dataset $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t)\}_{t=1}^T$, where $\boldsymbol{p}^t = (p_1^t, \dots, p_n^t) \in \mathbf{R}^n_{++}$ is a positive price vector, and $\boldsymbol{x}^t = (x_1^t, \dots, x_n^t) \in \mathbf{R}^n_+ \setminus \{\mathbf{0}\}$ is the consumer's demand bundle observed at time t, given the available budget $b_t \in \mathbf{R}_+$. The inner product $\boldsymbol{p}^t \cdot \boldsymbol{x}^t = \displaystyle\sum_{i=1}^n p_i^t x_i^t$ represents the total expenditure at time t, which we also refer to as the (revealed) cost of bundle $\boldsymbol{x}^t$ under prices $\boldsymbol{p}^t$. We assume $\boldsymbol{p}^t \cdot \boldsymbol{x}^t = b_t$. The dataset D is *rationalized* by a utility function U in the sense that for all $t \leq T$, $\boldsymbol{x}^t$ maximizes U over $\{\boldsymbol{x} \mid \boldsymbol{p}^t \cdot \boldsymbol{x} \leq \boldsymbol{p}^t \cdot \boldsymbol{x}^t\}$. The basic question raised by Afriat is whether the dataset is rationalized by a locally non-satiated utility function U.

A dataset D satisfies WARP if and only if, for each pair of distinct bundles $\boldsymbol{x}^i$, $\boldsymbol{x}^j$, $i, j \leq T$ with $\boldsymbol{p}^i \cdot \boldsymbol{x}^i \geq \boldsymbol{p}^i \cdot \boldsymbol{x}^j$, it is not the case that $\boldsymbol{p}^j \cdot \boldsymbol{x}^j \geq \boldsymbol{p}^j \cdot \boldsymbol{x}^i$. We also say that the consumer's behavior satisfies Generalized Axiom of Revealed Preference (GARP) if $(\boldsymbol{p}^{t_1}, \boldsymbol{x}^{t_1}), (\boldsymbol{p}^{t_2}, \boldsymbol{x}^{t_2}), \ldots, (\boldsymbol{p}^{t_m}, \boldsymbol{x}^{t_m})$ satisfying $\boldsymbol{p}^{t_k} \cdot \boldsymbol{x}^{t_k} \geq \boldsymbol{p}^{t_k} \cdot \boldsymbol{x}^{t_{k+1}}$ $(k = 1, \ldots, m-1)$ for all $t_1, \ldots, t_m \leq T$, we have $\boldsymbol{p}^{t_m} \cdot \boldsymbol{x}^{t_1} \geq \boldsymbol{p}^{t_m} \cdot \boldsymbol{x}^{t_m}$. A utility function U is said to *rationalize* the observed dataset D if $U(\boldsymbol{x}^t) \geq U(\boldsymbol{x})$ for all $\boldsymbol{x}$ such that $\boldsymbol{p}^t \cdot \boldsymbol{x}^t \geq \boldsymbol{p}^t \cdot \boldsymbol{x}$.

Theorem 2.1 (Afriat's Theorem [1], [18]): *The following four statements are equivalent:*
(a) *The dataset D can be rationalized by a locally non-satiated utility function U.*
(b) *The dataset D satisfies GARP.*
(c) *There is a positive solution ϕ, λ to the set of linear inequalities $\lambda_j \leq \phi_i + \lambda_i \boldsymbol{p}^i \cdot (\boldsymbol{x}^j - \boldsymbol{x}^i)$ for all i, j.*
(d) *The dataset D can be rationalized by a continuous, concave, strictly monotone increasing utility function U.* □

Let V and A be finite sets. A *directed graph* $G = (V, A)$ consists of a set V of vertices and a set A of arcs, whose elements are ordered pairs of distinct vertices. Throughout this paper, all graphs are assumed to be directed unless otherwise stated. If $a \in A$, $i, j \in V$, and $a = (i, j)$, then we say that *a joins* i to j. We also call i the *tail* of a and j the *head* of a. A *path* in $G = (V, A)$ is a sequence $P = (i_1, \ldots, i_\ell)$ of distinct vertices i_k $(k = 1, \ldots, \ell)$ with $\ell \geq 2$, such that $(i_k, i_{k+1}) \in A$ $(k = 1, \ldots, \ell - 1)$. The end vertices of this path are i_1 and i_ℓ, and the path is said to be an (i_1, i_ℓ)*-path*. If P is an (i_1, i_ℓ)-path in $G = (V, A)$ and $a \in A$ is an arc that joins i_ℓ to i_1, then $C = (i_1, \ldots, i_\ell, i_1)$ is called a *cycle*.

A graph $G' = (W, B)$ is called a *subgraph of* $G = (V, A)$ if $W \subseteq V$ and $B \subseteq A$. For a vertex subset $W \subseteq V$, the subgraph $G[W]$ of G whose vertex set is W and whose arc set consists of the arcs of G joining vertices of W is called the subgraph of G induced by W. We denote by $G - a$ the graph obtained from $G = (V, A)$ by deleting the arc $a \in A$. Furthermore, if $B \subseteq A$, we denote by $G - B$ the graph obtained by deleting the arcs in B. A (sub)graph H is said to be *strongly connected* if for every two vertices i, j in the graph H there exists a path in H from i to j. A maximal strongly connected subgraph of a graph $G = (V, A)$ is called a *strongly connected component* of the graph G. G is decomposed into its strongly connected components $H_k = (V_k, A_k)$ $(k \in K)$ where $\{V_k \mid k \in K\}$ is a partition of V. An algorithm by Tarjan ([16]) finds a partition in linear time, $O(|V| + |A|)$.

To evaluate revealed preferences, we consider the matrix of expenditure differences between observed bundles.

Definition 2.2 (Data Matrix): *Given a dataset $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t)\}_{t=1}^T$, the data matrix $D_T = [d_{ij}] \in \mathbb{R}^{T \times T}$ is defined by*

$$d_{ij} := \boldsymbol{p}^i \cdot (\boldsymbol{x}^j - \boldsymbol{x}^i).$$

Here, d_{ij} represents the revealed cost difference of bundle $\boldsymbol{x}^j$ relative to $\boldsymbol{x}^i$, evaluated under price vector $\boldsymbol{p}^i$.

This matrix forms the basis for a graph-theoretic representation of revealed preference relations.

Definition 2.3 (**Difference-based Revealed Preference Graph**): *Let* $D_T = [d_{ij}]$ *be the data matrix defined above. We define the difference-based revealed preference graph* $G_D^{\leq 0} = (V_T, A^{\leq 0})$ *by:*

- $V_T = \{1, 2, \ldots, T\}$,
- $A^{\leq 0} = \{(i, j) \in V_T \times V_T \mid i \neq j,\, d_{ij} \leq 0\}$.

Each directed arc $(i, j) \in A^{\leq 0}$ *is assigned the length* d_{ij}.

We say that bundle $\boldsymbol{x}^i$ is revealed to be preferred to $\boldsymbol{x}^j$ in the difference-based sense if $(i, j) \in A^{\leq 0}$. A cycle in a directed graph is called a *negative-length cycle* if the total length, defined as the sum of the arc lengths along the cycle, is strictly negative.

Example 2.4: *Let* $t = 1, \ldots, 6$. *Suppose that the dataset* $D = \left\{(\boldsymbol{p}^t, \boldsymbol{x}^t)\right\}_{t=1}^{6}$ *is given by*

$$\{((3,5,5),(26,25,29)),((4,5,4),(21,34,24)),((4,4,5),(31,28,22)),((4,5,5),(29,25,24)), ((3,5,6),(26,27,24)),((4,5,6),(30,27,24))\}$$

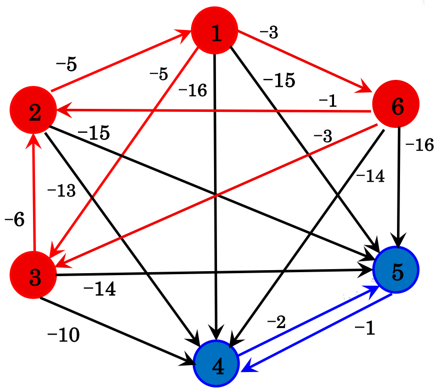
and that the consumer has a budget of $b_1 = 348$, $b_2 = 350$, $b_3 = 346$, $b_4 = 361$, $b_5 = 357$, $b_6 = 399$,

$$\begin{pmatrix} 3 & 5 & 5 \\ 4 & 5 & 4 \\ 4 & 4 & 5 \\ 4 & 5 & 5 \\ 3 & 5 & 6 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 26 & 21 & 31 & 29 & 26 & 30 \\ 25 & 34 & 28 & 25 & 27 & 27 \\ 29 & 24 & 22 & 24 & 24 & 24 \end{pmatrix} = \begin{pmatrix} 348 & 353 & 343 & 332 & 333 & 345 \\ 345 & 350 & 352 & 337 & 335 & 351 \\ 349 & 340 & 346 & 336 & 332 & 348 \\ 374 & 374 & 374 & 361 & 359 & 375 \\ 377 & 377 & 365 & 356 & 357 & 369 \\ 403 & 398 & 396 & 385 & 383 & 399 \end{pmatrix}.$$

Hence corresponding data matrix D_T *is*

$$D_T = \begin{pmatrix} 0 & 5 & -5 & -16 & -15 & -3 \\ -5 & 0 & 2 & -13 & -15 & 1 \\ 3 & -6 & 0 & -10 & -14 & 2 \\ 13 & 13 & 13 & 0 & -2 & 14 \\ 20 & 20 & 8 & -1 & 0 & 12 \\ 4 & -1 & -3 & -14 & -16 & 0 \end{pmatrix}.$$

Based on the above data matrix D_T, *we construct the subgraph* $G_D^{\leq 0}$, *which consists of all arcs with non-positive lengths. The resulting graph is illustrated below.*



$G_D^{\leq 0}$ contains two strongly connected components. □

Proposition 2.5 ([7]): *The following three statements are equivalent:*
(a) *The data matrix D_T satisfies GARP.*
(b) *Every cycle C in the graph $G_D^{\leq 0}$ satisfies $\boldsymbol{p}^i \cdot (\boldsymbol{x}^j - \boldsymbol{x}^i) = 0$ for all $(i, j) \in C$.*
(c) *Every strongly connected component $H_k = (V_k, A_k)$ of the graph $G_D^{\leq 0}$ satisfies $\boldsymbol{p}^i \cdot (\boldsymbol{x}^j - \boldsymbol{x}^i) = 0$ for all $(i, j) \in A_k$.* □

GARP is equivalent to what Afriat called cyclical consistency (Proposition 2.2 (b)). The cyclic consistency plays a fundamental role in the various literature on revealed preference (Dziewulski, Lanier, and Quah [4]). Algorithms for fast verification of GARP have been developed; see, e.g., Nobibon, Smeulders, and Spieksma [15].

While the difference-based graph structure derived from D_T is central to the verification of GARP, it may not be the most suitable representation for evaluating the degree of inconsistency in a scale-invariant manner. To address this, we consider a ratio-based representation of the same dataset.

Definition 2.6 (Ratio-based Cost Matrix): *Given a dataset $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t)\}_{t=1}^T$, we define the* ratio-based cost matrix *$R_T = [r_{ij}]$ of size $T \times T$ by*

$$r_{ij} := \frac{\boldsymbol{p}^i \cdot \boldsymbol{x}^j}{\boldsymbol{p}^i \cdot \boldsymbol{x}^i}, \quad \text{for } 1 \leq i, j \leq T.$$

Note that $r_{ii} = 1$ for all i. □

Each entry r_{ij} in the matrix represents the relative cost of choosing bundle $\boldsymbol{x}^j$ instead of $\boldsymbol{x}^i$, both evaluated under the price vector $\boldsymbol{p}^i$. This matrix provides the foundation for constructing the graph $RG_D^{\leq e}$, whose acyclicity is related to e-GARP, as will be shown in Proposition 3.4.

This interpretation enables us to define a directed graph that captures efficiency-based revealed preference relations.

Graph-based representations of revealed preference relations, particularly those based on cost ratios, have been discussed in several studies such as [14], and further developed in [8]. Following a similar idea, we define arc weights using relative expenditure ratios aiming to evaluate revealed preferences in a scale-invariant manner. This leads to the following graph structure:

Definition 2.7 (**Ratio-based Revealed Preference Graph**)**:** *Let $R_T = [r_{ij}]$ be the ratio-based cost matrix defined in Definition 2.6. For a given efficiency level $e \in [0, 1]$, we define the directed graph $RG_D^{\leq e} = (V_T, A^{\leq e})$ by:*

- $V_T = \{1, 2, \ldots, T\}$
- $A^{\leq e} = \{(i, j) \in V_T \times V_T \mid i \neq j,\ r_{ij} \leq e\}$

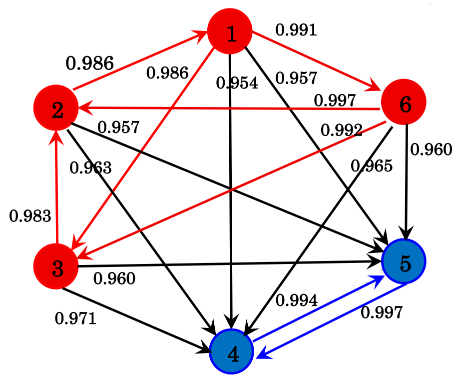
We say that bundle $\boldsymbol{x}^i$ is revealed to be preferred to $\boldsymbol{x}^j$ at efficiency level e if there is a directed arc from i to j in $RG_D^{\leq e}$. □

This graph captures efficiency-adjusted revealed preference relations: An arc (i, j) indicates that bundle $\boldsymbol{x}^i$ is revealed to be preferred to $\boldsymbol{x}^j$, up to an efficiency level e. The following example illustrates how the ratio-based cost matrix R_T and the corresponding revealed preference graph $RG_D^{\leq e}$ are constructed in practice. To facilitate comparison, we use the same dataset as in Example 2.4.

Example 2.8: *The ratio-based cost matrix R_T corresponding to Example 2.4 is given by:*

$$R_T = \begin{pmatrix} 1.000 & 1.014 & 0.986 & 0.954 & 0.957 & 0.991 \\ 0.986 & 1.000 & 1.006 & 0.963 & 0.957 & 1.003 \\ 1.009 & 0.983 & 1.000 & 0.971 & 0.960 & 1.006 \\ 1.036 & 1.036 & 1.036 & 1.000 & 0.994 & 1.039 \\ 1.056 & 1.056 & 1.022 & 0.997 & 1.000 & 1.034 \\ 1.010 & 0.997 & 0.992 & 0.965 & 0.960 & 1.000 \end{pmatrix}.$$

Based on the above ratio-based cost matrix R_T, we construct the subgraph $RG_D^{\leq 1}$, consisting of all directed arcs (i, j) such that $r_{ij} \leq 1$. The resulting graph is shown below.



Each entry $r_{ij} = \frac{p_i \cdot x_j}{p_i \cdot x_i}$ represents the relative expenditure ratio of bundle x_j with respect to x_i, evaluated under the price vector p_i. All entries are rounded to three decimal places. □

Each arc represents a revealed preference at efficiency level e. If a negative-length cycle exists in $RG_D^{\leq e}$ then e-GARP is violated. The maximum value of e for which $RG_D^{\leq e}$ contains no such cycle corresponds to Afriat's Critical Efficiency Index (ACEI). The scalar index considered in this paper, ACEI, originates from Afriat's extension of his earlier work [2]. While the foundational conditions for rationalizability were established in his 1967 paper [1], the concept of a scalar efficiency index was introduced in his 1973 formulation. This index was later studied under the name " Afriat's efficiency index " by Varian [17] and Smeulders et al. [13]. For clarity, we will use the term *ACEI* throughout this paper.

3 Afriat's Critical Efficiency Index and e-GARP

For a given dataset, the revealed preference tests yield a binary outcome: rationalizable or not. However, we are often interested in the degree of violation of rationality. A variety of goodness-of-fit measures for rationality have been proposed. In this section, we focus on Afriat's Critical Efficiency Index (ACEI) and its related concept, Generalized Axiom of Revealed Preference at efficiency level e (***e*-GARP**). Other measures, such as MCI, will be introduced in a later section.

Afriat [2] introduced a partial efficiency index by relaxing the strict requirement of revealed preference. For a given efficiency level e with $0 \leq e \leq 1$, we say that bundle $\boldsymbol{x}^t$ is directly revealed preferred to $\boldsymbol{x}$ if $e\boldsymbol{p}^t \cdot \boldsymbol{x}^t \geq \boldsymbol{p}^t \cdot \boldsymbol{x}$. This idea naturally extends to sequences of such relations, leading to a relaxed version of GARP, called ***e*-GARP**, which allows for a uniform degree of inefficiency across all comparisons. We formally define this condition below.

Definition 3.1 (*e*-**GARP):** *Let $e \in [0,1]$ be a given efficiency level. A finite dataset $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t)\}_{t=1}^T$ is said to satisfy the Generalized Axiom of Revealed Preference at efficiency level e (e-GARP) if for any finite sequence of observations*

$$(\boldsymbol{p}^{t_1}, \boldsymbol{x}^{t_1}), (\boldsymbol{p}^{t_2}, \boldsymbol{x}^{t_2}), \ldots, (\boldsymbol{p}^{t_m}, \boldsymbol{x}^{t_m})$$

satisfying

$$e\boldsymbol{p}^{t_k} \cdot \boldsymbol{x}^{t_k} \geq \boldsymbol{p}^{t_k} \cdot \boldsymbol{x}^{t_{k+1}} \quad (k = 1, \ldots, m-1),$$

we have

$$\boldsymbol{p}^{t_m} \cdot \boldsymbol{x}^{t_1} \geq e\boldsymbol{p}^{t_m} \cdot \boldsymbol{x}^{t_m}.$$

This condition requires that no strictly inefficient cycle exists when evaluating expenditures at a uniform efficiency level e. If $e = 1$, then the condition coincides with the standard GARP, and if $e = 0$, then the condition is trivially satisfied. Hence there is some critical level e^* where the data just satisfy e-GARP. We now define the largest such efficiency level at which e-GARP holds, known as Afriat's Critical Efficiency Index (ACEI).

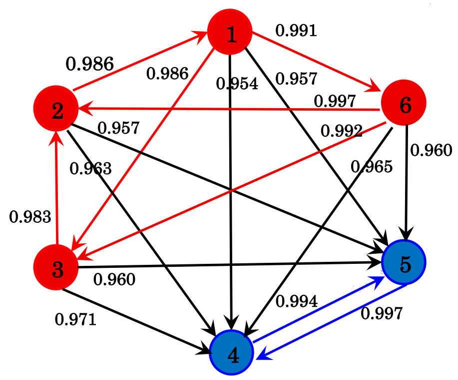
Definition 3.2: (Afriat's Critical Efficiency Index)
For a dataset $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t)\}_{t=1}^T$, Afriat's Critical Efficiency Index (ACEI) *is defined as follows:*

$$\mathrm{ACEI}(D) = \sup_{0 \leq e \leq 1} \left\{ e \,\middle|\, D \,\text{satisfies } e\text{−GARP} \right\}.$$

To illustrate the implications of e-GARP and ACEI, we visualize the revealed preference graph using ratio-based weights defined as

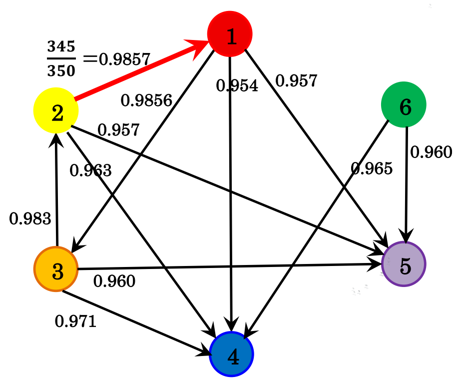
$$r_{ij} := \frac{\boldsymbol{p}^i \cdot \boldsymbol{x}^j}{\boldsymbol{p}^i \cdot \boldsymbol{x}^i},$$

Example 3.3: *Consider the dataset used in Example 2.4 and Example 2.8.*



The graph $RG_D^{\leq 1}$ contains two strongly connected components ($V_1 = \{1,2,3,6\}$, $V_2 = \{4,5\}$), which implies that it includes at least two directed cycles. This shows that the ACEI is strictly less than 1.

If $e = \frac{345}{350}$, then $RG_D^{\leq e}$ contains the arc (2,1) and the cycle $(1,3,2,1)$.



If $e < \frac{345}{350}$, then $RG_D^{\leq e}$ is acyclic. Hence $\mathrm{ACEI}(D) = \dfrac{345}{350} \approx 0.986$ □

We now explore the relationship between e-GARP and the structural properties of the ratio-based revealed preference graph $RG_D^{\leq e}$. The following proposition shows that e-GARP imposes equality on the efficiency ratios along any directed cycle in the graph.

Proposition 3.4: *Let $e \in [0,1]$. If a dataset $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t)\}_{t=1}^T$ satisfies e-GARP, then every directed cycle C in the graph $RG_D^{\leq e}$ consists only of arcs $(i,j) \in C$ for which:*

$$\boldsymbol{p}^i \cdot \boldsymbol{x}^j = e\boldsymbol{p}^i \cdot \boldsymbol{x}^i.$$

Proof: Assume that D satisfies e-GARP. Let C be a directed cycle:

$$i_1 \to i_2 \to \cdots \to i_\ell \to i_1.$$

Since each arc $(i_k, i_{k+1}) \in C$ (with $i_{\ell+1} := i_1$) belongs to $RG_D^{\leq e}$, we have:

$$\boldsymbol{p}^{i_k} \cdot \boldsymbol{x}^{i_{k+1}} \leq e\boldsymbol{p}^{i_k} \cdot \boldsymbol{x}^{i_k} \quad \text{for all } k = 1, \ldots, \ell.$$

Meanwhile, applying the definition of e-GARP to the time series $i_1 \to i_2 \to \cdots \to i_\ell$, we obtain:

$$\boldsymbol{p}^{i_\ell} \cdot \boldsymbol{x}^{i_1} \geq e\boldsymbol{p}^{i_\ell} \cdot \boldsymbol{x}^{i_\ell}.$$

Combining these inequalities and $e \in [0, 1]$ gives:

$$\boldsymbol{p}^{i_\ell} \cdot \boldsymbol{x}^{i_1} \leq e\boldsymbol{p}^{i_\ell} \cdot \boldsymbol{x}^{i_\ell} \leq \boldsymbol{p}^{i_\ell} \cdot \boldsymbol{x}^{i_1},$$

hence equality must hold throughout:

$$\boldsymbol{p}^{i_\ell} \cdot \boldsymbol{x}^{i_1} = e\boldsymbol{p}^{i_\ell} \cdot \boldsymbol{x}^{i_\ell}.$$

By cyclically shifting the cycle and applying the same reasoning, we conclude that for every arc $(i, j) \in C$,

$$\boldsymbol{p}^i \cdot \boldsymbol{x}^j = e\boldsymbol{p}^i \cdot \boldsymbol{x}^i.$$

□

Proposition 3.4 establishes only the forward direction. As noted in Lanier and Quah [8], while e-GARP always implies e-acyclicity, the converse implication does not hold in general. Nonetheless, the Afriat's Critical Efficiency Index (ACEI) can be characterized using the acyclicity condition, as shown below.

Definition 3.5 (Ratio-based Revealed Preference Graph 2): *Let $R_T = [r_{ij}]$ be the ratio-based cost matrix defined in Definition 2.6. For a given efficiency level $e \in [0, 1]$, we define the directed graph $RG_D^{<e} = (V_T, A^{<e})$ by:*

- $V_T = \{1, 2, \ldots, T\}$
- $A^{<e} = \{(i, j) \in V_T \times V_T \mid i \neq j,\ r_{ij} < e\}$

We say that bundle $\boldsymbol{x}^i$ is revealed to be preferred to $\boldsymbol{x}^j$ at efficiency level e if there is a directed arc from i to j in $RG_D^{<e}$. □

We say that D is e-acyclic if the directed graph $RG_D^{<e} = (V_T, A^{<e})$ contains no cycles, where $A^{<e} := \{(i, j) \in V_T \times V_T \mid i \neq j,\ \boldsymbol{p}^i \cdot \boldsymbol{x}^j < e\boldsymbol{p}^i \cdot \boldsymbol{x}^i\}$.

Proposition 3.6 ([8]): *Let $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t)\}_{t=1}^T$ be a dataset and let $R_T = [r_{ij}]$ be the ratio-based cost matrix defined in Definition 2.6. We denote by*

$$R := \{r_{ij} \in [0, 1] \mid i \neq j\}$$

the set of all off-diagonal entries of R_T, corresponding to pairwise cost ratios. Then

$$ACEI(D) = \max\left\{e \in R \mid RG_D^{<e} \text{ is acyclic}\right\} = \min\left\{e \in R \mid RG_D^{\leq e} \text{ contains a directed cycle}\right\}.$$

Equivalently,

$$\mathrm{ACEI}(D) \; = \; \min_{C} \; \max_{(i,j)\in C} r_{ij},$$

where the minimum is taken over all directed cycles C in $RG_D^{\leq 1}$.

This result follows the structural perspective of Proposition 12 of Lanier and Quah [8], which characterizes the critical efficiency as the largest e for which the strict graph $RG_D^{<e}$ is acyclic. In a finite dataset, the emergence of cycles occurs only at finitely many ratio values $e \in R$, so the formulation based on the supremum of all acyclic thresholds and the one based on the minimum value at which a cycle first appears are equivalent. The bottleneck characterization $\min_C \max_{(i,j)\in C} r_{ij}$ therefore provides a constructive way to compute or bound the ACEI.

4 Critical components

In this section, we propose the concept of *critical components*— arcs within strongly connected components that are iteratively removed to render the graph acyclic, with the aim of keeping the total violation length as small as possible. The formal connection to Dean and Martin's Minimum Cost Index (MCI) is discussed in Section 5. We begin by illustrating the procedure through examples.

We denote by $A_{SCC}(G)$ the set of arcs in the strongly connected components of a directed graph G. We first decompose the graph $RG_D^{\leq 1}$ into its strongly connected components. If all arcs $(i, j) \in A_{SCC}(RG_D^{\leq 1})$ satisfy $\dfrac{\boldsymbol{p}^i \cdot \boldsymbol{x}^j}{\boldsymbol{p}^i \cdot \boldsymbol{x}^i} = 1$ (equivalently, $\boldsymbol{p}^i \cdot (\boldsymbol{x}^j - \boldsymbol{x}^i) = 0$), then GARP is satisfied. Otherwise, there exists at least one arc with $\dfrac{\boldsymbol{p}^i \cdot \boldsymbol{x}^j}{\boldsymbol{p}^i \cdot \boldsymbol{x}^i} < 1$, indicating a violation.

Definition 4.1 (**Critical Component**): *Let $G = (V, A)$ be a directed graph representing a ratio-based revealed preference structure. An arc $(i, j) \in A$ is called a critical component if it satisfies the following conditions:*

1. *The arc (i, j) is a member of the strongly connected components, i.e., $(i, j) \in A_{SCC}(G)$.*
2. *Among all such arcs, it has the maximum relative cost ratio:*

$$\frac{\boldsymbol{p}^i \cdot \boldsymbol{x}^j}{\boldsymbol{p}^i \cdot \boldsymbol{x}^i}.$$

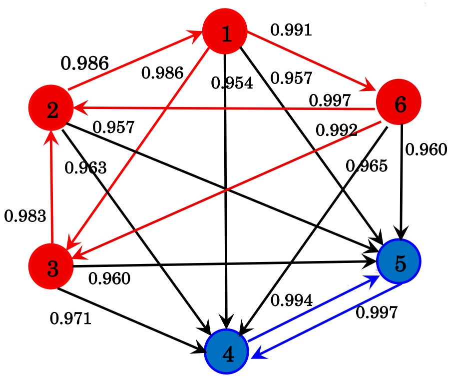
3. If multiple arcs attain the maximum ratio, the tie is resolved by:

- *Selecting the arc with the largest value of* $\boldsymbol{p}^i \cdot (\boldsymbol{x}^j - \boldsymbol{x}^i)$*;*
- *If still tied, the arc with the smallest tail index* i*;*
- *If still tied, the arc with the smallest head index* j*.*

The selected arc is referred to as a critical component and is denoted $a_n = (i_n, j_n)$ *at the* n*-th iteration.* □

The following example demonstrates how a critical component is selected according to the above definition.

Example 4.2: *Consider the dataset used in Example 2.8,*

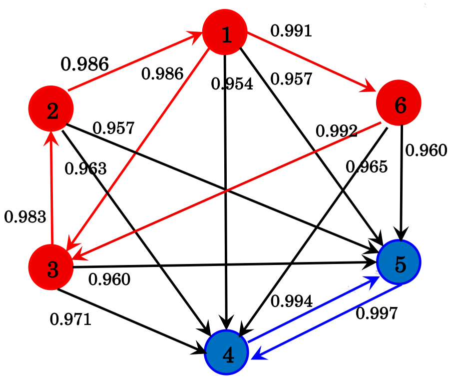


The strongly connected components of $RG_D^{\leq 1}$ *contain the following arcs:*

$$A_{SCC}(RG_D^{\leq 1}) = \{(1,3),(1,6),(2,1),(3,2),(4,5),(5,4),(6,2),(6,3)\}.$$

Among these arcs, the highest ratio is achieved by $(6,2)$*:*

$$\frac{\boldsymbol{p}^6 \cdot \boldsymbol{x}^2}{\boldsymbol{p}^6 \cdot \boldsymbol{x}^6} = \frac{398}{399} \approx 0.9975.$$



Therefore, $(6,2)$ *is selected as the first critical component, i.e.,* $a_1 = (6,2)$. $\square$

Next, decompose $RG_D^{\leq 1} - a_1$ into strongly connected components. If $RG_D^{\leq 1} - a_1$ satisfies GARP, then $a_1 = (i_1, j_1)$ is the unique critical component of $RG_D^{\leq 1}$, and the algorithm terminates. Otherwise, choose one arc $a_2 = (i_2, j_2) \in A_{SCC}(RG_D^{\leq 1} - a_1)$ based on the same rules as when we chose a_1. Repeat these steps until $RG_D^{\leq 1} - \bigcup_{h=1}^{n} \{a_h\}$ satisfies GARP. Then critical components of $RG_D^{\leq 1}$ are $\{a_1, \ldots, a_n\}$.

Hereafter, we assume the nontrivial case that $RG_D^{\leq 1}$ contains at least one arc with $r_{ij} < 1$ inside a strongly connected component; equivalently, the data matrix D_T has at least one negative entry $d_{ij} < 0$ with $(i,j) \in A_{SCC}(G_D^{\leq 0})$. Otherwise, Step 1 applies and the algorithm terminates with $e_{\max} = 1$.

Algorithm: Critical Component Selection Let $G := RG_D^{\leq 1}$ be the initial ratio-based revealed preference graph.

1. Decompose G into its strongly connected components. If all arcs $(i,j) \in A_{SCC}(G)$ satisfy $\dfrac{\boldsymbol{p}^i \cdot \boldsymbol{x}^j}{\boldsymbol{p}^i \cdot \boldsymbol{x}^i} = 1$, set $e_{\max} := 1$ and terminate.
2. Initialize: $n := 1$; $L := 0$ (total accumulated violation length); $e_{\max} := 1$
3. Repeat the following steps (a)–(h):
   (a) Select an arc $a_n = (i_n, j_n) \in A_{SCC}(G)$ with the largest ratio
   $$w_n := \frac{\boldsymbol{p}^{i_n} \cdot \boldsymbol{x}^{j_n}}{\boldsymbol{p}^{i_n} \cdot \boldsymbol{x}^{i_n}}$$
   (possibly $w_n = e_{\max}$; initially $e_{\max} = 1$) and apply the tie-breaking rules.
   (b) Update the current threshold: $e_{\max} := w_n$.
   (c) Define the violation length:
   $$\ell_n := \boldsymbol{p}^{i_n} \cdot (\boldsymbol{x}^{j_n} - \boldsymbol{x}^{i_n}).$$
   (d) Update graph: $G := G - a_n$.
   (e) Accumulate the cost: $L := L + |\ell_n|$
   (f) Recompute strongly connected components.
   (g) *Termination test:* If every SCC of G contains only arcs with ratio $r_{ij} = e_{\max}$ (vacuously true if $A_{SCC}(G) = \varnothing$), terminate. *Equivalently (finite setting), terminate when* $RG_D^{< e_{\max}}$ *is acyclic.*
   (h) Otherwise, set $n \leftarrow n + 1$.

Proposition 4.3: (e-GARP and Strongly Connected Components)
Let $D = \{(p_t, x_t)\}_{t=1}^{T}$ *be a finite dataset, and let* $RG_D^{\leq e} = (V_T, A^{\leq e})$ *denote the e-weighted*

ratio-based revealed preference graph defined by the arc set

$$A^{\leq e} := \{(i,j) \in V_T \times V_T \mid r_{ij} := \frac{\boldsymbol{p}^i \cdot \boldsymbol{x}^j}{\boldsymbol{p}^i \cdot \boldsymbol{x}^i} \leq e\}.$$

(i) *If the dataset D satisfies e-GARP, then every strongly connected component of $RG_D^{\leq e}$ contains only arcs (i,j) with $r_{ij} = e$.*
(ii) *Conversely, if every strongly connected component of $RG_D^{\leq e}$ contains only arcs with $r_{ij} = e$, then the strict-threshold graph $RG_D^{<e}$ is acyclic.*

Proof: (i) follows from Proposition 3.4 (equality on any directed cycle). (ii) If $RG_D^{<e}$ had a directed cycle, the arcs on that cycle (all with $r_{ij} < e$) would form an SCC of $RG_D^{\leq e}$ containing $r_{ij} < e$, contradicting the premise. □

Remark. Note that if a strongly connected component consists of a single node, then the condition is vacuously satisfied, since there are no arcs in the component.

Acyclicity of $RG_D^{<e}$ does not imply that D satisfies e-GARP in full generality (Lanier & Quah, 2024, Prop. 12). We therefore compute $ACEI(D)$ via the acyclicity characterization in Proposition 3.5. The critical component selection algorithm removes arcs in descending order of their revealed efficiency ratio r_{ij}, until the resulting graph satisfies e-GARP. Let $e_{\max}$ denote the final threshold value at which the algorithm terminates. The following theorem guarantees that this value corresponds exactly to Afriat's Critical Efficiency Index.

Theorem 4.4: *Let $e_{\max}$ be the final efficiency level recorded by the algorithm. Then $e_{\max} = ACEI(D)$.*

Proof: The algorithm iteratively removes arcs in descending order of their ratio r_{ij}, retaining only those with $r_{ij} \leq e_{\max}$. By construction, after the final deletion, every strongly connected component of the current graph G consists only of arcs with $r_{ij} = e_{\max}$ (vacuously true if $A_{SCC}(G) = \emptyset$). By Proposition 4.3(ii), the strict-threshold graph $RG_D^{<e_{\max}}$ is acyclic.

Since the candidate ratios form a finite set, this means that $RG_D^{\leq e}$ is acyclic for all $e < e_{\max}$, and any cycle in $RG_D^{\leq e_{\max}}$ (if present) must consist solely of arcs with $r_{ij} = e_{\max}$.

Hence $e_{\max}$ is the largest threshold for which acyclicity holds in the sense of Proposition 3.5. Therefore we conclude that

$$ACEI(D) = \max\{\, e \in R : \; RG_D^{<e} \text{ is acyclic} \,\} = e_{\max}.$$

□

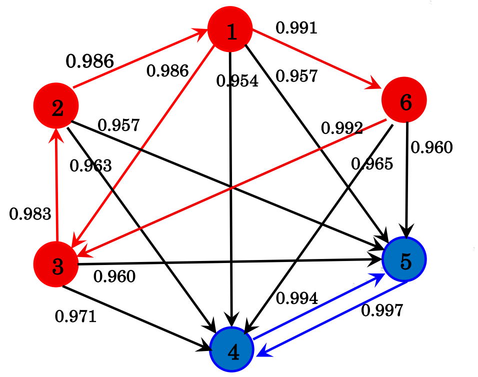
To illustrate how the critical component selection algorithm works in practice and how the value $e_{\max}$ is derived step by step, we present the following example based on the same dataset used earlier.

Example 4.5: *Continuing from Example 4.2, we consider again the dataset used in Example 2.8. The following table lists the arcs in the strongly connected components of $RG_D^{\leq 1}$, along* *with their efficiency ratios and violation lengths.*

arc (i,j)	$(6,2)$	$(5,4)$	$(4,5)$	$(6,3)$	$(1,6)$	$(2,1)$	$(1,3)$	$\cdots$	$(1,4)$
$\boldsymbol{p}^i \cdot \boldsymbol{x}^j$	398	356	359	396	345	345	343	$\cdots$	332
$\boldsymbol{p}^i \cdot \boldsymbol{x}^i$	399	357	361	399	348	350	348	$\cdots$	348
ratio	0.9975	0.9972	0.9945	0.9925	0.9914	0.9857	0.9856	$\cdots$	0.9540
length	-1	-1	-2	-3	-3	-5	-5	$\cdots$	-16

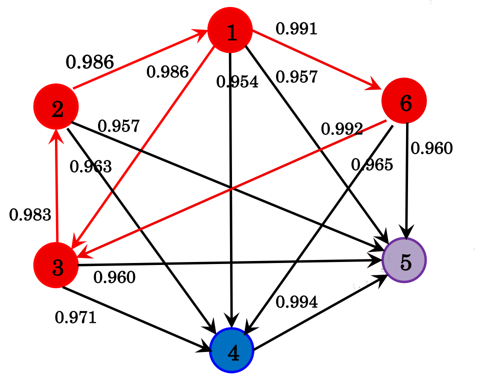
According to the critical component selection algorithm, we iteratively remove the arc with the largest ratio among the current set of arcs contained in strongly connected components. As mentioned in Example 4.2, $a_1 = (6,2)$*. After removal,*

$$A_{SCC}(RG_D^{\leq 1}) - \{a_1\} = \{(1,3),(1,6),(2,1),(3,2),(4,5),(5,4),(6,3)\}.$$



Among these arcs, the arc $(5,4)$ *has the highest ratio:* $\dfrac{\boldsymbol{p}^5 \cdot \boldsymbol{x}^4}{\boldsymbol{p}^5 \cdot \boldsymbol{x}^5} = \frac{356}{357} \approx 0.9972$*. Set* $a_2 = (5,4)$*.*

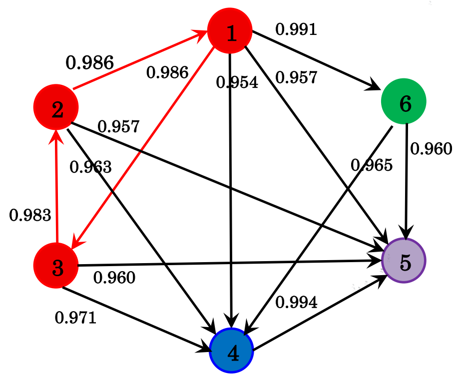
$$A_{SCC}(RG_D^{\leq 1}) - \{a_1, a_2\} = \{(1,3),(1,6),(2,1),(3,2),(6,3)\}.$$



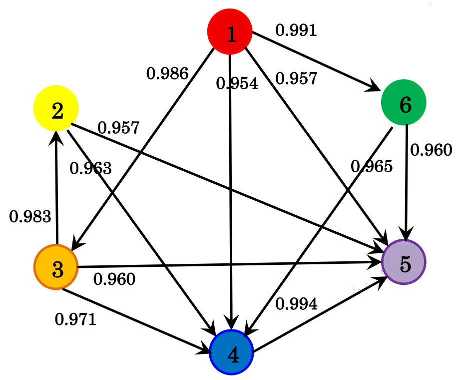
Among these arcs, the arc $(6,3)$ *has the highest ratio. Hence* $a_3 = (6,3)$*.*

After removing (6,2) and (6,3), node 6 becomes disconnected from any cycle.

$$A_{SCC}(RG_D^{\leq 1}) - \{a_1, a_2, a_3\} = \{(1,3), (2,1), (3,2)\}.$$



As a result, the arc (1,6) is no longer part of any strongly connected component. Among these arcs, the arc $(2,1)$ *has the highest ratio. Hence* $a_4 = (2,1)$.



Since $A_{SCC}(RG_D^{\leq 1}) - \{a_1, a_2, a_3, a_4\} = \emptyset$, *the graph* $RG_D^{<e_{\max}}$ *is acyclic. Hence, by Proposition 3.6,* $e_{\max} = r_{21} = 345/350 \approx 0.9857$. *The total violation length is* $L = 10$, *which will be used in the next section.* □

5 Goodness-of-fit indices and approximation measures

Dean and Martin [3] proposed a goodness-of-fit measure based on Afriat's cyclical consistency.

Definition 5.1: (*Dean and Martin's minimum cost index*)
For a dataset $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t) \mid t = 1, \dots, T\}$ *the minimum cost index* (MCI) *is defined as follows:*

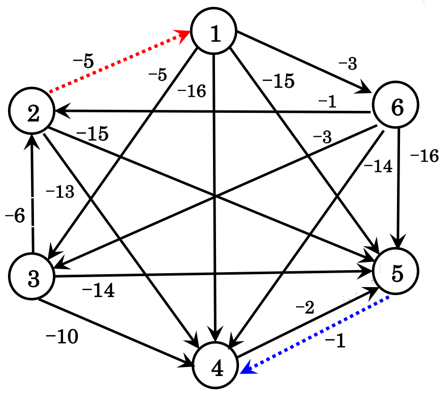
$$\mathrm{MCI}(D) = \min_{A' \subseteq A^{\leq 0}} \left\{ \frac{SA'}{\sum_{t=1}^{T} \boldsymbol{p}^t \cdot \boldsymbol{x}^t} \;\middle|\; G' = (V_T, A^{\leq 0} \setminus A') \text{ contains no directed cycle} \right\}$$

where $SA' = \sum_{(i,j) \in A'} \boldsymbol{p}^i \cdot (\boldsymbol{x}^i - \boldsymbol{x}^j)$ *and* $A^{\leq 0}$ *is the arc set of* $G_D^{\leq 0}$.

That is, MCI measures the minimum total cost of removing arcs that violate GARP in order to obtain an acyclic graph.

To illustrate the implications of MCI, we visualize the revealed preference graph using arc lengths defined as $\boldsymbol{p}^i \cdot (\boldsymbol{x}^j - \boldsymbol{x}^i)$, instead of the ratio-based weights.

Example 5.2: *Consider the dataset used in Example 2.4. If we delete* $\{(5,4), (2,1)\}$*, then* $G^{\leq 0}$ *does not contain a cycle, since every directed cycle in* $G^{\leq 0}$ *uses at least one of these arcs.*



Hence $\mathrm{MCI}(D) = \dfrac{1+5}{348+350+346+361+357+399} = \dfrac{6}{2161} \approx 0.00278.$ □

MCI can be considered to effectively use information on the length of arcs, that need to be modified minimally to satisfy GARP. However, it is known that computing MCI is NP-hard ([11]). In this section, we introduce a goodness-of-fit measure for GARP that uses the violation lengths of the arcs selected as critical components. Denote by L the sum of the absolute violation lengths of these arcs. We focus on L as the numerator. This value L can be obtained using the algorithm in the previous section. The index we propose uses the same denominator as MCI.

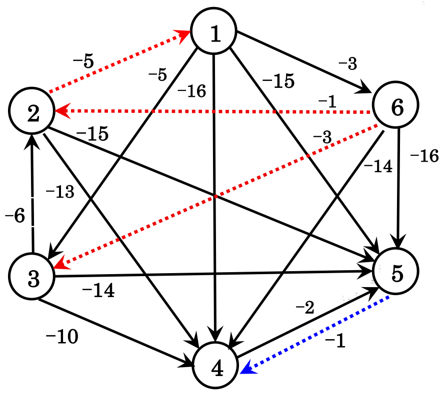
Definition 5.3: *For a dataset* $D = \{(\boldsymbol{p}^t, \boldsymbol{x}^t) \mid t = 1, \dots, T\}$*, Index(D) is defined as follows:*

$$\mathrm{Index}(D) = \frac{L}{\sum_{t \in T} \boldsymbol{p}^t \cdot \boldsymbol{x}^t}$$

where L is the total violation length, defined as the sum of the absolute violation lengths of the arcs selected as critical components.

$\text{Index}(D) = 0$ iff $G_D^{\leq 0}$ is acyclic (i.e., GARP holds). The index is scale-invariant in prices: scaling all $\boldsymbol{p}^t$ by $\alpha > 0$ scales both L and the denominator by α.

Example 5.4: *Consider the dataset used in Example 2.4. The arcs removed as critical components are* $\{(6,2),(5,4),(6,3),(2,1)\}$.



$$\text{Index}(D) = \frac{1+1+3+5}{2161} = \frac{10}{2161} \approx 0.00463. \qquad \square$$

While MCI directly seeks a minimum-cost subset of arcs whose removal restores GARP, it is known to be NP-hard to compute. On the other hand, Index(D) offers a tractable, approximation-based alternative: it measures the total violation cost incurred by removing the most critical arcs, as determined by their revealed efficiency ratios. This makes Index particularly useful in settings where computational efficiency is a priority, even if optimality is not guaranteed.

6 Conclusion

This paper proposed a graph-based algorithm for computing Afriat's Critical Efficiency Index (ACEI), in which arcs with high revealed efficiency ratios are incrementally removed from strongly connected components of the revealed preference graph. This method yields the critical efficiency threshold directly and provides an interpretable decomposition of revealed preference violations. The algorithm identifies a sequence of critical components—arcs within strongly connected components containing at least one arc with a revealed preference violation—and records the maximum expenditure ratio among these arcs. This maximum ratio reflects the most severe local violation and precisely corresponds to the ACEI.

Unlike existing approaches such as Varian's approximation method [17] and the exact algorithm by Smeulders et al. [13], which evaluate rationalizability by globally enumerating efficiency thresholds, our approach focuses on the internal structure of strongly connected components to avoid such exhaustive enumeration. Moreover, whereas Polisson and Quah [10] establish a theoretical equivalence between e-GARP and cost-rationalizability, our method attempts to identify the specific structural violations contributing to irrationality.

In this respect, our approach bridges the gap between computational efficiency and structural interpretability, and may serve as a practical tool for both evaluating and visualizing the degree of inconsistency in empirical consumption data. Furthermore, by utilizing the violation lengths of arcs selected as critical components, our method is expected to serve as a natural approximation to the Minimum Cost Index (MCI). While the MCI requires solving a global-optimization problem, our approach instead uses the total violation length L accumulated during the critical component deletion process as a tractable surrogate. This algorithm not only identifies the threshold level of inefficiency (ACEI) but also captures the cumulative cost of corrections required to restore rationalizability.

Acknowledgments

The author is deeply grateful to Satoru Fujishige for introducing this research topic and for many illuminating discussions on the graph-theoretic aspects of revealed preference, especially the role of strongly connected components. Any remaining errors are solely the author's responsibility.

References

[1] S. N. Afriat: The Construction of Utility Function from Expenditure Data. *International Economic Review* **8**(1) (1967) 67-77.

[2] S. N. Afriat: On a system of inequalities in demand analysis: An extension of the Classical Method. *International Economic Review* **14** (2) (1973) 460-472.

[3] M. Dean and D. Martin: Measuring Rationality with the Minimum Cost of Revealed Preference Violations. *Review of Economics and Statistics* **98**(3) (2016) 524-534.

[4] P Dziewulski, J. Lanier, and J. K. H. Quah: Revealed preference and revealed preference cycles: A survey. *Journal of Mathematical Economics* **113** (2024) 1-10.

[5] F. Forges and V. Iehlè Afriat's theorem for indivisible goods. *Journal of Mathematical Economics* **54** (2014) 1-6.

[6] S. Fujishige: *Submodular Functions and Optimization* : Second Edition (North-Holland, Amsterdam, 2005).

[7] S. Fujishige and Z. Yang: On revealed preference and indivisibilities. *Modern Economy* **3**(6) (2012) 752-758.

[8] J. Lanier and J. K.-H. Quah: Goodness-of-fit and utility estimation: what's possible and what's not. https://arxiv.org/abs/2405.08464 arXiv (2024) preprint

[9] T. Naitoh: New Goodness-of-Fit Measures for GARP and Critical Arcs. *The Hikone Ronso* **444** (Summer) (2025) 4-23. [Shiga University Faculty of Economics]

[10] M. Polisson and J. K.-H. Quah: Rationalizability, Cost-Rationalizability, and Afriat's Efficiency Index https://arxiv.org/abs/2406.10136 arXiv (2024) preprint

[11] K. Shiozawa: Note on the goodness-of-fit measure for GARP; NP-hardness of minimum cost index. *Economics Bulletin* **35**(4) (2015) 2455-2461.

[12] K. Shiozawa: Revealed preference test and shortest path problem; graph theoretic structure of the rationalizability test. *Journal of Mathematical Economics* **67** (2016) 38-48.

[13] B. Smeulders, F. C. R. Spieksma, L. Cherchye, and B. De Rock: Goodness-of-fit measures for Revealed Preference Tests: Complexity Results and Algorithms. *ACM Transactions on Economics and Computation archive* **2**(1), Article 3 (2014).

[14] B. Smeulders, Y. Crama, and F. C. R. Spieksma: Revealed preference theory: An algorithmic outlook. *European Journal of Operational Research* **272**(3) (2019) 803-815.

[15] F. Talla Nobibon, B. Smeulders, and F. C. R. Spieksma: A note on testing axioms of revealed preference. *Journal of Optimization Theory and Applications* **166**(3) (2015) 1063-1070.

[16] R. Tarjan: Depth-first search and linear graph algorithms. *SIAM Journal on Computing* **1**(2) (1972) 146-160.

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

[18] H. R. Varian: *Microeconomic Analysis* : Third Edition (W. W. Norton & Company, New York, 1992).