# Goodness of fit in optimizing consumer's model

Alcantud, José Carlos R. and Matos, Daniel L. and Palmero, Carlos R.

Universidad de Salamanca, Instituto Diego de Praves, Universidad de Valladolid

29 September 2009

# Goodness of Fit in Optimizing Consumer's Model

José Carlos R. Alcantud[a], Daniel L. Matos[b], Carlos R. Palmero[c,*]

[a] *Facultad de Economía y Empresa, Universidad de Salamanca, E-37008 Salamanca, Spain*

[b] *Instituto Diego de Praves, E-47005 Valladolid, Spain*

[c] *Facultad de Ciencias Económicas y Empresariales, Universidad de Valladolid, E-47011 Valladolid, Spain*

**Abstract**

We provide two new indices of efficiency for determining the degree of coherence in an agent's consumption decisions. We analyze to which extent they improve the efficiency displayed by Varian's [16] index. We report on the results of a Montecarlo experiment that confirms that *strict* improvements of Varian's vector-index appear on a regular basis.

## 1. Introduction

In the theory of consumer behavior many non-parametric tests are designed to check for an agent's optimizing behavior without any functional restriction on the demand. For finite sets of data, Afriat [1] and Varian [15] outstand among a large literature. After Afriat-Varian it is known that violations of the Generalized Axiom of Revealed Preference (GARP) mean violations of the usual neoclassical model of demand choice. In experimental economics this has permitted to find contradictions to the standard demand model (cf., e.g., Battalio et. al. [3], Koo [9], Sippel [12], and Mattei [10, 11]).

Nonetheless, apart from intrinsic lack of rationality many other factors may be a cause for inconsistency with the *exact* optimizing model: measurement errors, non-observability of all consumption choices, rationing in the available quantities, time-evolving preferences, .... Afriat's [2] seminal contribution investigates the degree of coherence in a finite list of demand observations. by using a *uniform* bound for goodness-of-fit of the agent's behavior. Varian [16] proposes the use of *vector-indices* instead, a line followed by e.g., Famulari [6], Gross [7], Swofford and Withney [13] among others.

*Corresponding author

*URL:* `jcr@usal.es` (José Carlos R. Alcantud), `ldlopez@educa.jcyl.es` (Daniel L. Matos), `cpalmero@eco.uva.es` (Carlos R. Palmero )

In this work we provide two new vector-indices of efficiency that allow to determine the degree of coherence in an agent's consumption decisions. Our construction is algorithmic and computationally efficient. These indices improve the performance of Varian's [16] index (further studied by Tsur [14]), in the following sense. Comparing uniform bounds is trivial: the higher, the closer to rational. Comparing vector-indices can be done through norms as proposed by Varian, but the choice of the norm can produce different conclusions. Here we take a more basic and less arguable position: we request that for all sets of data our vector-indices provide Pareto-improvements of Varian's, that improvement being strict for a non-negligible part of the problems. We perform a Montecarlo experiment to check how Tsur's test performs against our proposals, confirming that strict improvements of Varian's vector-index appear on a regular basis.

We organize our research as follows. In Section 2 we fix the notation and comment on the literature briefly. Section 3 gives our main results, and Section 4 reports on the conclusions from our Montecarlo experiment. Section 5 contains the proofs of our results.

## 2. Definitions and preliminary results

The pure theory of consumer's behavior aims at studying the structure of choices among bundles of goods by a rational agent, when he faces different price-income situations. We fix $k > 0$ goods. The agent can select non-negative amounts of every good. A demand vector $\bar{x} = (x_1, \cdots, x_k)$ is a $k$-dimensional vector whose $i$-th component captures the amount of good $i$ that the agent demands. Market prices are captured by $\bar{p} = (p_1, \cdots, p_k)$. Henceforth we assume that all prices are positive (we discard freely available goods from the analysis), that is, $\bar{p} \in \mathbb{R}^k_{++}$.

We fix a *finite* set of demand data, namely $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$. Each $(\bar{p}_t, \bar{x}_t) \in \mathbb{R}^k_{++} \times \mathbb{R}^k_{+}$ means that $\bar{x}^t$ has been demanded at *normalized* prices $\bar{p}_t$, thus $\bar{p}_t\bar{x}_t = 1$ throughout.

The possible rationality of this series of observations relates to the fulfilment of behavioral postulates that are typically expressed in terms of the following concepts.

**Definition 1.** *The consumer directly reveals that prefers $\bar{x}_t$ to $\bar{x}_s$, denoted $\bar{x}_t \; R^0 \; \bar{x}_s$, if $\bar{p}_t\bar{x}_t \geqslant \bar{p}_t\bar{x}_s$. He reveals that prefers $\bar{x}_t$ to $\bar{x}_s$, denoted $\bar{x}_t \; R \; \bar{x}_s$, if for some suitable bundles we have $\bar{x}_t \; R^0 \; \bar{x}_{t_1} \cdots \; R^0 \; \bar{x}_{t_k} \; R^0 \; \bar{x}_s$.*

The following postulate is necessary and sufficient for our list of demand observations to be generated by an agent through optimization of a non-satiated, continuous, concave and monotone utility (cf., Varian [15]):

**Definition 2.** *The set of consumption data $\{(p^t, x^t)\}_{t=1}^n$ agrees with the Generalized*

*Axiom of Revealed Preference (also GARP) if for each pair of observations $i, j$ such that $x^i R x^j$ one has $p^j x^j \leqslant p^j x^i$.*

GARP is an *exact* test of the rationality of demand choices. As has been mentioned, experimental studies confirm that relevant series of data do not pass such test. Does this mean that the data must be considered as fully irrational? It is agreed that the answer is no, and thus different models attempt to account for some inaccuracies in the specification of the observations that permit to fit the data into *approximate* rationality.

In the non-parametric approach that we follow in this paper the analyst keeps control of the inefficiency of the agent as an optimizer by introducing either a global parameter for the problem (namely, a number $0 \leqslant e \leqslant 1$), or a vector-index of efficiency $\bar{e} = (e_t)_{t=1}^n$ (with $0 \leqslant e_t \leqslant 1, \ \forall t$). In the latter, richer instance, each $e_j$ is interpreted as the "level of efficiency" of the agent in the $j$-th budget situation. Afriat's [2, pp. 467-8] index [1] belongs to the first class, while Varian's [16] proposal $\boldsymbol{i}$ belongs to the second one. A formal definition of revelation subject to a vector-index follows:

**Definition 3.** *An index-mapping $\boldsymbol{\xi}$ is a procedure that with each series of $m$ demand observations assigns a vector-index $\bar{\xi} = (\xi_1, \ldots, \xi_m)$, $0 \leqslant \xi_t \leqslant 1 \ \forall t$. Formally speaking: $\boldsymbol{\xi}(\{(\bar{q}_t, \bar{y}_t)\}_{t=1}^m) = \bar{\xi} = (\xi_1, \ldots, \xi_m)$, for each $\{(\bar{q}_t, \bar{y}_t)\}_{t=1}^m$.*

*Given $\bar{e} = (e_t)_{t=1}^n$ with $0 \leqslant e_t \leqslant 1 \ \forall t$, we say that the consumer directly reveals under the vector-index $\bar{e}$ that prefers $\bar{x}_t$ to $\bar{x}_s$, denoted $\bar{x}_t \ R^0(\bar{e}) \ \bar{x}_s$, if $e_t \bar{p}_t \bar{x}_t \geqslant \bar{p}_t \bar{x}_s$. He reveals under the vector-index $\bar{e}$ that prefers $\bar{x}_t$ to $\bar{x}_s$, denoted $\bar{x}_t \ R(\bar{e}) \ \bar{x}_s$, if there are $\bar{x}_{t_1}, \ldots, \bar{x}_{t_k}$ such that $\bar{x}_t \ R^0(\bar{e}) \ \bar{x}_{t_1} \cdots \ R^0(\bar{e}) \ \bar{x}_{t_k} \ R^0(\bar{e}) \ \bar{x}_s$.*

From the point of view of intuition, Varian [16] explains that if for example $e_t = 0.9$ we only count bundles $\bar{x}_s$ whose cost is less than 90% of the price paid for $\bar{x}_t$ as candidates for being revealed "worse" than choice $\bar{x}_t$. It is intuitively clear that this imposes less restrictions on the conditions for optimization that must be verified [2] and therefore yields a non-exact concept of rationality in the form of the next postulate (cf., [16]):

**Definition 4.** *The demand data $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ agree with the* Generalized Axiom of Revealed Preference *under $\bar{e}$, henceforth $GARP(\bar{e})$, if for each pair of observations $t, s$ it is true that $\bar{x}_t \ R(\bar{e}) \ \bar{x}_s$ entails $e_t \bar{p}_t \bar{x}_t \leqslant \bar{p}_t \bar{x}_s$.*

[1] It represents the minimal percentage of money (in unit terms) that the agent can waste in *every* decision so that the data fit the standard optimization model. Thus its main drawback is that it does not inform of which observations cause the possible lack of consistency. It is worth mentioning that Houtmann and Maks [8] propose an efficient algorithm to compute Afriat's index that is based on the bisection method.

[2] Formally: if we let $\bar{0} = (0, \ldots, 0)$ and $\bar{1} = (1, \ldots, 1)$ then for every vector $\bar{e} \in [\bar{0}, \bar{1}]$ the relations $R^0(\bar{e}) \subseteq R^0(\bar{1}) = R^0$ and $R(\bar{e}) \subseteq R(\bar{1}) = R$ must hold true.

For our purposes only index-mappings that perform well with respect to such approximately rational behavior are worth studying. This can be achieved in two related senses.

**Definition 5.** *The index-mapping* $\boldsymbol{\xi}$ *is efficient (resp., strongly efficient) if for each finite set of data* $\{(\bar{q}_t, \bar{y}_t)\}_{t=1}^m$ *with associated* $\bar{\xi} = \boldsymbol{\xi}(\{(\bar{q}_t, \bar{y}_t)\}_{t=1}^m)$ *the following holds* [3]*:*

$$\{(\bar{q}_t, \bar{y}_t)\}_{t=1}^m \text{ verifies } GARP(\bar{\zeta}),\ \forall \bar{\zeta} \in C_m \text{ with } \bar{\zeta} \ll \bar{\xi} \text{ (resp., with } \bar{\zeta} \leqslant \bar{\xi}\text{)}$$

We intend to build on Varian's proposal $\boldsymbol{i}$. This is axiomatically based on Samuelson's overcompensation function, which yields interesting economic insights. We do not need such interpretation here, but rather its alternative algorithmic construction [4]:

**Algorithm 1.** Pseudo-code for computing Varian's index

**Input:** Cost matrix associated with$\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$

1. `begin`
2. `for` $j := 1$ `to n do` $e(j) \leftarrow 1$
3. `for` $i, j := 1$ `to n if` $C(i,i) \geqslant C(i,j)$ `then` $R^0(i,j) \leftarrow 1$ `else` $R^0(i,j) \leftarrow 0$
4. `Compute the transitive closure` $R$ `of` $R^0$ `and check for GARP`
5. `if` $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ `does not agree with GARP then`
   `for` $j := 1$ `to n do` $e(j) \leftarrow min\left\{\frac{C(j,i)}{C(j,j)} : \bar{x}_i R \bar{x}_j\right\}$
6. `return`

**Output:** $\bar{i} = \boldsymbol{i}(\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n)$

It is trivial that strong efficiency implies efficiency. The converse implication does not hold: Afriat's index-mapping is efficient but not strongly efficient [5]. Varian's index-mapping is strongly efficient. In what follows all index-mapping are at least efficient.

For series of demand observations that do not agree with the exact optimizing model, Varian proposed to use the Euclidean norm of $\bar{1} - \bar{i}$ as a measure of the goodness-of-fit of the data with such standard model. He argued that there is a positive correlation between the degree of coherence in the decisions and $\|\bar{i}\|$. Alternatively, in order to compare the

[3]By $C_m$ we mean the unit $m$-cube. Also, $(a_1, ..., a_m) \ll (b_1, ..., b_m)$ means $a_i < b_i$ for each $i$, and $(a_1, ..., a_m) < (b_1, ..., b_m)$ means $a_i \leqslant b_i$ for each $i$ but $(a_1, ..., a_m) \neq (b_1, ..., b_m)$.

[4]For any $\{(\bar{q}_t, \bar{y}_t)\}_{t=1}^m$, its cost matrix is $C = (c_{ij})_{i,j \leqslant m}$ defined by $c_{ij} = \bar{q}_i \bar{y}_j$ for each $i, j$.

[5]Consider the following counterexample: $\bar{x}_1 = (8, 1, 8), \bar{x}_2 = (5, 5, 6), \bar{x}_3 = (5, 6, 5), \bar{x}_4 = (8, 8, 1)$, for respective prices $\bar{p}_1 = (1, 1, 0.5), \bar{p}_2 = (1, 1, 1.5), \bar{p}_3 = (1, 0.5, 1), \bar{p}_4 = (1, 2, 2)$. Afriat's index is $\bar{e}^* = (1, 1, 1, 1)$ and because the data violate GARP we can not have strong efficiency.

performance of different index-mappings with regard to efficiency we introduce Definition 6 below (implications as to comparing in Varian's sense are straightforward).

**Definition 6.** *Let* $\boldsymbol{\xi}$ *and* $\boldsymbol{\zeta}$ *be index-mappings. We say that* $\boldsymbol{\xi}$

1. *describes the demand behavior quasi-better than* $\boldsymbol{\zeta}$ *if for each* $\mathscr{D} = \{(\bar{q}_t, \bar{y}_t)\}_{t=1}^m$ *it is true that* $\boldsymbol{\xi}(\mathscr{D}) \geqslant \boldsymbol{\zeta}(\mathscr{D})$ *and at least for one such case the inequality is strict.*
2. *describes the demand behavior better than* $\boldsymbol{\zeta}$ *if for each* $\mathscr{D} = \{(\bar{q}_t, \bar{y}_t)\}_{t=1}^m$ *it is true that* $\boldsymbol{\xi}(\mathscr{D}) > \boldsymbol{\zeta}(\mathscr{D})$.

## 3. Main theoretical results

Next we put forward two strongly efficient index-mappings that describe the behavior of an agent quasi-better than Varian's index-mapping (proofs are given in Section 5). They are defined algorithmically and yield the output in polinomial time.

**Definition 7 (**IM1). *Let* $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ *be demand data and* $\{\bar{\upsilon}_k\}_{k=0}^{n-1}$ *the family of n-vectors* $\bar{\upsilon}_k = (\upsilon_{k1}, \cdots, \upsilon_{kn}) \in \mathbb{R}_{++}^n$ *associated with it through the following recurrent equation* [6] *initialized at* $\bar{\upsilon}_0 = (1, \cdots, 1)$*:* $\bar{\upsilon}_k = \bar{\upsilon}_{k-1} \odot \bar{\upsilon}_k^*$*, where* $\bar{\upsilon}_k^* = (\upsilon_{k1}^*, \cdots, \upsilon_{kn}^*)$ *is a vector in* $\mathbb{R}_{++}^n$ *such that* $\upsilon_{kt}^* = max\left\{\frac{\bar{p}_t\bar{x}_s}{\upsilon_{tk-1}\bar{p}_t\bar{x}_t} < 1 : \bar{x}_s R(\bar{\upsilon}_{k-1})\bar{x}_t\right\}$ *if* $\bar{x}_t$ *violates* $GARP(\bar{\upsilon}_{k-1})$ *and* $\upsilon_{kt}^* = 1$ *otherwise.*

*We denote by* $\bar{\upsilon}$ *the vector-index* $\bar{\upsilon} = (\upsilon_1, \cdots, \upsilon_n) \in \mathbb{R}_{++}^n$ *such that* $\upsilon_t = \upsilon_{(n-1)t}$ *for each observation* $t$.

**Algorithm 2.** Pseudo-code for computing IM1

1. `begin`
2. `for` $i, j := 1$ `to n do` $v(j) \leftarrow 1$`;` $C_v(i,j) \leftarrow C(i,j)$
3. `While` [ $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ `does not verify` $GARP(v)$ ]
4. `for` $j := 1$ `to n do` $C_v(j,j) \leftarrow C(j,j)$`;` $C_v(j,j) \leftarrow v(j) \cdot C(j,j)$
5. `for` $i, j := 1$ `to n if` $C_v(i,i) \geqslant C_v(i,j)$ `then` $R_v^0(i,j) \leftarrow 1$ `else` $R_v^0(i,j) \leftarrow 0$
6. `Compute the transitive closure` $R_v$ `of the relation` $R_v^0$
7. `for` $j := 1$ `to n do` $G_v(\bar{x}_j) \leftarrow \{\bar{x}_i : \bar{x}_i R_v \bar{x}_j$ `and` $C_v(j,j) > C_v(j,i)\}$
8. `for` $j := 1$ `to n if` $G_v(\bar{x}_j) \neq \varnothing$ `then` $v(j) \leftarrow v(j) \cdot max\left\{\frac{C(j,i)}{C(j,j)} < 1 : \bar{x}_i \in G_v(\bar{x}_j)\right\}$
9. `return`

[6] If $\bar{v}, \bar{w} \in \mathbb{R}^n$, we denote $\bar{z} = \bar{v} \odot \bar{w} \in \mathbb{R}^n$ when $z_t = v_t \cdot w_t$ for each $t$.

In order to define our second index-mapping we need an auxiliary definition. With $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ and a family $\{\bar{\xi}_k\}_{k\in I}$ of $n$-vectors $\bar{\xi}_k = (\xi_{k1}, \cdots, \xi_{kn}) \in \mathbb{R}^n_{++}$ we define the family $\{\bar{\zeta}_k\}_{k\in I}$ of $n$-vectors $\bar{\zeta}_k = (\zeta_{k1}, \cdots, \zeta_{kn}) \in \mathbb{R}^n_{++}$ associated with $\{\bar{\xi}_k\}_{k\in I}$ through: $\zeta_{kt} = max\left\{\frac{\bar{p}_t\bar{x}_s}{\xi_{tk-1}\bar{p}_t\bar{x}_t} < 1 : \bar{x}_s R(\bar{\xi}_{k-1})\bar{x}_t\right\}$ if $\bar{x}_t$ violates $GARP(\bar{\xi}_{k-1})$ and $\zeta_{kt} = 1$ otherwise.

**Definition 8** (IM2). *Let $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ be demand data and let $\{\bar{\vartheta}_k\}_{k=1}^{n^2-n}$ be a family of $n$-vectors $\bar{\vartheta}_k = (\vartheta_{k1}, \cdots, \vartheta_{kn}) \in \mathbb{R}^n_{++}$ obtained by the following recurrent equation initialized at $\bar{\vartheta}_1 = (1, \cdots, 1)$: $\bar{\vartheta}_{k+1} = \bar{\vartheta}_k \odot \bar{\vartheta}^*_k$, where $\{\bar{\zeta}_k\}_{k=1}^{n^2-n}$ is the family of $n$-vectors associated with $\{\bar{\vartheta}_k\}_{k=1}^{n^2-n}$ and $\bar{\vartheta}^*_k = (\vartheta^*_{k1}, \cdots, \vartheta^*_{kn}) \in \mathbb{R}^n_{++}$ is such that $\vartheta^*_{kt} = \zeta_{kt}$ if $\bar{x}_t$ is not consistent with $GARP(\bar{\vartheta}_{k-1})$, $\zeta_{kt} = \zeta^*_k$ and $\nexists \bar{x}_s$ con $s < t$ such that $\zeta_{ks} = \zeta^*_k$, where $\zeta^*_k = max\{\zeta_{ks} : \zeta_{ks} < 1\}$ and $\vartheta^*_{kt} = 1$, otherwise.*

*We denote by $\bar{\vartheta}$ the vector-index $\bar{\vartheta} = (\vartheta_1, \cdots, \vartheta_n) \in \mathbb{R}^n_{++}$ such that $\vartheta_t = \vartheta_{n(n-1)t}$ for each observation $t$.*

**Algorithm 3.** Pseudo-code for computing IM2

1. `begin`
2. `for` $i, j := 1$ `to n do` $\vartheta(j) \leftarrow 1;$ $C_\vartheta(i,j) \leftarrow C(i,j)$
3. `While` [ $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ `does not verify` $GARP(\vartheta)$ ]
4. `for` $j := 1$ `to n do` $C_\vartheta(j,j) \leftarrow C(j,j);$ $C_\vartheta(j,j) \leftarrow \vartheta(j) \cdot C(j,j)$
5. `for` $i, j := 1$ `to n if` $C_\vartheta(i,i) \geqslant C_\vartheta(i,j)$ `then` $R^0_\vartheta(i,j) \leftarrow 1$ `else` $R^0_\vartheta(i,j) \leftarrow 0$
6. `Compute the transitive closure` $R_\vartheta$ `of the relation` $R^0_\vartheta$
7. `for` $j := 1$ `to n do` $G_\vartheta(\bar{x}_j) \leftarrow \{\bar{x}_i : \bar{x}_i R_\vartheta \bar{x}_j$ `and` $C_\vartheta(j,j) > C_\vartheta(j,i)\}$
8. `for` $j := 1$ `to n if` $G_\vartheta(\bar{x}_j) \neq \varnothing$ `then` $Pert^*(j) \leftarrow max\left\{\frac{C(j,i)}{C(j,j)} < 1 : \bar{x}_i \in G_\vartheta(\bar{x}_j)\right\}$
9. $v^* \leftarrow max\{Pert^*(j) : G_\vartheta(\bar{x}_j) \neq \varnothing\};$ $\omega^* \leftarrow min\{j : \bar{x}_j \in G_\vartheta(\bar{x}_j) : Pert^*(j) = v^*\}$
10. `for` $j := 1$ `to n if` $(G_\vartheta(\bar{x}_j) \neq \varnothing; Pert^*(j) = v^*; j = \omega^*)$ `then` $\vartheta(j) \leftarrow v^* \cdot \vartheta(j)$
11. `return`

**Example 1.** *Let us consider the demand data with the following cost matrix:*

$$\begin{pmatrix} 1.00 & 1.25 & 0.91 & 1.50 \\ 0.67 & 1.00 & 0.85 & 0.95 \\ 0.60 & 0.80 & 1.00 & 1.25 \\ 0.40 & 0.70 & 0.75 & 1.00 \end{pmatrix}$$

*It can be checked that Varian's index is* $(0.91, 0.67, 0.6, 0.4)$*, IM1 is* $(0.91, 0.85, 0.6, 0.7)$ *and IM2 is* $(0.91, 0.67, 0.6, 1)$*.*

*The respective cost matrices associated with IM1 and IM2 are the following:*

$$\begin{pmatrix} \mathbf{0.91} & 1.25 & 0.91 & 1.50 \\ 0.67 & \mathbf{0.85} & 0.85 & 0.95 \\ 0.60 & 0.80 & \mathbf{0.60} & 1.25 \\ 0.40 & 0.70 & 0.75 & \mathbf{0.70} \end{pmatrix} \qquad \begin{pmatrix} \mathbf{0.91} & 1.25 & 0.91 & 1.50 \\ 0.67 & \mathbf{0.67} & 0.85 & 0.95 \\ 0.60 & 0.80 & \mathbf{0.60} & 1.25 \\ 0.40 & 0.70 & 0.75 & \mathbf{1.00} \end{pmatrix}$$

*By using the standard representation of binary relations by oriented graphs, we can visualize this information as in Figure 1. Here the transition cost from vertex (observation)* $i$ *to vertex* $j$ *is* $c_{ij} - c_{ii}$ *when this amount is lesser or equal than 0. Observe that originally there were cycles of negative length, which can not appear under the respective vector-indices because they stem from strongly efficient index-mappings.*

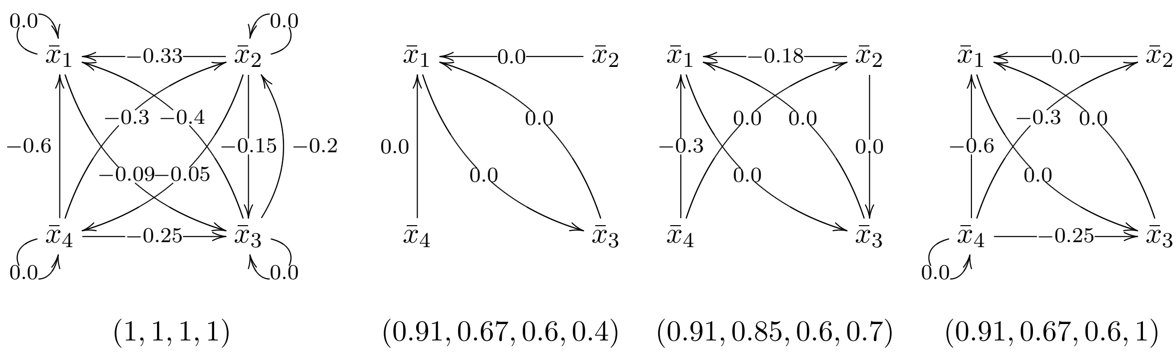


Figure 1: A representation of the respective relations $R^0$, $R^0(\bar{\imath})$, $R^0(\bar{\upsilon})$, $R^0(\bar{\vartheta})$

## 4. Montecarlo experiment

Next we run a Montecarlo experiment in order to analyze the goodness-of-fit of the optimizing model as well as the statistical significance of the violations of GARP. We generate the data using the *Almost Ideal Demand System (AIDS)* model by Deaton and Muellbauer [4, 5]. We introduce perturbations both in prices and demanded amounts through random variables with a normal logarithmic distribution. Our series of data have $n = 20$ observations with $k = 8$ goods.

Table 1 summarizes the results of our Montecarlo experiment. In all cases GARP is violated and $\bar{\vartheta} > \bar{\imath}$, although we must point out that examples are known where both indices coincide, and the inequality $\bar{\upsilon} > \bar{\imath}$ holds in 99.966% of the simulations. This information complements the main theoretical results about comparisons of the different indices. The proportion of simulations where IM2 is "better" than IM1, in the sense of Definition 6, is 1.82%, while the remaining cases end up in incomparability: 98.18% of the cases yield $\bar{\vartheta} \not> \bar{\upsilon}$ and $\bar{\upsilon} \not> \bar{\vartheta}$.

**Number of simulations:** 4500 (20 demand observations of 8 components)
**Average Afriat index:** 0.562
**Percentage simulations violating GARP:** 100.00%
**Percentage simulations with** $\bar{\vartheta} > \bar{\upsilon}$**:** 1.82%
**Percentage simulations with** $\bar{\vartheta} \not> \bar{\upsilon} \wedge \bar{\upsilon} \not> \bar{\vartheta}$**:** 98.18%

| **Index** | **Euclidean Norm (Average)** | **Percentage of simulations with** $\|\|\bar{\xi}\|\| > \|\|\bar{i}\|\|$ |
|---|---|---|
| $\bar{\upsilon}$ | 3.421927808 | 99.966% |
| $\bar{\vartheta}$ | 3.721796885 | 100.00% |

Table 1: Results of the Montecarlo experiment.

Another fact that must be stressed concerns the statistical significance of the violations of GARP. Our experiment produced 4 cases where Tsur's test concludes that such violation is structural while our indices do not concur to this rejection of rationality at suitable levels, but rather small random perturbations explain such apparent irrationality.

## 5. Appendix: proofs

Given $\{(\bar{p}_1, \bar{x}_1), ..., (\bar{p}_n, \bar{x}_n)\}$ arbitrary demand data, we say that $\bar{x}_t$ is *strongly inconsistent with* GARP if there is $\bar{x}_s$ such that $\bar{x}_s R \bar{x}_t$ and $\bar{p}_t \bar{x}_t > \bar{p}_t \bar{x}_s$. In terms of the description given in Example 1: $\bar{x}_t$ lies in a cycle of negative length where the transition cost from it is strictly negative.

**IM1 is strongly efficient:** Take $\{(\bar{p}_1, \bar{x}_1), ..., (\bar{p}_n, \bar{x}_n)\}$ arbitrary demand data that are inconsistent with GARP. We need to prove that its associated $\bar{\upsilon}$ verifies $GARP(\bar{\upsilon})$ By absurdum, assume that there are bundles $\bar{x}_t$ and $\bar{x}_{t_j}$ such that $\bar{x}_{t_j} R(\bar{\upsilon}) \bar{x}_t$ and $\upsilon_t \bar{p}_t \bar{x}_t > \bar{p}_t \bar{x}_{t_j}$. Since $\bar{x}_t$ is strongly inconsistent with GARP there are $\bar{x}_{t_1}, \cdots, \bar{x}_{t_{r_t}}$, $1 \leqslant r_t \leqslant n-1$ such that $\bar{x}_{t_i} R \bar{x}_t$ and $\bar{p}_t \bar{x}_t > \bar{p}_t \bar{x}_{t_i}$ for each $1 \leqslant i \leqslant r_t$. Without loss of generality we can reorder the prior demanded bundles in such way that $\bar{p}_t \bar{x}_{t_i} \geqslant \bar{p}_t \bar{x}_{t_{i+1}}$ for $1 \leqslant i \leqslant r_t - 1$. Besides, when $k \geqslant 1$ the definition of $\{\bar{\upsilon}_k\}_{k=0}^{n-1}$ entails $\bar{\upsilon}_k \geqslant \bar{\upsilon}$, thus $\bar{x}_{t_j} R(\bar{\upsilon}_k) \bar{x}_t$ and $\upsilon_{kt} \bar{p}_t \bar{x}_t > \bar{p}_t \bar{x}_{t_j}$. In particular, $\bar{x}_{t_j} R(\bar{\upsilon}_h) \bar{x}_t$ and $\upsilon_{ht} \bar{p}_t \bar{x}_t > \bar{p}_t \bar{x}_{t_j}$ for all $h \leqslant j$. From this we deduce $\upsilon_{jt} = \frac{\bar{p}_t\bar{x}_{jt}}{\bar{p}_t\bar{x}_t}$. Furthermore, $\bar{\upsilon}_j \geqslant \bar{\upsilon}$ implies $\upsilon_{jt}\bar{p}_t\bar{x}_t \geqslant \upsilon_t\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_{t_j}$, hence $\bar{p}_t\bar{x}_{t_j} \geqslant \upsilon_t\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_{t_j}$, which in turn implies $\bar{p}_t\bar{x}_{t_j} > \bar{p}_t\bar{x}_{t_j}$, an absurd conclusion that proves the claim. ■

**IM2 is strongly efficient:** Take $\{(\bar{p}_1, \bar{x}_1), ..., (\bar{p}_n, \bar{x}_n)\}$ arbitrary demand data that are inconsistent with GARP. We need to prove that its associated $\bar{\vartheta}$ verifies $GARP(\bar{\vartheta})$. By absurdum, assume that there are bundles $\bar{x}_t$ and $\bar{x}_{t_j}$ such that $\bar{x}_{t_j}R(\bar{\vartheta})\bar{x}_t$ and $\vartheta_t\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_{t_j}$. The fact that $R(\bar{\vartheta})$ is a subrelation of $R$ entails that $\bar{x}_t$ is strongly inconsistent with GARP thus there are $\bar{x}_{t_1}, \cdots, \bar{x}_{t_{r_t}}$, $r_t \geqslant 1$ with $\bar{x}_{t_i}R\bar{x}_t$ and $\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_{t_i}$, $\forall i = 1, \cdots, r_t$. Without loss of generality we can assume that $\bar{p}_t\bar{x}_{t_i} \geqslant \bar{p}_t\bar{x}_{t_{i+1}}$ $\forall\, i = 1, \cdots, r_t - 1$. Besides, for all $k \geqslant 1$ the definition of $\{\bar{\vartheta}_k\}_{k=1}^{n^2-n}$ yields $\bar{\vartheta}_k \geqslant \bar{\vartheta}$, which implies $\bar{x}_{t_j}R(\bar{\vartheta}_k)\bar{x}_t$ because $R(\bar{\vartheta})$ is a subrelation of $R(\bar{\vartheta}_k)$, and furthermore $\vartheta_{kt}\bar{p}_t\bar{x}_t \geqslant \vartheta_t\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_{t_j}$, which in turn yields the existence of $k_j > 1$ such that $\vartheta_{k_jt} = \frac{\bar{p}_t\bar{x}_{t_j}}{\bar{p}_t\bar{x}_t}$. Further, $\bar{\vartheta}_{k_j} \geqslant \bar{\vartheta}$ implies $\vartheta_{k_jt}\bar{p}_t\bar{x}_t \geqslant \vartheta_t\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_{t_j}$, therefore $\bar{p}_t\bar{x}_{t_j} \geqslant \vartheta_t\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_{t_j}$ and $\bar{p}_t\bar{x}_{t_j} > \bar{p}_t\bar{x}_{t_j}$, which is absurd. This proves the claim. ■

**IM1 describes the demand behavior quasi-better than Varian's index-mapping:** In view of Example 1 we only need to check that $\bar{\upsilon} = \bar{\upsilon}_{\mathscr{D}} \geqslant \bar{i}_{\mathscr{D}} = \bar{i}$ –we drop subindices for convenience– for every finite list of demand observation $\mathscr{D} = \{(\bar{p}_1, \bar{x}_1), ..., (\bar{p}_n, \bar{x}_n)\}$ that is inconsistent with GARP.

For every fixed bundle $\bar{x}_t$, if it is not strongly inconsistent with GARP, i.e., if $\nexists\, \bar{x}_s$ such that $\bar{x}_sR\bar{x}_t$ and $\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_s$, then $i_t = \upsilon_t = 1$ by construction. If $\bar{x}_t$ is strongly inconsistent with GARP there exist $\bar{x}_{t_1}, \cdots, \bar{x}_{t_{r_t}}$, $r_t \geqslant 1$, such that $\bar{x}_{t_i}R\bar{x}_t$ and $\bar{p}_t\bar{x}_t > \bar{p}_t\bar{x}_{t_i}$, $\forall i = 1, \cdots, r_t$. We do not lose generality by ordering its precedent bundles in such way that $\bar{p}_t\bar{x}_{t_i} \geqslant \bar{p}_t\bar{x}_{t_{i+1}}$, $\forall\, i = 1, \cdots, r_t - 1$. Suppose that $\bar{x}_t$ is strongly inconsistent with $GARP(\bar{\upsilon}_k)$ for every $k = 1, \cdots, r_t - 1$, and also that $\bar{x}_t$ is consistent with $GARP(\bar{\upsilon}_{r_t})$. Then $\upsilon_{k't} = \upsilon_{r_tt}$, $\forall\, k' = r_t, \cdots, n-1$, thus $\upsilon_t = \upsilon_{r_tt}$. Besides, as $\upsilon_{kt} = max\left\{\frac{\bar{p}_t\bar{x}_s}{\bar{p}_t\bar{x}_t\upsilon_{k-1t}} < 1 \,:\, \bar{x}_sR(\bar{\upsilon}_{k-1t})\bar{x}_t\right\}$, $\forall\, k = 1, \cdots, r_t$ the following equality holds:

$$\upsilon_t = \upsilon_{r_tt} = \frac{\prod_{i=1}^{r_t}\bar{p}_t\bar{x}_{t_i}}{\prod_{i=1}^{r_t-1}\bar{p}_t\bar{x}_i} \cdot \frac{1}{\bar{p}_t\bar{x}_t} = \frac{\bar{p}_t\bar{x}_{t_{r_t}}}{\bar{p}_t\bar{x}_t} = i_t$$

If there is $k \in \mathbb{N}$ ($k \leqslant r_t - 1$) such that $\bar{x}_t$ violates $GARP(\bar{\upsilon}_h)$ for every $h < k$ and $\bar{x}_t$ is consistent with $GARP(\bar{\upsilon}_k)$, then

$$\upsilon_t = \upsilon_{kt} = \frac{\prod_{i=1}^{k}\bar{p}_t\bar{x}_{t_i}}{\prod_{i=1}^{k-1}\bar{p}_t\bar{x}_i} \cdot \frac{1}{\bar{p}_t\bar{x}_t} \Longrightarrow \upsilon_t^k = i_t \cdot \frac{\prod_{i=k}^{r_t-1}\bar{p}_t\bar{x}_i}{\prod_{i=k+1}^{r_t}\bar{p}_t\bar{x}_{t_i}} = i_t \cdot \frac{\bar{p}_t\bar{x}_k}{\bar{p}_t\bar{x}_{r_t}} \geqslant i_t$$

Consequently, for every $t$ the inequality $\upsilon_t \geqslant i_t$ holds true. ■

**IM2 describes the demand behavior quasi-better than Varian's index-mapping:** In view of Example 1 we only need to check that $\bar{\vartheta} = \bar{\vartheta}_{\mathscr{D}} \geqslant \bar{i}_{\mathscr{D}} = \bar{i}$ –we drop subindices for convenience– for every finite list of demand observation $\mathscr{D} = \{(\bar{p}_1, \bar{x}_1), ..., (\bar{p}_n, \bar{x}_n)\}$ that is inconsistent with GARP.

For any bundle $\bar{x}_t$ we denote by $RI(\bar{x}_t)$ the set formed by all the demanded bundles $\bar{x}_s$ such that $\bar{x}_s R \bar{x}_t$ and $\bar{p}_t \bar{x}_t > \bar{p}_t \bar{x}_s$. Observe that for a given $\bar{x}_t$, if it is not strongly inconsistent with GARP, that is, if $RI(\bar{x}_t) = \varnothing$, by construction one has $\vartheta_t = i_t$. Therefore in order to prove the claim it suffices to check $\vartheta_t \geqslant i_t$ for every $\bar{x}_t$ that is strongly inconsistent with GARP. For one such bundle, there are $\bar{x}_{t_1}, \cdots, \bar{x}_{t_{r_t}}$, $r_t \geqslant 1$, such that $\bar{x}_{t_i} R \bar{x}_t$ and $\bar{p}_t \bar{x}_t > \bar{p}_t \bar{x}_{t_i}$, $\forall i = 1, \cdots, r_t$. We do not lose generality by ordering its precedent bundles in such way that $\bar{p}_t \bar{x}_{t_i} \geqslant \bar{p}_t \bar{x}_{t_{i+1}}$, $\forall\, i = 1, \cdots, r_t - 1$. Let $k^* = \sum_{t=1}^{n} \sharp RI(\bar{x}_t)$, then for every $\bar{x}_t$ and $k > k^*$ with $k^* < k \leqslant n(n-1)$ one must have $\vartheta_{kt} = \vartheta_{k^*t}$ thus $\vartheta_t = \vartheta_{k^*t}$. If $\bar{x}_t$ is strongly inconsistent with $GARP(\bar{\vartheta}_k)$ for every $k < k^*$, then

$$\vartheta_{k^*t} = \vartheta_t = \frac{\prod_{h=1}^{k^*} \vartheta_{ht}}{\prod_{h=1}^{k^*-1} \vartheta_{ht}} = \frac{\prod_{i=1}^{r_t} \bar{p}_t \bar{x}_{t_i}}{\prod_{i=1}^{r_t-1} \bar{p}_t \bar{x}_i} \cdot \frac{1}{\bar{p}_t \bar{x}_t} = \frac{\bar{p}_t \bar{x}_{t_{r_t}}}{\bar{p}_t \bar{x}_t} = i_t$$

Otherwise, there is $\tilde{k} < k^*$ such that $\bar{x}_t$ is strongly inconsistent with $GARP(\bar{\vartheta}_h)$ $\forall\, h < \tilde{k}$ and $\bar{x}_t$ verifies $GARP(\bar{\vartheta}_{\tilde{k}})$, thus

$$\vartheta_t = \vartheta_{\tilde{k}t} = \frac{\prod_{h=1}^{\tilde{k}} \vartheta_{ht}}{\prod_{h=1}^{\tilde{k}-1} \vartheta_{ht}} \Longrightarrow \vartheta_{\tilde{k}t} = i_t \cdot \frac{\prod_{h=\tilde{k}}^{k^*-1} \vartheta_{ht}}{\prod_{h=\tilde{k}+1}^{k^*} \vartheta_{ht}}$$

Furthermore, because $\bar{\vartheta}_h \leqslant \bar{\vartheta}_k$ for every $h > k$, one has $\prod_{h=\tilde{k}}^{k^*-1} \vartheta_{ht} \geqslant \prod_{h=\tilde{k}+1}^{k^*} \vartheta_{ht}$, which yields $\vartheta_{\tilde{k}t} \geqslant i_t$ and thus $\vartheta_t \geqslant i_t$ for each $t$. ■

**Acknowledgements**

José C. R. Alcantud gratefully acknowledges financial support by the Spanish *Ministerio de Ciencia e Innovación* under Project ECO2009-07682, and by *Consejería de Educación* de la *Junta de Castilla y León* under Project SA024A08. Carlos R. Palmero gratefully acknowledges financial support by *Ministerio de Ciencia e Innovación* under Project ECON2009-10231, and by *Consejería de Educación* de la *Junta de Castilla y León* under Project VA081A07.

## References


[1] Afriat, S.N. (1967) The construction of utility functions from expenditure data. International Economic Review, 8(1), 67-77.

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

[3] Battalio, R.C., Kagel, J.H., Winkler, R.C., Fisher, E.C. Jr., Basmann, R.L. & Krasner, L. (1973) A test of consumer demand theory using observations of individual consumer purchases. Western Economic Journal XI, 411-428.

[4] Deaton, A. & Muellbauer, J. (1980) An Almost Ideal Demand System. American Economic Review, 70(3), 312-336.

[5] Deaton, A. & Muellbauer, J. (1980) Economics and Consumer Behavior. Cambridge, England: Cambridge University Press.

[6] Famulari, M. (1995) A household-based, nonparametric test of demand theory. The Review of Economics and Statistics 77, 372-382.

[7] Gross, J. (1991) On expenditure indices in revealed preference tests. Journal of Political Economy, 98, 416-419.

[8] Houtman, M. & Maks, J.A. (1985) Determining all maximal data subsets consistent with revealed preference. Kwantitatieve Methoden, 19, 89-104.

[9] Koo, A. (1963) An empirical test of revealed preference theory. Econometrica, 31(4), 646-664.

[10] Mattei, A. (1994) La coherence des choix des consommateurs suisses. Revue Suisse d'Economie Politique et de Statistique, 130, 3-20.

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

[12] Sippel, R. (1997) An experiment on the pure theory of consumer's Behaviour. The Economic Journal, 107, 1431-44.

[13] Swofford, J. & Whitney, G. (1987) Nonparametric tests of utility maximization and weak separability for consumption, leisure and money. Review of Economics and Statistics, 69(3), 458-64.

[14] Tsur, Y. (1989) On testing for revealed preference conditions. Economics Letters, 31, 359-362.

[15] Varian, H. (1982) The nonparametric approach to demand analysis. Econometrica, 50(4), 945-971.

[16] Varian, H. (1990) Goodness-of-fit in optimizing models. Journal of Econometrics, 46, 125-140.

[17] Varian, H. (1993) Goodness-of-fit for revealed preference tests. CREST Working Paper. Department of Economics. University of Michigan.