# A MEASURE OF RATIONALITY AND WELFARE*

JOSE APESTEGUIA† AND MIGUEL A. BALLESTER‡

Abstract. There is evidence showing that individual behavior often deviates from the classical principle of maximization. This evidence raises at least two important questions: (i) how severe the deviations are and (ii) which method is the best for extracting relevant information from choice behavior for the purposes of welfare analysis. In this paper we address these two questions by identifying from a foundational analysis a new measure of the rationality of individuals that enables the analysis of individual welfare in potentially inconsistent subjects, all based on standard revealed preference data. We call such measure minimal index.

**Keywords:** Rationality; Individual Welfare; Revealed Preference.
**JEL classification numbers:** D01; D60.

## 1. Introduction

The standard model of individual behavior is based in the maximization principle, in which the alternative chosen by the individual is the one that maximizes a well-behaved preference relation over the menu of available alternatives. The consensus concerning this principle can be explained by at least two of its main features. The first is that it provides a simple and versatile account of individual behavior. It is difficult to conceive of a simpler and more operational model with such a large predictive power. Its second main feature is that it suggests the maximized preference relation as a tool for individual welfare analysis. That is, the standard approach allows the policy-maker

*Date*: September, 2011.
* We thank Larbi Alaoui, Christopher Chambers, Tugce Cuhadaroglu, Eddie Dekel, Jerry Green, Paul Milgrom, Mauro Papi, Larry Samuelson, Karl Schlag, and three referees for helpful comments. Financial support by the Spanish Commission of Science and Technology (ECO2009-12836, ECO2008-04756, ECO2008-01768), the Barcelona GSE research network, and the Government of Catalonia is gratefully acknowledged.
† ICREA and Universitat Pompeu Fabra. E-mail: `jose.apesteguia@upf.edu`.
‡ Universitat Autonoma de Barcelona. E-mail: `miguelangel.ballester@uab.es`.

aiming to reproduce the decisions that, given the chance, the individual would have made of her own volition.

Over the last decades, however, the research has produced increasing amounts of evidence documenting systematic and predictable deviations from the notion of rationality implied in the maximization principle. Some phenomena that have attracted a great deal of empirical and theoretical attention and prove difficult, if not impossible, to accommodate within the classical theory of choice are framing effects, menu effects, dependence on reference points, cyclic choice patterns, choice overload effects, etc.[1] The violation in some instances of the maximization principle raises at least two important questions:

**Q.1:** how severe are the deviations from the classical theory?

**Q.2:** what is the best way to extract relevant information from the choices of the individual for the purposes of welfare analysis?

By properly addressing Q.1, it would be possible to evaluate whether the classical theory of individual behavior is a reasonable way to describe behavior. That is, the validity of the theory should not be based on whether or not individuals violate the maximization principle in a given situation, but on how close their behavior is with respect to this benchmark. Moreover, the availability of a reliable tool to assess the distance between actual behavior and behavior consistent with the maximization of a preference relation will enable interpersonal comparisons. This, in turn, may improve our understanding of individual behavior and may also prove crucial in the development of future choice models, which may take the distribution of the degree of consistency of the individuals as one of their primitives. Furthermore, the possibility of performing meaningful comparisons of rationality will allow evaluation of deviations between various alternative models of choice and, hence, provide a tool to give some structure to the rapidly growing literature on alternative individual decision-making models that are expanding the classical notion of rationality (see footnote 1 for some recent examples).

[1] As with the empirical findings see, respectively, Tversky and Kahneman (1981), Tversky and Simonson (1993), Thaler (1980), May (1954), and Iyengar and Lepper (2000). Some theoretical accounts inspired by the above findings expanding the classical notion of rationality and adopting a revealed preference approach are Bossert and Sprumont (2003), Masatlioglu and Ok (2005), Manzini and Mariotti (2007), Salant and Rubinstein (2008), Masatlioglu and Nakajima (2008), and Masatlioglu, Nakajima and Ozbay (2010).

By dealing with Q.2, it would be possible to identify, from an external perspective, the good or bad alternatives for the individual even when the behavior of the individual is not fully compatible with the maximization principle. This is of course of prime relevance since welfare analysis is at the core of economics.

In principle, there is a wide range of different possible procedures in order to address these two questions. In this paper we first introduce the notion of an inconsistency index that evaluates the inconsistency associated to a collection of revealed choices. We then suggest a set of foundational conditions on inconsistency indices that identify a new class of indices to measure the rationality of individuals that, at the same time, enables individual welfare analysis using potentially inconsistent subjects. We call it the class of minimal indices.

Minimal indices measure the inconsistency of the data by identifying the preference relation that is closest to observed behavior. Distance is evaluated by additively considering every menu of alternatives when there is a divergence between the choice of alternative dictated by the maximization principle and the one actually chosen by the individual, and then by weighting such divergence by means of a collection of weights that may depend on the preference relation, the menu of available alternatives, and the chosen alternative. Minimal indices are therefore flexible enough to evaluate the inconsistency of choice not only by counting the number of menus in which the individual deviates from the maximization principle, but also by pondering each inconsistent observation on the basis of the possible underlying values of the alternatives and the various menus of alternatives, judged by the closest preference relation to the data.

Given the general formulation of weights in minimal indices, it follows that this class of indices is a broad one encompassing a number of special cases. In this paper we study in detail a version of the well-known Varian's index, and two new ones that we call the minimal swaps index and the minimal loss index. Since each of them has its own attractive features, we aim to gain a deeper understanding by establishing the properties that, on top of those characterizing the class of minimal indices, uniquely identify them.[2] This exercise should be informative to the analyst when choosing one index over the others in applications.

Varian (1990) refines the first suggestion in the literature measuring the consistency of the data due to Afriat (1973). In a consumer setting, Afriat proposes to measure

[2]We also show, as a corollary, that the Houtman-Mak's index is a minimal index, and discuss its characterization.

the amount of relative wealth adjustment required in each budget constraint to avoid all violations of the maximization principle. The suggestion of Varian is to consider potentially different levels of wealth adjustments in the different observations, and then measure the total wealth adjustment that is necessary in order to rationalize all the choices. In spite of the fact of its very different formulations, we show that in our setting the Varian's index belongs to the class of minimal indices. This exemplifies the compatibility of minimal indices with the use of exogenous information to the revealed choices.

The minimal swaps index is the first of the new indices belonging to the broad class of minimal indices that we introduce in this paper. The minimal swaps index evaluates the inconsistency of observations unexplained by a preference relation by enumerating the number of available alternatives in the menus that are above the chosen ones according to the preference relation. That is, it counts in each menu the number of alternatives that need to be swapped with the chosen alternative in order to explain the choices of the individual. Hence, the minimal swaps index is given by the preference relation that minimizes the total number of swaps in all the observations. Importantly, this index judges the inconsistency in the choices of an individual by using exclusively the information that arises endogenously from the revealed preference that is closest to the data.

The minimal loss index, the second new index belonging to the class of minimal indices that we put forward in this paper, also uses the information arising endogenously from the revealed choices of the individual, in addition to some exogenous cardinal information. Suppose that the analyst has information on the cardinal utility values of the different alternatives, based on their position in the ranking. Then, minimal loss indices evaluate an inconsistent observation by attending to the difference in the utility values associated to the maximal available alternative in the menu and to the chosen alternative. This can be interpreted as the utility loss for the inconsistent choice in that observation. Then, the minimal loss index is given by the preference relation that minimizes the sum of utility losses.

The structure of minimal indices immediately suggests a way to deal with individual welfare. Minimal indices focus on the preference relation that minimizes the inconsistency with the data and that can be interpreted as the best approximation of the individual's actual choices. It is then natural to use this preference relation as the instrument to make welfare analysis. The main advantages of such an approach are that the relevant preference relation is identified from a foundational analysis, and that the preference relation is well-behaved in the sense of being a linear order and hence its maximization offers a unique alternative in any possible menu of alternatives, thus providing the policy maker with a clear guideline.

We then contrast our approach to welfare analysis with two prominent suggestions in the literature, Bernheim and Rangel (2009) and Green and Hojman (2009). This exercise illustrates that the ranking of any two alternatives in the preference resulting from our approach is constructed taking into consideration the whole data set, and not just those observations where the two alternatives are present. In principle, therefore, the welfare criterion that we suggest endogenizes all the consequences of ranking one alternative over another, and hence it may be more informative for the policy maker.

Our paper draws on two significant strands of literature. The first is the literature on revealed preference tests of the maximization principle. These tests are typically run on consumer behavior data, with no axiomatic foundation. As we mentioned above, the first to propose this type of test was Afriat (1973), that suggested measuring the amount of adjustment required in each budget constraint to avoid any violation of the maximization principle. Chalfant and Alston (1988) and Varian (1990) further developed Afriat's approach. Another proposal is to count the number of violations of consistency with the maximization principle detected in the data. In this respect, see the work of Swofford and Whitney (1987) and Famulari (1995). Yet a third approach computes the maximal subset of the data that is consistent with the maximization principle. Papers following this approach are Houtman and Maks (1985) and Banker and Maindiratta (1988). In a recent paper, Dean and Martin (2010) provide a powerful algorithm to compute such a maximal subset of the data. We devote section 2 to review these tests within the framework we propose in this paper.

The second includes a growing number of papers dealing with individual welfare analysis, even when the individual's behavior is inconsistent. Bernheim and Rangel (2009) add to the standard revealed choice data the notion of ancillary conditions, or frames. Ancillary conditions are assumed to be observable and may affect individual choice, but are irrelevant in terms of the welfare associated with the chosen alternative. Bernheim and Rangel suggest a Pareto-type welfare preference relation that ranks an alternative as welfare-superior to another only if the latter is never chosen when the former is available. It is clear that this method may provide an incomplete preference relation, and hence in occasions will be uninformative. Chambers and Hayashi (2009) characterize an extension of Bernheim and Rangel's model to probabilistic settings, providing a complete welfare ranking. Manzini and Mariotti (2009) offer a critical assessment of Bernheim and Rangel. A different view on the question of behavioral welfare analysis is discussed in Koszegi and Rabin (2007) where welfare analysis is proposed to be done by studying the information regarding the cognitive process that underlies the choices of the individual. In this line, Rubinstein and Salant (2011) propose the welfare relation that is consistent with a set of preference relations in the sense that all the preference relations in the set could have been generated by the cognitive process distorting that welfare relation. Also consistent with this view, Masatlioglu, Nakajima and Ozbay (2010) suggest a welfare preference based on their limited attention model of decision-making. Alternatively, Green and Hojman (2009) suggest identifying a list of conflicting selves, which, when aggregated, induce the revealed choices, and then using the aggregation rule to make the individual welfare-analysis. Finally, Baldiga and Green (2010) analyze the conflict between preference relations in terms of their disagreement on choice. They then use their measures of conflict between preference relations together with Green and Hojman's notion of multiple-selves to find the list of multiple-selves with the minimal internal conflict that explain a given choice data, and suggest this as a welfare measure.

The rest of the paper is organized as follows. Section 2 formally presents our environment, reviews the most prominent classical measures that have been suggested in the literature, and introduces the class of minimal indices. In section 3 we lay down the axiomatic characterization of minimal indices. Then, section 4 is devoted to the study of various especially prominent cases of minimal indices: Varian's, Houtman-Mak's, minimal swaps and minimal loss indices. Section 5 introduces our approach to individual welfare with potentially inconsistent individuals, compares it with two prominent suggestions in the literature, Bernheim and Rangel (2009) and Green and Hojman (2009), and discusses the interpretation of minimal indices in light of these two approaches. Section 6 deals with the question of computing the index in practice and obtaining the optimal preference relation from the data. Section 7 concludes. All the proofs are contained in the appendix.

## 2. Measures of Rationality

Let $X$ be a finite set of $k$ alternatives. An observation is a pair $(A, a)$, where $A \subseteq X$ is a non-empty menu of alternatives and $a \in A$ is the chosen alternative. Denote by $\mathcal{O}$ the set of all possible observations. A collection of observations is a mapping $f : \mathcal{O} \to \mathbb{Z}_+$, indicating the number of times each observation $(A, a)$ occurs. Note that we allow for the possibility that the same menu $A$ may be in $f$ more than once, with the same or different associated chosen elements. Denote by $\mathcal{F}$ the set of all possible collections of observations.

A preference relation $P$ is a strict linear order on $X$, that is, an asymmetric, transitive, and connected binary relation. Denote by $\mathcal{P}$ the set of all possible linear orders on $X$. The collection of observations $f$ is rationalizable if there exists a preference relation $P$ that explains the collection of observations, i.e., $m(P, A) = a$ whenever $f(A, a) > 0$, where $m(P, A)$ represents the maximal element in $A$ according to $P$. That is, $f$ is rationalizable if every single observation that is registered at least once in $f$ can be explained by the maximization of the same preference relation. Clearly, not every collection of observations is rationalizable. An inconsistency index is a mapping $I : \mathcal{F} \to \mathbb{R}_+$ that measures how inconsistent (or how far away from rationalizability) a collection of observations is. We now formally introduce in our framework some inconsistency indices that have been suggested in the literature, and a new family of indices that we call minimal indices.

2.1. **Classical Indices of Inconsistency.** To the best of our knowledge, Afriat (1973) proposes the first method measuring the inconsistency of behavior. In a consumer setting, Afriat suggests to measure the amount of relative wealth adjustment required in each budget constraint to avoid all violations of the maximization principle. The idea is that when a portion of the wealth is considered, the budget set shrinks eliminating some revealed information, and hence some inconsistencies in the data may vanish. Then, the degree of inconsistency of a collection of observations that Afriat proposes is associated to the minimal necessary wealth adjustment that makes all the data consistent with the maximization principle.

In order to represent Afriat's measure in our setting, we consider the notion of the level of attention the individual pays to the menu of alternatives, along the lines suggested by Masatlioglu, Nakajima and Ozbay (2010). An attention mapping $\phi$ assigns to every menu $A$ the alternatives that are considered at the attention level $e \in [0, 1]$. The mapping $\phi$ satisfies the following conditions: (1) $\phi(A, 0) = \emptyset$, (2) $\phi(A, 1) = A$, and (3) $\phi(A, e)$ is increasing in $e$. That is, when there is no attention, $e = 0$, no alternative is considered, while when there is full attention, $e = 1$, all the alternatives in the menu are considered. The higher the attention level, the more alternatives are considered in

$A$. In this context, the question is how much inattention on the part of the individual one needs to assume in order to rationalize all her choices.

It is important to make clear here that the severity of the violations is captured using information unrelated to the revealed preference. In our formulation, the severity is captured by the exogenous attention mapping $\phi$, which plays the role of the budget set adjustments in the original Afriat's consumer setup. That is, both, in the consumer setting and in our general formulation, the way of capturing the severity of a violation makes use of some exogenous structure of the menus of alternatives.

For any attention mapping $\phi$, collection of observations $f$, and level of attention $e \in [0,1]$, consider the revealed preference defined by $xR^{\phi}_{f,e}y \Leftrightarrow \exists(A,x)$ with $f(A,x) > 0$ and $y \in \phi(A,e)$. That is, $x$ is revealed preferred to $y$ if, given the attention mapping $\phi$, there is an observation where $x$ is chosen and $y$ is observed at the attention level $e$. If the level of attention is equal to 1, this is merely the standard revealed preference relation. If the level of attention is equal to 0, there is no revelation from the data since no alternative is observed. The less attention is assumed, the less revealed information and the less likely inconsistencies are to appear. There is then a limit value of $e$ at which the revealed preference $R^{\phi}_{f,e}$ does not contain any cycle and hence, it is compatible with the principle of maximization. This suggests Afriat's measure of rationality:

$$I_{Af}(f) = \inf_{e:R^{\phi}_{f,e} \text{ is acyclic}} (1-e).$$

The inconsistency value of the collection of observations $f$ according to the index $I_{Af}$ is the lowest possible level of inattention $(1-e)$ such that the revealed preference at attention level $e$ is acyclic. We now illustrate this measure and some of its limitations by using a simple example.

Let $X = \{x,y,z\}$ and let the perception of a menu of alternatives depend on the individual's level of attention, as follows. If her level of attention is low, the individual will perceive no alternative. If her level of attention is intermediate, the individual will perceive alternatives $y$ and $z$, whenever these are available. If her level of attention is sufficiently high, she will perceive all the available alternatives. More concretely:

$$\phi(A,e) = \begin{cases} \emptyset & \text{if } e \leq 1/4; \\ A \cap \{y,z\} & \text{if } 1/4 < e \leq 3/4; \\ A \cap \{x,y,z\} & \text{if } 3/4 < e. \end{cases}$$

In order to be perceived, alternatives $x, y$ and $z$ require attention levels above $3/4$, $1/4$ and $1/4$, respectively. Consider now the collection of observations $f$ such that the only observations are $f(\{x,y\},x)=2$ and $f(\{y,z\},y)=f(\{x,z\},z)=f(\{x,y,z\},x)=1$. That is, $f$ reveals a cycle in the binary menus of alternatives, and another cycle given by $(\{x,z\},z)$ and $(\{x,y,z\},x)$. Note that both cycles disappear if alternative $x$ is not considered in the menu $(\{x,z\},z)$. Given $\phi$, this is guaranteed at the level of attention $e=3/4$. At such level of attention the observation $(\{x,z\},z)$ no longer reveals that $z$ is preferred to $x$, i.e. it is not the case that $zR^{\phi}_{f,3/4}x$. It then follows that $I_{Af}(f)=1-3/4=1/4$.

Consider now a collection of observations $g$ consisting of the same observations in $f$ except that the choice from $\{x,y,z\}$ is alternative $y$ instead of $x$. Notice that on top of the same cycle involving the binary menus of alternatives, $g$ contains the cycles generated by $g(\{x,y\},x)=2$ and $g(\{x,y,z\},y)=1$. The collection of observations $g$ can be rationalized if alternative $x$ is neither observed from $(\{x,z\},z)$ nor from $(\{x,y,z\},y)$. This follows at the level of attention $e=3/4$ and consequently, $I_{Af}(g)=1-3/4=1/4=I_{Af}(f)$.

We conclude therefore, that at the light of Afriat's index, both $f$ and $g$ are identical in terms of their inconsistency value. It may be natural to argue, however, that a reasonable measure of rationality should assign a higher inconsistency value to $g$ than to $f$, since one needs to assume the individual ignores alternatives more often in the former than in the latter.

In our setting, Varian's (1990) refinement of Afriat's measure of inconsistency can be presented as follows. Varian addresses the question of how much aggregated inattention on the part of the individual one needs to assume in order to rationalize all her choices. In order to do so, he considers potentially different levels of attention in the different observations $\mathbf{e}=\{e_{(A,a)}\}$, instead of a single level of attention $e$ for all the observations. Then, for any attention mapping $\phi$, collection of observations $f$, and vector of levels of attention $\mathbf{e}$, consider the revealed preference defined by $xR^{\phi}_{f,\mathbf{e}}y \Leftrightarrow \exists(A,x)$ with $f(A,x)>0$ and $y\in\phi(A,e_{(A,x)})$. We are interested in the vector of attention levels $\mathbf{e}$, closest to 1, that makes $R^{\phi}_{f,\mathbf{e}}$ acyclic. We can therefore define Varian's index as:[3]

$$I_V(f)=\inf_{\mathbf{e}:R^{\phi}_{f,\mathbf{e}}\text{ is acyclic}}\sum_{(A,a)} f(A,a)(1-e_{(A,a)}).$$

[3]Varian's original definition considers any possible norm. For tractability, we here use the linear version.

To illustrate this index consider again the above example. It is easy to see that the optimal vector $\mathbf{e}$ for $f$ is given by $e_{(\{x,z\},z)} = 3/4$ and $e_{(\{x,y\},x)} = e_{(\{y,z\},y)} = e_{(\{x,y,z\},x)} = 1$, implying a Varian's degree of inconsistency equal to $1/4$. The optimal vector $\mathbf{e}'$ for $g$ is given by $e'_{(\{x,z\},z)} = e'_{(\{x,y,z\},y)} = 3/4$ and $e'_{(\{y,z\},y)} = e'_{(\{x,y\},x)} = 1$, implying a Varian's degree of inconsistency equal to $I_V(g) = 1/2 > 1/4 = I_V(f)$.

Hence, unlike in the case of Afriat's index, since Varian's index is sensitive to both, the number and the severity of unexplained observations, it discriminates between collections of observations $f$ and $g$.

Houtman and Maks (1985) propose considering the minimal subset of observations that needs to be eliminated from the data in order to make the remainder rationalizable. The cardinality of this minimal subset to be discarded suggests itself as a measure of inconsistency:

$$I_{HM}(f) = \min_{g \leq f : f-g \text{ is rationalizable}} \sum_{(A,a)} g(A,a).$$

The following example illustrates this measure of rationality. Consider the collections of observations $f_1$ and $f_2$ given by $f_1(\{x,y\},x) = f_1(\{y,z\},y) = f_2(\{x,y\},x) = f_2(\{y,z\},y) = 3$ and $f_1(\{x,y,z\},y) = f_2(\{x,y,z\},z) = 1$. That is, $f_1$ and $f_2$ share the same menus of alternatives, and also the same choices, except for menu $\{x,y,z\}$. It follows that the maximal subsets consistent with the maximization principle of collections $f_1$ and $f_2$ coincide. In both cases, it involves eliminating the observation related to menu $\{x,y,z\}$. The inconsistency index associated to both collections of observations is therefore equal to 1.

Notice that the revealed preference associated to the maximal subsets of $f_1$ and $f_2$ consistent with the maximization principle establishes, in both cases, that $x$ is better than $y$, and $y$ better than $z$. In this light, it seems logical to argue that the inconsistent choice in $f_1$, namely $(\{x,y,z\},y)$, is less severe than that in $f_2$, $(\{x,y,z\},z)$, since the latter selects the worst possible alternative, while the former selects the intermediate one. However, $I_{HM}$, by only counting the number of menus that cannot be rationalized, ignores the severity of the inconsistencies.

Finally, rationality has also been measured by counting the number of times in the data a consistency property is violated (see, e.g., Swofford and Whitney, 1987). Depending on the setting, the consistency property may be WARP, or GARP, or

SARP, or IIA. This measure has exactly the same limitation as the one above, in that it takes into consideration the number of violations, but not their severity.[4]

2.2. **Minimal Indices.** We now propose a new and broad family of indices that we call minimal indices. The basic idea of minimal indices is to consider every possible inconsistency between an observation $(A, a)$ and a preference relation $P$ through a weight that may depend on the nature of the menu of alternatives $A$, the chosen alternative $a$, and the preference relation $P$. Then, for a given collection of observations $f$, the inconsistency index takes the form of the minimum total inconsistency across all preference relations. This approach represents a very flexible way of measuring the inconsistency of an individual and at the same time it highlights the preference relation that is closest to the data, enabling welfare analysis.

In order to formally introduce the class of minimal indices consider first a mapping $w : \mathcal{P} \times \mathcal{O} \to \mathbb{R}_+$ such that $w(P, A, a) = 0$ if and only if $a = m(P, A)$. That is, observations that are explained by the preference $P$ receive a null weight, while any other observation receives a strictly positive weight. For given weights $w$ an inconsistency index is a minimal index if

$$I_M(f) = \min_{P \in \mathcal{P}} \sum_{(A,a)} f(A, a) w(P, A, a).$$

The generality of the weights $w(P, A, a)$ allows us to consider various types of information on the measurement of inconsistencies. Minimal indices may consider the nature of the menus of alternatives, the nature of the chosen alternatives, and all this is judged on the basis of the closest preference relation to the data. Hence, minimal indices may make use of the endogenous information arising from the revealed choices in different ways, and also of different types of exogenous information related to the characteristics of the chosen alternatives or the menus of alternatives. Consequently, minimal indices encompass a broad class of special cases, each one with its own characteristics, and its own way of using exogenous or endogenous information to the revealed choices. We illustrate this by highlighting four special cases.

We start the analysis of special cases of minimal indices by showing that, despite its very different formulation, the Varian's index belongs to it. This exemplifies the

[4]Echenique, Lee and Shum (2010) make use of the monetary structure of budget sets to suggest a version of this notion, the money pump index, that captures also the severity of each violation. They consider the total wealth lost in all revealed cycles.

possible use of external information to the revealed choices that are nevertheless at the core of our measure of rationality.

**Proposition 1.** *For every attention mapping, the associated Varian's index is a minimal index.*

The weight associated to observation $(A, a)$ and preference $P$ that makes the Varian's index a minimal index takes the following form: it is one minus the largest level of attention for observation $(A, a)$ for which none of the alternatives that dominates $a$ according to $P$ is observed by the individual.

Indeed, in our setting, it is not difficult to see that the inconsistency index $I_{HM}$ is but a special case of a Varian's index when the attention mapping takes the form $\phi(A, e) = A$ for all $e > 0$. Hence, given Proposition 1, the Houtman-Maks' index is also a minimal index. We omit the formal proof of this immediate fact.

**Proposition 2.** *The Houtman-Maks' index is a minimal index.*

But the family of minimal indices allows for more general weights than those considered in the previous two indices. We now introduce two new special cases that focus in different ways on the endogenous information that emerges from the revealed choices. We call them minimal swaps indices and minimal loss indices. While the former exclusively uses the information contained in the revealed choices, the latter makes use also of exogenous information on the cardinal values of the alternatives.

Let us start by introducing the minimal swaps indices. Consider an observation $(A, a)$ that is inconsistent with the maximization of a given preference relation $P$. This implies that there is a number of alternatives in $A$ that are preferred to the chosen alternative $a$, according to $P$, but that are nevertheless ignored. A natural inconsistency measure of $P$ with regard to $(A, a)$ entails counting the number of alternatives in $A$ above the chosen one. These are the alternatives that need to be swapped with the chosen alternative in order to explain the choice. Then, the minimal swaps index is given by the preference $P$ that minimizes the total number of swaps in all the observations:

$$I_S(f) = \min_{P \in \mathcal{P}} \sum_{(A,a)} f(A, a)|\{x \in A : xPa\}|.$$

To illustrate this index, consider again the example we introduced in the discussion of index $I_{HM}$. The optimal preference for both $f_1$ and $f_2$ is given by $xPyPz$, that only fails to explain the observations involving the grand set. The upper contour set of alternative $y$ in $\{x, y, z\}$ for such a $P$ is composed by a unique alternative, $x$, and hence, $I_S(f_1) = 1$. However, the upper contour set of alternative $z$ in $\{x, y, z\}$ for such a $P$ is composed by alternatives $x$ and $y$ and therefore, $I_S(f_2) = 2 > I_S(f_1)$.

In the case of our second new special case of minimal indices, the minimal loss indices, the measurement of the severity of each inconsistency uses the revealed preference, together with some cardinal information on it. Consider again an observation $(A, a)$ and a given preference relation $P$ whose maximization in $A$ is inconsistent with $a$. In addition, suppose that one has information on the cardinal utility value of the alternatives, based on their positions in the ranking.[5] Then, one may value the inconsistency of $P$ with regard to $(A, a)$ by the difference in the cardinal values associated to the maximal element in $A$ and the one associated to the chosen element $a$. Formally, denote by $\widehat{m}(P, A)$ and $\widehat{a}(P)$ the rankings of alternatives $m(P, A)$ and $a$ in $P$. Now, given a vector of real-valued weights $u = (u_1, \ldots, u_k)$ such that $u_1 > u_2 > \cdots > u_k$, an inconsistency index is a minimal loss index if

$$I_L(f) = \min_{P \in \mathcal{P}} \sum_{(A,a)} f(A, a)(u_{\widehat{m}(P,A)} - u_{\widehat{a}(P)}).$$

It is immediate that in the above mentioned example, the minimal loss index evaluates the collections $f_1$ and $f_2$ as $I_L(f_2) = u_1 - u_3 > u_1 - u_2 = I_L(f_1)$.

It therefore emerges that minimal indices represent a prominent and broad class of inconsistency indices. It encompasses the Houtman-Maks' index, together with other indices that take into consideration both the number of inconsistencies and their severity, like the Varian's index, the minimal swaps index and the minimal loss index. Furthermore, by allowing to measure the inconsistency of the data on the basis of the closest preference relation to the revealed choices, it proposes a tool to do welfare analysis. Consequently, we aim to gain a deeper understanding of the class of minimal indices. For this purpose, we now provide a complete characterization of them.

[5] For example, the cardinal utility values may be modeled as the random variables with a common distribution, and then one may sort the realizations in decreasing order of magnitude to focus on the order statistics. The expected value of the order statistics may be taken as the cardinal values of the alternatives based on their ranking, as in Apesteguia, Ballester and Ferrer (2011).

## 3. A Characterization of Minimal Indices

In this section we propose four conditions that an inconsistency index $I$ should desirably satisfy. The four conditions shape the treatment that an inconsistency index $I$ should ideally give to various sorts of collections of observations. We then show that the minimal indices, and only them, satisfy this set of four properties. Our first characterizing property is the following.

**Rationality (RAT).** Let $f \in \mathcal{F}$, $I(f) = 0 \Leftrightarrow f$ is rationalizable.

Rationality imposes that a collection of observations has an inconsistency value of 0 if and only if the collection is rationalizable. Since by definition an inconsistency index $I$ assigns non-negative numbers to any possible collection of observations, Rationality imposes a weak monotonicity requirement. In line with the maximization principle, the minimal inconsistency level is reached only when every single choice in the collection of observations can be explained by maximizing a preference relation. The property normalizes the smallest inconsistency level by assigning a value of 0 to this case and a strictly higher inconsistency level to any other case.

In order to introduce our next property, consider first the following definition. Denote by $r$ a rationalizable collection of observations where all the binary menus are observed and by $\mathcal{R}$ the set of all such collections of observations. Clearly, for every $r \in \mathcal{R}$ there exists a unique preference relation, that we denote by $P^r$, that rationalizes $r$. Given $f$ and $r$, we say that $f$ is $r$-invariant if $I(f) = I(f + r)$. That is we say that $f$ is $r$-invariant whenever the addition of $r$ to $f$ leaves the inconsistency level associated to $f$ unchanged.

**Invariance (INV).** For every $f \in \mathcal{F}$, there exists $r \in \mathcal{R}$ such that $f$ is $r$-invariant.

Invariance implies that every possible collection of observations can be related to some rationalizable collection of observations involving the binary menus of alternatives. No matter the nature of the collection of observations $f$, there is an $r$ that when added to $f$ does not increase the inconsistency associated to $f$. In other words, Invariance imposes that any $f$ has at least one close enough rationalizable collection of observations that can be added without consequences on the total inconsistency value.

**Attraction (ATTR).** For every $f \in \mathcal{F}$ and every $r \in \mathcal{R}$, there exists a positive integer $z$ such that $f + zr$ is $r$-invariant.

Consider any collection of observations $f$ and any rationalizable collection $r$. Attraction establishes that if some structured data $r$ is added sufficiently often to $f$, extra information confirming the structured data becomes inessential, in the sense of leaving the inconsistency of the data unchanged. While Invariance imposes the existence of a rationalizable collection $r$ that makes $f$ $r$-invariant, Attraction says that $f+zr$ will be $r$-invariant as long as $r$ is made sufficiently important in the database.

**Separability (SEP).** For every $f,g\in\mathcal{F}$, $I(f+g)\geq I(f)+I(g)$, with equality if and only if $f$ and $g$ are $r$-invariant for some $r\in\mathcal{R}$.

Take any two collections of observations $f$ and $g$. Separability judges that the sum of the inconsistency of $f$ and that of $g$ can never be greater than the inconsistency associated to the collection of observations resulting from $f$ and $g$ together. To illustrate, suppose that $f$ and $g$ are both rationalizable when taken separately. Clearly, the conjunction of $f$ and $g$ does not need to be rationalizable, and hence arguably the collection of observations $f+g$ can only take the same or a larger inconsistency value than the sum of the inconsistency values of the two collections separately. The same idea applies when either $f$ or $g$ or both are not rationalizable. The sum of $f$ and $g$ can only generate the same or more frictions with the rationalizability principle, and hence should get the same or a larger inconsistency value. Further, Separability also states that in the special case where both $f$ and $g$ are related to the same rationalizable collection of observations $r$, then the inconsistency of $f+g$ should be the same than the sum of the inconsistencies of $f$ and $g$ taken separately. The rationale for this second implication of Separability is that if $f$ and $g$ share a common structure, that is both are $r$-invariant with the same $r$, the two collections are analogous enough so that they can be aggregated without further consequences on the inconsistency value.

Theorem 1 shows that these four properties uniquely characterize the class of minimal indices.

**Theorem 1.** *An inconsistency index $I$ satisfies (RAT), (INV), (ATTR) and (SEP) if and only if it is a minimal index.*

The intuition of the only if part of the proof can be grasped by considering the following four steps.

(1) **Categorization:** We start by relating any collection of observations $f$ with one of the rationalizable collections of observations $r$ involving all binary menus of alternatives, such that $f$ is $r$-invariant. This provides the preference relation $P^r$ that will be shown to be the closest to the rationalization of $f$. Invariance guarantees that such a rationalizable collection $r$ exists.

(2) **Separation:** We then show that the four properties guarantee that we can write the inconsistency of an $r$-invariant collection of $n$ observations, $f$, as $I(f) = \sum_{(A,a)} f(A,a) I(\mathbf{1}_{(A,a)} + z_n^r r)$. That is, we can separate the inconsistency of $f$ on the basis of the different observations $(A,a)$ and weight each one of these observations by $I(\mathbf{1}_{(A,a)} + z_n^r r)$. The latter denotes the inconsistency value assigned to the collection formed by the observation $(A,a)$, $\mathbf{1}_{(A,a)}$, together with $z_n^r$ times the rationalizable collection $r$. The integer $z_n^r$ guarantees, by Attraction, that any possible collection of $n$ or fewer observations is $r$-invariant.

(3) **Representation:** Next, we define the weighting mapping $w$ as $w(P^r, A, a) = 0$ whenever $P^r$ rationalizes the observation $(A,a)$, and $w(P^r, A, a) = I(\mathbf{1}_{(A,a)} + z_1^r r)$ otherwise. By using (SEP) and (RAT) we can equate the values $I(\mathbf{1}_{(A,a)} + z_1^r r)$ and $I(\mathbf{1}_{(A,a)} + z_n^r r)$, and then it follows that $I(f) = \sum_{(A,a)} f(A,a) w(P^r, A, a)$.

(4) **Minimality:** Finally, for any other preference $P'$ we can consider the associated collection $r'$. By (ATTR) and the previous representation, for $z$ large enough we know that $I(f + zr') = \sum_{(A,a)} f(A,a) w(P', A, a)$. (SEP) guarantees that this value is larger or equal than $I(f) + I(zr')$, and (RAT) leads to $I(f) + I(zr') = I(f) = \sum_{(A,a)} f(A,a) w(P^r, A, a)$, as desired.

## 4. Special Cases of Minimal Indices

We have argued that the class of minimal indices is a prominent one. It is a general and flexible class encompassing a number of inconsistency indices, some of them sharing an attractive feature: being sensitive to both the number of inconsistencies and their severity. We now study in detail four special cases of minimal indices, the Houtman-Maks', Varian's, minimal swaps and minimal loss indices, by providing a characterization for each one of the cases. The characterization results are all based on the system of properties characterizing the minimal indices, and add extra structure in order to capture the specific nature of the different classes. This exercise enhances the understanding of the different indices by stressing the additional structure that is needed in order to obtain the desired representation, facilitating the evaluation of the similarities and differences between them and helping the observer to decide which of these indices to use in applications.

4.1. **Varian's Index.** We introduce two new properties on an inconsistency index, that together with the four properties we discuss in the previous section characterize the Varian's index. In order to grasp the need of the first new property, Upper Consistency, note that the Varian's index is only concerned with those alternatives which are above the chosen one in the optimal preference relation. These are the alternatives that the individual ignores when making a choice. In other words, we need to focus on the upper contour sets of the chosen alternatives. Denote by $U_P^A(x)$ the upper contour set of alternative $x$ in $A$ with respect to the preference $P$, i.e., $U_P^A(x) = \{y \in A : yPx\}$.

**Upper Consistency (UC).** For every, $f, g, \mathbf{1}_{(A,x)}, \mathbf{1}_{(A,y)} \in \mathcal{F}$ and every $r, r' \in \mathcal{R}$, if $f$ and $f + \mathbf{1}_{(A,x)}$ are $r$-invariant, $g$ and $g + \mathbf{1}_{(A,y)}$ are $r'$-invariant, $U_{P^r}^A(x) = U_{P^{r'}}^A(y)$, and $I(f) = I(g)$ then $I(f + \mathbf{1}_{(A,x)}) = I(g + \mathbf{1}_{(A,y)})$.

Upper Consistency imposes conditions on when one can be sure that the addition of two observations, $(A, x)$ and $(A, y)$, to two collections of observations, $f$ and $g$, respectively, has no consequences on the evaluation of the inconsistency levels associated to the resulting two collections of observations, $f + \mathbf{1}_{(A,x)}$ and $g + \mathbf{1}_{(A,y)}$. First, both $f$ and $f + \mathbf{1}_{(A,x)}$ and $g$ and $g + \mathbf{1}_{(A,y)}$ must be part of a common structure, since they must share the same rationalizable collection of observations, $r$ and $r'$, respectively. Second, the upper contour sets of the chosen alternatives $x$ and $y$ in the menu $A$ with respective preferences $P^r$ and $P^{r'}$ must be identical. If in addition to these conditions, the inconsistency index judges $f$ and $g$ as equally inconsistent, then Upper Consistency imply that the inconsistency levels associated to $f + \mathbf{1}_{(A,x)}$ and $g + \mathbf{1}_{(A,y)}$ must also coincide. Note, therefore, that Upper Consistency is a weak property, since it imposes structure on the consequences of adding new data, only under a number of requirements that impose a great deal of similarity between the observations.

Our second new property for the characterization of the Varian's index aims to capture another main feature of the index. Note that among all the information contained in the upper contour set of the chosen alternative in an inconsistent observation, the Varian's index only considers the alternative with the highest associated inattention so that it is ignored in the revealed choice. Clearly, if such an alternative is ignored, all other alternatives in the upper contour set are also ignored. In order to introduce our second property, Preference Irrelevance, let $r_{x,a} \in \mathcal{R}$ denote the collection of rationalizable binary observations such that $x$ is always chosen whenever it is present, and $a$ is also always chosen whenever present except against $x$.

**Preference Irrelevance (PI).** For every $r\mathcal{R}$ and every $\mathbf{1}_{(A,a)} \in \mathcal{F}$ with $U_{P^r}^A(a) \neq \emptyset$, there exists a positive integer $z$ such that $I(\mathbf{1}_{(A,a)} + zr) = \max_{x \in U_{P^r}^A(a)} I(\mathbf{1}_{(A,a)} + zr_{x,a})$.

Take any observation $(A, a)$ and any rationalizable collection $r$ such that there is at least one alternative in $A$ that is preferred to the chosen alternative $a$ according to the preference relation $P^r$. That is, judged on the basis of $P^r$, the choice of $a$ in $A$ is inconsistent. Preference Irrelevance says that there always exist a large enough integer $z$ such that the inconsistency of $\mathbf{1}_{(A,a)} + zr$ is exactly the one associated to the inconsistency of the same observation $(A, a)$, associated the same number of times $z$, to the rationalizable collection of observations $r_{x,a}$. In the collection $r_{x,a}$ the maximal alternative $x$ in $A$ according to $P^r$ is always chosen whenever it is present and the same for $a$, except when $x$ is present. That is, according to Preference Irrelevance, given an observation $(A, a)$ and a rationalizable collection $r$, to evaluate the inconsistency of $\mathbf{1}_{(A,a)} + zr$ one can focus on the rationalizable collection $r_{x,a}$ that places the maximal alternative in $P^r$, alternative $x$, and the chosen alternative $a$, in the top two positions of a preference relation $P^{r_{x,a}}$.

Theorem 2 shows that these two properties together with Rationality, Invariance, Attraction and Separability uniquely characterize a positive scalar transformation of the Varian's index. The requirement of the positive scalar transformation is necessary to make the inconsistency values of certain collections of observations lie within the unit interval, and hence construct the attention mapping $\phi$, with image in the unit interval, on the basis of them.

**Theorem 2.** *An inconsistency index $I$ satisfies (RAT), (INV), (ATTR), (SEP), (UC) and (PI) if and only if $I$ is a positive scalar transformation of the Varian's index.*

We argued above that the Houtman-Maks' index $I_{HM}$, counting the minimal subset of the data that needs to be eliminated in order to make the remainder of the data consistent with the maximization principle, is but a special case of the Varian's index with attention mapping $\phi(A, e) = A$ for every $e > 0$. Consequently, the characterization of $I_{HM}$ must build on that of $I_V$, and impose some additional structure. Notice that the particular attention mapping of the Houtman-Maks' index treats all alternatives equally. That is, it does not discriminate between them, and hence any relabeling of the alternatives should have no effect on the level of inconsistency. Indeed, it turns out that we simply need to add a property imposing that the inconsistency index should be independent of the names of the alternatives. In order to formally introduce this notion, consider the following definition. Given a permutation $\sigma$ over the set of alternatives $X$, for any collection of observations $f$ we denote by $\sigma(f)$ the permuted collection of observations such that $\sigma(f)(A, a) = f(\sigma(A), \sigma(a))$.

**Neutrality (NEU).** For every permutation $\sigma$ and every $f \in \mathcal{F}$, $I(f) = I(\sigma(f))$.

The following corollary of Theorem 2, that we state without proof, makes formal the claim on $I_{HM}$.

**Corollary 1.** *An inconsistency index $I$ satisfies (RAT), (INV), (ATTR), (SEP), (UC), (PI) and (NEU) if and only if $I$ is a positive scalar transformation of the Houtman-Maks' index.*

4.2. **Minimal Swaps Index.** The minimal swaps index measures how many binary comparisons must be swapped in each observation in order to explain all the choices. Note, therefore, that for the minimal swaps index all the relevant information in each observation is again collected in those alternatives that, endogenously, become better than the chosen one. Hence, Upper Consistency, the property we introduced above with the purpose of focusing on the upper contour sets is needed in the case of the minimal swaps index. Also, note that unlike the Varian's index but in line with the Houtman-Maks' index, the minimal swaps index treats all the alternatives symmetrically. That is, for the minimal swaps index the identity of the alternatives is inessential. Hence, the property of Neutrality above introduced must be satisfied too. Preference Irrelevance, on the other hand, is clearly violated by the minimal swaps index since now all the alternatives in the upper contour set not only become equally relevant, but are considered additively. Consequently, what is needed here is another property imposing structure on when two observations can be merged with no consequences on the inconsistency value, and that it respects the importance of the number of alternatives in the upper contour sets. One can capture this feature with the following property.

**Disjoint Composition (DC).** For every, $f, \mathbf{1}_{(A,x)}, \mathbf{1}_{(B,x)} \in \mathcal{F}$ and every $r \in \mathcal{R}$, if $f + \mathbf{1}_{(A,x)} + \mathbf{1}_{(B,x)}$ is $r$-invariant and $A \cap B = \{x\}$ then $I(f + \mathbf{1}_{(A,x)} + \mathbf{1}_{(B,x)}) = I(f + \mathbf{1}_{(A \cup B,x)})$.

Disjoint Composition establishes that under very special circumstances, two observations can be merged into one without affecting the inconsistency level. Take two observations $(A, x)$ and $(B, x)$ that share the same chosen alternative $x$ and nothing else. That is, other than $x$ the two menus $A$ and $B$ are disjoint. Suppose that these two observations are added to a collection $f$ resulting in a collection sharing a common structure represented by the rationalizable collection $r$. Then, Disjoint Composition implies that the observations $(A, x)$ and $(B, x)$ can be merged into one single observation respecting the choice of $x$, $(A \cup B, x)$, with no consequences on the value of the inconsistency.

Theorem 3 states the characterization result. In this case the requirement of the positive scalar transformation is necessary to guarantee that the weight of every alternative in every menu of alternatives placed above the chosen one according to any preference relation gets a value of 1.

**Theorem 3.** *An inconsistency index $I$ satisfies (RAT), (INV), (ATTR), (SEP), (UC), (NEU) and (DC) if and only if it is a positive scalar transformation of the minimal swaps index.*

4.3. **Minimal Loss Index.** The main distinctive feature of minimal loss indices is that they care not only about the alternatives that are above the chosen one in the closest preference relation, but about the ranking of all these alternatives. This immediately implies that Upper Consistency and Preference Irrelevance are not satisfied since they ignore this type of information. Furthermore, minimal loss indices also fail to satisfy Disjoint Composition since merging two observations of the type dictated by Disjoint Composition reduces the inconsistency in the minimal loss indices, as the loss associated to the merged observation can indeed be tracked down to one of the observations; the one in which the maximal alternative according to the optimal preference is located. Instead of Disjoint Composition, what is needed here is another property establishing when and how two observations can be merged and that incorporates information on the ranking of alternatives. In order to introduce such a property, Composition, let us denote by $APB$ the case where all the alternatives in menu $A$ are preferred by $P$ to all the alternatives in $B$.

**Composition (COM).** For every, $f, \mathbf{1}_{(A,x)}, \mathbf{1}_{(B\cup\{x\},y)} \in \mathcal{F}$ and every $r \in \mathcal{R}$, if $f + \mathbf{1}_{(A,x)} + \mathbf{1}_{(B\cup\{x\},y)}$ is $r$-invariant and $AP^rB$ then $I(f + \mathbf{1}_{(A,x)} + \mathbf{1}_{(B\cup\{x\},y)}) = I(f + \mathbf{1}_{(A\cup B,y)})$.

Composition implies that to the inconsistency index $I$ it is equivalent to observe a given inconsistency at once, or separated into two very particular hypothetical stages. That is, suppose that the collection of observations formed by $(A, x)$, $(B \cup \{x\}, y)$ and $f$ share a common structure such that are $r$-invariant, and that all the alternatives in $A$ are above all the alternatives in $B$ according to the preference $P^r$. Now, from $(A, x)$ it may be the case that the chosen element $x$ is different from the maximal element in $A$ according to $P^r$. Since all the alternatives in $A$ are better according to $P^r$ than all the alternatives in $B$, it follows that $x$ is the maximal alternative from the menu $B \cup \{x\}$. Again, it may be the case that the chosen alternative in this menu, $y$, is different from the maximal alternative, $x$. If, on the other hand, all of the alternatives $A \cup B$ are presented simultaneously and the choice is $y$, Composition implies that the two possible inconsistencies in the previous scenario are equivalent to the one in the latter. Note that while Disjoint Composition and Composition share a common intuition, both properties are independent. Neither one implies the other. Indeed, it is easy to see that all the previous indices fail to satisfy this property.

Since what matters to the minimal loss index is the ranking of the alternatives but not the their nature, it clearly follows that Neutrality is satisfied. Theorem 4 establishes that these two properties, Neutrality and Composition, on top of the axiomatic structure of minimal indices, characterize the minimal loss indices.

**Theorem 4.** *An inconsistency index $I$ satisfies (RAT), (INV), (ATTR), (SEP), (NEU) and (COM) if and only if it is a minimal loss index.*

## 5. Individual Welfare

Welfare analysis with inconsistent individuals is problematic, since there is no preference relation compatible with all the revealed choices of the individual. There are two main approaches to the question of how to extract welfare relevant information in these situations. The first approach involves a choice model-free view, while the second builds on a particular model of decision-making consistent with the behavior of the individual.

In the choice model-free approach, the analyst is agnostic about the underlying model the individual may follow and focus exclusively on the observable choice data. That is, no particular assumption on the sources of the inconsistency of the choices is done, and revealed choices are taken as the unique source of information to infer individual welfare. This is the main perspective adopted by Bernheim and Rangel (2009). Alternatively, the analyst may assume a particular choice model consistent with the revealed choices of the individual, and use this model in order to infer individual welfare. The underlying assumption is that the causes of the inconsistencies may provide valuable information about how to fix the welfare ranking. This is the approach taken by Green and Hojman (2009).[6]

Interestingly, Bernheim and Rangel (2009) and Green and Hojman (2009), following these two different approaches, independently suggest the use of the same welfare notion. Let us denote by $\overline{P}$ the Bernheim-Rangel-Green-Hojman preference, defined as $x\overline{P}y$ if and only if there is no observation $(A, y)$ with $x \in A$ such that $f(A, y) > 0$. In other words, $y$ is never chosen when $x$ is available. Bernheim and Rangel show that whenever every menu $A$ in $X$ is observed at least once, $\overline{P}$ is acyclic, and hence it is consistent with the maximization principle.

We now wonder which is the relationship between $\overline{P}$ and the optimal preference relation of a minimal index, $P^*$. It turns out to be the case that the two welfare relations are fundamentally different. That $P^*$ is not contained in $\overline{P}$ follows immediately, since $P^*$ is a linear order, while $\overline{P}$ is acyclic and hence incomplete in general. In the other direction, and more importantly, note that while $\overline{P}$ evaluates the ranking of two alternatives $x$ and $y$ by attending exclusively to those menus of alternatives where both $x$ and $y$ are available, to establish the ranking of two alternatives $P^*$ takes into consideration all the observations together. Hence, $P^*$ and $\overline{P}$ may rank two alternative in opposite directions. A simple example illustrates this point.

Consider a collection $f$ composed by the following observations: $f(\{x, y\}, x) = f(\{y, z\}, y) = 3$ and $f(\{x, y, z\}, y) = f(\{x, z\}, z) = 1$. Clearly, $z\overline{P}x$ since $x$ is never chosen in the presence of $z$. However, to evaluate the ranking of alternatives $x$ and $z$, the minimal index considers the whole collection $f$. Observations $f(\{x, y\}, x) = f(\{y, z\}, y) = 3$ signify a strong argument for the preference $xPyPz$. This preference implies a mistake in the observations $f(\{x, y, z\}, y) = f(\{x, z\}, z) = 1$, but rationalizes the more frequent evidence of $f(\{x, y\}, x) = f(\{y, z\}, y) = 3$. Preference $P$ is in fact the optimal preference relation for the minimal swaps index and for the minimal loss indices, independently of the vector of weights $u$. The argument can be extended to any minimal index.

**Theorem 5.** *For every weighting mapping $w$, there exists a collection of observations $f$ for which the optimal preference of the minimal index is not an extension of the Bernheim-Rangel-Green-Hojman preference.*

[6]Other proposals within this second approach are Köszegi and Rabin (2007), Masatlioglu, Nakajima and Ozbay (2010), and Rubinstein and Salant (2011). See Green and Hojman (2009, section 4.2) for a discussion of the two approaches.

We close this section with a discussion on the interpretation of minimal indices, in light of the two approaches mentioned above. Our view here is that minimal indices are flexible enough to be compatible with the two approximations to behavioral welfare analysis. On the one hand, the analyst may adopt the weights $w(P, A, a)$ that she may consider the most appropriate from an axiomatic approach without adhering to a specific model of choice. This approach is in line with the choice model-free approach. Alternatively, the analyst may have in mind a particular model of choice and this may dictate the weights $w(P, A, a)$ to be considered. For example, the minimal loss index is consistent with a model of decision making where the individual has a clear ranking of the alternatives in mind $u$, but in occasions the available alternatives may experience shocks that makes the individual to select an ex-ante suboptimal alternative. Then, this particular model of choice would suggest the preference that minimizes the minimal index, since this is ex-ante the best way to accommodate individual welfare.

## 6. Computation of the Optimal Preference Relation

We now deal with the question of obtaining in practice the preference relation that minimizes the inconsistency associated to a collection of observations. Given that we have imposed no restriction whatsoever on the nature of the collections of observations, it is not surprising that finding the optimal preference relation may be computationally complex. Fortunately, we can draw upon existing techniques that address computational problems that can be related to ours, and offer reasonably good solutions. In this section we briefly describe two of such techniques. We further suggest a third method that involves the restriction of the data sets to special domains.

**Globally-Optimal Techniques.** The first approach we stress here involves the search for the globally optimal preference relation. Clearly, this may involve the use of computationally heavy techniques, but that are orders of magnitude superior to the naive approach of evaluating all the possible linear orders.

Houtman and Maks (1984) offer one such algorithm for the computation of the maximal subset of the data consistent with the maximization of a preference relation. It follows immediately that if such a subset is found, the computation of the preference relation consistent with it is very simple. Dean and Martin (2010) show that the minimum set covering problem (MSCP) is equivalent to the problem at hand, and then argue that one can use the wide variety of algorithms that have been designed in the operations research literature to solve the MSCP. This allows Dean and Martin to offer a powerful algorithm.

The same strategy can be adopted for other cases of minimal indices. To illustrate, consider the well-known linear ordering problem (LOP). Formally, the integer LOP problem over the set of vertices $Y$ and directed weighted edges that connect all vertices $x$ and $y$ with (integer) cost $c_{xy}$, consists of finding the linear order relation over the set of vertices that minimizes the total aggregated cost. That is, if we denote by $\Pi$ the set of all mappings from $Y$ to $\{1, 2, \ldots, |Y|\}$, the LOP involves solving $\arg\min_{\pi \in \Pi} \sum_{\pi(x)<\pi(y)} c_{xy}$.

The LOP has been related to a wide variety of problems, including various economic problems, particularly to the triangularization of input-output matrices for the study of the hierarchical structures of the productive sectors in an economy. The literature offers a good array of algorithms to solve this problem (see, e.g., Grötschel, Jünger, and Reinelt (1984); see also Chaovalitwongse et al (2010) for a good introduction to the LOP, a review of the algorithmic literature solving the LOP, and for the analysis of one such algorithm). As the following result shows, we can directly use the algorithms developed to solve the LOP to compute the optimal preference relation for the minimal swaps index.

**Theorem 6.**

(1) *For every $f \in \mathcal{F}$ one can define a LOP with vertices in $X$, the solution of which provides the optimal preference for the minimal swaps index.*

(2) *For any LOP with vertices in $X$ one can define an $f \in \mathcal{F}$, its optimal preference being the solution to the LOP.*

Intuitively, the linear orders in the LOP are the preference relations in our setting and the cost of having one alternative before the other is the inconsistency that arises from revealed data. Note that the evaluation of the inconsistency associated to having one alternative $a$ over the other $b$ is very simple for the case of the minimal swaps index. It involves counting the number of observations where the chosen alternative is $b$ but $a$ is present in the menu. Then, the computation of the optimal preference relation requires attending all the inconsistency values associated to having one alternative over the other, exactly like in the LOP. This makes that the algorithms that the literature has offered to solve the LOP, or equivalently the triangularization of input-output matrices, can be directly used in the case of the minimal swaps index.

Although this approach can also be used for the case of the minimal loss index or the Varian's index, the analysis becomes more cumbersome. When considering the inconsistency associated to having one alternative above the other in the ranking, in the minimal loss index it is necessary to evaluate whether $a$ is not only present but maximal in the menu, while in the Varian's index it is necessary to consider the attention levels to determine whether alternative $a$ is observed at each stage.

**Iterative Improvement Strategies.** Algorithms computing the globally optimal solution may be, on occasions, too heavy computationally speaking. For this reason the literature also offers methods that whilst they do not compute the globally optimal solution are much lighter in terms of the required computational intensity, giving good approximations. The LOP is not an exception and there is a variety of techniques searching for locally optimal solutions. The iterative improvement strategies represent one prominent example (see Brusco, Kohn and Stahl (2008) for a good introduction and relevant references). This technique has been intensively used in the problem of the triangularization of input-output matrices (see, e.g., Korte and Oberhofer 1970, Fukui 1986, and Howe 1991).

The main idea of this technique in the context of our setting involves the consideration of an initial preference $P$ and its associated inconsistency value. Then, it consists of the interchange of sets of consecutive alternatives in the original preference $P$, resulting into a new preference $P'$ with a lower inconsistency value. This process is iterated until no more interchanges further improve the explanation of the data. Clearly, exactly the same considerations with regard to the minimal swaps, the minimal loss and the Varian's index we argue above for the globally optimal techniques apply here.

**Efficient Computations in Relevant Domains.** A third approach is to consider special choice domains where one can identify the optimal preference relation efficiently. We illustrate this methodology by suggesting a domain of data sets that although it is general enough to include the classical domains considered in the literature and many others, it has a special structure that allows the computation of the optimal preference relation for the minimal loss index practically costless.

We say that a collection $f$ is balanced if all the menus of alternatives of the same cardinality are observed the same number of times. That is, $|A| = |B| \Rightarrow \sum_{x\in A} f(A,x) = \sum_{y\in B} f(B,y)$. The class of balanced collections of observations encompasses classical domains, such as the universal domain, where all possible menus are observed once, or the binary domain, where all possible menus of two alternatives are observed once, etc. It also includes replicas of these domains, thus enabling the study of variability in choices from the same menu.

We show below that in these domains, for the case of the minimal loss index, the preference relation that places the most frequently chosen alternatives higher in the ranking is the optimal one. We call such preference relations basic. Formally, given a collection $f$, a basic preference relation $P^B(f)$ is any preference relation such that $\sum_{(A,a):a=x} f(A,a) > \sum_{(A,a):a=y} f(A,a) \Rightarrow xP^B(f)y$. Clearly, such a preference relation is extremely easy to compute. Hence, the following result is of substantial interest.

**Theorem 7.** *For any balanced collection of observations $f$, and for any vector $u$, $P$ is an optimal preference relation for the minimal loss index if and only if $P$ is a basic preference relation for $f$.*

## 7. Conclusions

It is now widely accepted that individual behavior often deviates from the maximization principle. It is therefore crucial to assign a proper meaning to such deviations, since this would not only give a sense of the reliability of classical choice theory in different environments, but would also provide useful information to theorists for the future development of the field. In this paper we identify from a foundational analysis a particular measure of the rationality of individuals, the class of minimal indices. We argue that minimal indices have several attractive characteristics. Minimal indices evaluate the inconsistency of the data by identifying the preference relation that is closest to observed behavior. For this purpose, it additively evaluates all the observations where the individual's revealed choice does not correspond to the maximization of the identified preference relation. The weight of each inconsistent choice may depend on the nature of the available menu of alternatives, the chosen alternative and the preference under consideration.

Given the generality of the weighting of the inconsistencies, minimal indices represent a broad class of inconsistency measures. We explore in detail various special cases, in particular the well-known Varian's index, and two new indices that we suggest here, the minimal swaps index and the minimal loss index. Each one of these indices, although sharing the structure of minimal indices, are fundamentally different and approach the problem of measuring the inconsistency of the data from different angles. In order to provide a deep understanding of each of these three indices, we provide an axiomatic characterization of them.

Our measure of rationality provides a tool for the conduction of welfare analysis even with inconsistent decision makers. Clearly, if there is no preference relation compatible with all the revealed choices, classical welfare analysis is a challenge. We suggest the preference relation minimizing the minimal index as a sound tool for this purpose. The main advantages of this are that it is the closest preference relation to the data, it is well behaved in the sense of identifying a single alternative in any possible situation, and it is axiomatically founded.

We see several avenues for future research. On the theoretical side, it is important to elucidate the power of minimal indices, in the sense of providing the measure of inconsistency contingent on the maximum degree attainable (see Bronars (1987) and Beatty and Crawford (2010) for treatments of specific measures). On the empirical side, one would like to see the application of minimal indices for the actual computation of inconsistency in various settings.

## Appendix A. Proofs

**Proof of Proposition 1:** Let $\phi$ be an attention mapping and $I_V$ the associated Varian's index. Define for every preference relation $P$ and every observation $(A, a)$ the weight $w(P, A, a) = 1 - \sup_{U_P^A(a) \cap \phi(A,e)=\emptyset} e$. The weight $w(P, A, a)$ is therefore determined by the largest value of $e$ at which no alternative that dominates $a$ according to $P$ is observed by the individual. We now show that for every collection of observations $f$, it is $I_V(f) = \min_{P \in \mathcal{P}} \sum_{(A,a)} f(A, a)w(P, A, a)$.

First, let $\mathbf{e}$ be a vector of attention levels such that (1) $I_V(f) = \sum_{(A,a)} f(A, a)(1 - e_{(A,a)})$, and (2) $R^{\phi}_{f,\mathbf{e}'}$ is acyclic for every vector of attention levels $\mathbf{e}'$ with $e'_{(A,a)} \leq e_{(A,a)}$, where there is at least an $(A, a)$ with $f(A, a) > 0$ and $e_{(A,a)} > 0$ for which the inequality is strict. Consider a preference relation $P^*$ that is a linear extension of all the acyclic preferences $R^{\phi}_{f,\mathbf{e}'}$. Given the definition of the Varian's index, $\mathbf{e}$ exists and $P^*$ is well-defined.

We now prove that $I_V(f) = \sum_{(A,a)} f(A, a)w(P^*, A, a)$. We need to show that for every $(A, a)$ such that $f(A, a) > 0$, it must be $1 - e_{(A,a)} = w(P^*, A, a) = 1 - \sup_{U_{P^*}^A(a) \cap \phi(A,e)=\emptyset} e$. Suppose otherwise. If $1 - e_{(A,a)} < 1 - \sup_{U_{P^*}^A(a) \cap \phi(A,e)=\emptyset} e$, we can consider a vector $\mathbf{e}'$ equal to $\mathbf{e}$ in all its components except that $e_{(A,a)} > e'_{(A,a)} > \sup_{U_{P^*}^A(a) \cap \phi(A,e)=\emptyset} e$. Hence, there exists an element $x$ with $xP^*a$ that is observed in menu $A$ for level of attention $e'_{(A,a)}$, implying that $aR^{\phi}_{f,\mathbf{e}'}x$. This contradicts the definition of $P^*$. Suppose, then, that there is an observation $(A, a)$ in $f$ such that

$1-e_{(A,a)} > 1-\sup_{U^A_{P*}(a)\cap\phi(A,e)=\emptyset} e$. Then we can consider a vector $\mathbf{e}'$ equal to $\mathbf{e}$ in all its components except that $e_{(A,a)} < e'_{(A,a)} < \sup_{U^A_{P*}(a)\cap\phi(A,e)=\emptyset} e$. The observation $(A,a)$ can only reveal alternative $a$ to be preferred to alternatives below $a$ in $P^*$, if any. This contradicts the optimality of $\mathbf{e}$ in the Varian's index.

Second, we prove that, for every $P \in \mathcal{P}$, $I_V(f) \leq \sum_{(A,a)} f(A,a)w(P,A,a)$. Suppose otherwise and let $P'$ be a preference relation such that $I_V(f) > \sum_{(A,a)} f(A,a)w(P',A,a)$. Define the vector of attention levels $\mathbf{e}'$ by $e'_{(A,a)} = 1-w(P',A,a)$. Consider andy vector of attention levels $\mathbf{e}''$ with $e''_{(A,a)} \leq e'_{(A,a)}$, where there is at least an $(A,a)$ with $f(A,a) > 0$ and $e'_{(A,a)} > 0$ for which the inequality is strict. Then, for every such $\mathbf{e}''$ the preference $R^{\phi}_{f,\mathbf{e}''}$ only contains information of the following form: the chosen alternative $a$ in $A$ is revealed preferred to alternatives in $A$ that are below $a$ according to $P'$. Obviously, all these revealed preferences can be extended to $P'$. Hence, $\inf_{\mathbf{e}:R^{\phi}_{f,\mathbf{e}} \text{ is acyclic}} \sum_{(A,a)} f(A,a)(1-e_{(A,a)}) \leq \sum_{(A,a)} f(A,a)w(P',A,a) < I_V(f)$ a contradiction with the definition of the Varian's index.■

**Proof of Theorem 1:** It is easy to see that any minimal index $I_M$ satisfies the axioms. We prove the converse statement by way of five lemmas.

**Lemma 1.** *For every $r \in \mathcal{R}$, $f \in \mathcal{F}$, and $z \in \mathbb{Z}_{++}$, (1) $zr$ is $r$-invariant, and (2) if $f$ is $r$-invariant, then $f+zr$ is $r$-invariant.*

**Proof of Lemma 1:** Consider any $r \in \mathcal{R}$, $f \in \mathcal{F}$, and $z \in \mathbb{Z}_{++}$. By (RAT), we know that $I(zr) = I((z+1)r) = 0$. Hence, $zr$ is $r$-invariant. This shows the first claim. Now, since $zr$ and $(z+1)r$ are $r$-invariant, if $f$ is $r$-invariant (SEP) implies that $I(f+zr) = I(f)+I(zr)$ and $I(f+(z+1)r) = I(f)+I((z+1)r)$. By (RAT), we know that $I(zr) = I((z+1)r) = 0$. Hence, $I(f+zr) = I(f) = I(f+(z+1)r)$ and then $f+zr$ is $r$-invariant. This shows the second claim.□

**Lemma 2.** *For every $r \in \mathcal{R}$ and $n \in \mathbb{Z}_{++}$, there exists a positive integer $z^r_n$ such that if $\sum_{(A,a)} f(A,a) \leq n$, then $f+z^r_n r$ is $r$-invariant.*

**Proof of Lemma 2:** Take any $r \in \mathcal{R}$ and $n \in \mathbb{Z}_{++}$. Define $z^r_n = \max_{f:\sum_{(A,a)} f(A,a)\leq n} z^r_f$ where $z^r_f$ is the minimal value in $\mathbb{Z}_{++}$ such that $f+z^r_f r$ is $r$-invariant. Notice that (ATTR) guarantees that $z^r_f$ always exists and the finiteness of $X$ guarantees that $z^r_n$ is well-defined. Then, if $\sum_{(A,a)} f(A,a) \leq n$, it follows that $z^r_n \geq z^r_f$ and by Lemma 1, $f+z^r_n r$ is $r$-invariant, as desired.□

**Lemma 3.** *For every $r \in \mathcal{R}$ and $n \in \mathbb{Z}_{++}$, if $f$ is $r$-invariant and $\sum_{(A,a)} f(A,a) = n$, then $I(f) = \sum_{(A,a)} f(A,a) I(\mathbf{1}_{(A,a)} + z_n^r r)$.*

**Proof of Lemma 3:** Consider any $r \in \mathcal{R}$ and $n \in \mathbb{Z}_{++}$. Let $f$ be $r$-invariant with $\sum_{(A,a)} f(A,a) = n$. Then, for any observation $(A,a)$ with $f(A,a) > 0$, the collections of observations $f - \mathbf{1}_{(A,a)}$ and $\mathbf{1}_{(A,a)}$ have both less than $n$ observations. By Lemmas 1 and 2, the collections of observations $f - \mathbf{1}_{(A,a)} + (n-1)z_n^r r$ and $\mathbf{1}_{(A,a)} + z_n^r r$ are both $r$-invariant. Hence, (SEP) guarantees that $I(f + nz_n^r r) = I(f - \mathbf{1}_{(A,a)} + (n-1)z_n^r r) + I(\mathbf{1}_{(A,a)} + z_n^r r)$. The repeated application of this argument leads to $I(f + nz_n^r r) = \sum_{(A,a)} f(A,a) I(\mathbf{1}_{(A,a)} + z_n^r r)$. Since $f$ is $r$-invariant, the repeated use of Lemma 1 guarantees that $I(f) = I(f + nz_n^r r)$ and hence, $I(f) = \sum_{(A,a)} f(A,a) I(\mathbf{1}_{(A,a)} + z_n^r r)$, as desired.$\square$

Now define the following weighting function. Given any $P \in \mathcal{P}$, consider the rationalizable collection $r \in \mathcal{R}$ such that $P = P^r$ and define $w(P,A,a) = 0$ if $a = m(P,A)$, and $w(P,A,a) = I(\mathbf{1}_{(A,a)} + z_1^r r)$ otherwise.

**Lemma 4.** *If $f$ is $r$-invariant, then $I(f) = \sum_{(A,a)} f(A,a) w(P^r, A, a)$.*

**Proof of Lemma 4:** Consider a collection of observations $f$ such that $f$ is $r$-invariant with $\sum_{(A,a)} f(A,a) = n$. By Lemma 3, $I(f) = \sum_{(A,a)} f(A,a) I(\mathbf{1}_{(A,a)} + z_n^r r)$. By (RAT), if $a = m(P^r, A)$, then $I(\mathbf{1}_{(A,a)} + z_n^r r) = 0 = w(P^r, A, a)$. Otherwise, by definition of $z_1^r$, we know that $\mathbf{1}_{(A,a)} + z_1^r r$ is $r$-invariant. Since by construction $z_1^r \leq z_n^r$, Lemma 1 guarantees that $w(P^r, A, a) = I(\mathbf{1}_{(A,a)} + z_1^r r) = I(\mathbf{1}_{(A,a)} + z_n^r r)$. Hence, $I(f) = \sum_{(A,a)} f(A,a) w(P^r, A, a)$, as desired.$\square$

**Lemma 5.** *For every $f \in \mathcal{F}$, $P \in \mathcal{P}$, $I(f) \leq \sum_{(A,a)} f(A,a) w(P,A,a)$.*

**Proof of Lemma 5:** Consider any $f \in \mathcal{F}$ and $P \in \mathcal{P}$. Let $r'$ such that $P^{r'} = P$. By (ATTR) there exists $z \in \mathbb{Z}_{++}$ such that $f + zr'$ is $r'$-invariant. By Lemma 4, $I(f + zr') = \sum_{(A,a)} (f + zr')(A,a) w(P,A,a) = \sum_{(A,a)} f(A,a) w(P,A,a)$. By (SEP) and (RAT), $I(f) = I(f) + I(zr') \leq I(f + zr')$. Hence, $I(f) \leq \sum_{(A,a)} f(A,a) w(P,A,a)$ as desired.$\square$

Lemmas 4 and 5 together imply that $I$ is a minimal index.$\blacksquare$

**Proof of Theorem 2:** It is easy to see that any positive scalar transformation of a Varian's index $I_V$ satisfies the axioms. To prove the converse, suppose that an inconsistency index $I$ satisfies all the axioms.

By Theorem 1, we know that any inconsistency index satisfying (RAT), (INV), (ATTR) and (SEP) is a minimal index. Let $\Psi$ denote the maximal value of $I(\mathbf{1}_{(A,a)} + zr_{x,a})$ across all observations $(A, a)$ and collections $r_{x,a}$ with $z$ such that $\mathbf{1}_{(A,a)} + zr_{x,a}$ is $r_{x,a}$-invariant. Clearly, $\Psi$ is a well-defined positive real number. Define the attention mapping $\phi$ by $\phi(A, 0) = \emptyset$ and for every $e > 0$ $\phi(A, e) = \{x \in A : \exists a \in A \setminus \{x\}$ such that $1 - I(\mathbf{1}_{(A,a)} + zr_{x,a})/\Psi \leq e\}$ for some $a \in A$. Clearly, $\phi$ is well-defined. By (UC), this definition is independent of the chosen element $a$ and the specific rationalizable collection of observations $r_{x,a}$.

We can construct the values $w(P, A, a)$ by looking at $I(\mathbf{1}_{(A,a)} + zr)$ where $P^r = P$ and $z$ large enough. If $a = m(P, A)$, we know that $w(P, A, a) = 0$. Otherwise, by (PI), $w(P, A, a) = \max_{x \in U_P^A(a)} I(\mathbf{1}_{(A,a)} + zr_{x,a})$. Given the above definition of $\phi$, one can directly rewrite $w(P, A, a) = \Psi(1 - \sup_{U_P^A(a) \cap \phi(A,e) = \emptyset} e)$. By Proposition 1, this construction is a positive scalar transformation of the Varian's index.■

**Proof of Theorem 3:** It is easy to see that any positive scalar transformation of the minimal swaps index satisfies the axioms. To prove the converse, suppose that an inconsistency index satisfies the axioms. For any collection of observations $f$, we know from Theorem 1 that $I(f) = \min_{P \in \mathcal{P}} \sum_{(A,a)} f(A, a)w(P, A, a)$. Consider any observation $(A, a)$ with $f(A, a) > 0$, and the rationalizable collection $r$ such that $P^r = P$. Recall from the proof of Theorem 1 that $w(P, A, a) = I(\mathbf{1}_{(A,a)} + z_1^r r)$. The repeated application of (DC), (SEP) and Lemma 1 leads to $w(P, A, a) = I(\mathbf{1}_{(A,a)} + |A - 1|z_1^r r) = \sum_{x \in A \setminus \{a\}} I(\mathbf{1}_{(\{x,a\},a)} + z_1^r r)$. By (RAT) and Theorem 1, this is equivalent to $w(P, A, a) = \sum_{x \in A : xP^r a} I(\mathbf{1}_{(\{x,a\},a)} + z_1^r r) = \sum_{x \in A : xP^r a} w(P, \{x, a\}, a)$.

We now prove that $w(P, \{x, a\}, a) = w(P', \{y, b\}, b)$ whenever $xPa$ and $yP'b$. (ATTR) and (UC) imply that $w(P, \{x, a\}, a) = w(P^{r_{x,a}}, \{x, a\}, a)$, and that $w(P', \{y, b\}, b) = w(P^{r_{y,b}}, \{y, b\}, b)$. (NEU) implies that $w(P^{r_{x,a}}, \{x, a\}, a) = w(P^{r_{y,b}}, \{y, b\}, b)$. Hence, $w(P, \{x, a\}, a) = w(P', \{y, b\}, b)$ and we can write $w(P, A, a) = \sum_{x \in A : xP^r a} w(P, \{x, a\}, a) = \Theta|\{x \in A : xP^r a\}|$, with $\Theta > 0$. This clearly represents a positive scalar transformation of the minimal swaps index.■

**Proof of Theorem 4:** It is easy to see that minimal loss indices satisfy the axioms, to prove the converse statement assume that these hold. From Theorem 1, we are left to prove that there exists a vector of real-valued weights $u = (u_1, \ldots, u_k)$ with $u_1 > u_2 > \cdots > u_k$ such that for every $P \in \mathcal{P}$ and every $(A, a)$ with $a \neq m(P, A)$,

$w(P,A,a) = u_{\widehat{m}(P,A)} - u_{\widehat{a}(P)}$. Let the rationalizable collection $r$ be such that $P^r = P$. Let $A_1 = \{x \in A : xP^r a\}$ and $A_2 = \{m(P^r, A), a\} \cup \{x \in A : aP^r x\}$. By (COM), (SEP), (RAT), Lemma 1 and the definition of $z_n^r$ we can conclude that $I(\mathbf{1}_{(A,a)} + 3z_2^r r) = I(\mathbf{1}_{(A_1,m(P^r,A))} + \mathbf{1}_{(A_2,a)} + 3z_2^r r) = I(\mathbf{1}_{(A_1,m(P^r,A))} + z_2^r r) + I(\mathbf{1}_{(A_2,a)} + 2z_2^r r) = I(\mathbf{1}_{(A_2,a)} + 2z_2^r r)$. Now, if $A_2$ is composed by only two elements we know from Lemma 1 that $I(\mathbf{1}_{(A_2,a)} + z_2^r r) = I(\mathbf{1}_{(A_2,a)} + (z_2^r + 1)r) = \cdots = I(\mathbf{1}_{(A_2,a)} + 2z_2^r r)$. If $A_2$ is composed by three elements or more, consider the binary set $A_3 = \{m(P^r, A), a\}$ and the set $A_4 = \{a\} \cup \{x \in A : aP^r x\}$. By (COM), (SEP) and (RAT) it follows that $I(\mathbf{1}_{(A_2,a)} + 2z_2^r r) = I(\mathbf{1}_{(A_3,a)} + \mathbf{1}_{(A_4,a)} + 2z_2^r r) = I(\mathbf{1}_{(A_3,a)} + z_2^r r) + I(\mathbf{1}_{(A_4,a)} + z_2^r r) = I(\mathbf{1}_{(A_3,a)} + z_2^r r)$. Hence, in any case, $w(P,A,a) = I(\mathbf{1}_{(A,a)} + 3z_2^r r) = I(\mathbf{1}_{(\{m(P^r,A),a\},a)} + z_2^r r) = w(P, \{m(P^r, A), a\}, a)$.

We now prove that for any three alternatives $b, c$ and $d$ such that $bPcPd$, we have $w(P,\{b,c\},c) + w(P,\{c,d\},d) = w(P,\{b,c,d\},d)$. To see this, consider a rationalizable collection of observations $r$ such that $P^r = P$. Now, (COM) and (SEP) guarantee that $w(P,\{b,c,d\},d) = I(\mathbf{1}_{(\{b,c,d\},d)} + 2z_2^r r) = I(\mathbf{1}_{(\{b,c\},c)} + \mathbf{1}_{(\{c,d\},d)} + 2z_2^r r) = I(\mathbf{1}_{(\{b,c\},c)} + z_2^r r) + I(\mathbf{1}_{(\{c,d\},d)} + z_2^r r) = w(P,\{b,c\},c) + w(P,\{c,d\},d)$. Hence, we can define weights $u_{P,x}$, such that $xPy$ implies that $u_{P,x} > u_{P,y}$, and $w(P,A,a) = u_{P,m(P,A)} - u_{P,a}$. (NEU) guarantees that these weights only depend on the position that the chosen alternative $a$ occupies in $P$, and hence $w(P,A,a) = u_{\widehat{m}(P,A)} - u_{\widehat{a}(P)}$, as desired.■

**Proof of Theorem 5:** Consider the collection of observations $f$ defined by $f(\{x,y\},x) = f(\{y,z\},y) = m \geq 1$ and $f(\{x,y,z\},y) = f(\{x,z\},z) = 1$. According to the Bernheim-Rangel-Green-Hojman preference $\overline{P}$, $z\overline{P}x$ since $x$ is never chosen in the presence of $z$. Now, note that for any preference $P'$ such that $zP'x$, it must be either $yP'x$ or $zP'y$ and hence $\sum_{(A,a)} f(A,a)w(P',A,a) \geq m \min\{w(P',\{x,y\},x), w(P',\{y,z\},y)\}$. For the preference $P$ given by $xPyPz$ we have $\sum_{(A,a)} f(A,a)w(P,A,a) = w(P,\{x,y,z\},y) + w(P,\{x,z\},z)$. For $m$ sufficiently large, it follows that $I_M(f) \leq \sum_{(A,a)} f(A,a)w(P,A,a) < \sum_{(A,a)} f(A,a)w(P',A,a)$ and hence, the optimal preference of the minimal index must place $x$ above $z$, in contradiction with the Bernheim-Rangel-Green-Hojman preference.■

**Proof of Theorem 6:** To see the first part, consider the collection of observations $f$ and define, for every pair of alternatives $x$ and $y$, the weight $c_{xy} = \sum_{(A,a):y=a,x\in A} f(A,a)$. Clearly, $\sum_{\pi(x)<\pi(y)} c_{xy} = \sum_{\pi(x)<\pi(y)} \sum_{(A,a):y=a,x\in A} f(A,a) = \sum_{(A,a)} f(A,a)|\{x \in A :$

$\pi(x) < \pi(a)\}|$, and hence solving the LOP provides the optimal preference for the minimal swaps index. To see the second part, consider the LOP given by weights $c$. Define the collection of observations $f$ given by $f(\{x,y\},y) = c_{xy}$. Since $f$ is defined over binary problems, $\sum_{(A,a)} f(A,a)|\{x \in A : \pi(x) < \pi(a)\}| = \sum_{(\{x,y\},y):\pi(x)<\pi(y)} f(\{x,y\},y) = \sum_{\pi(x)<\pi(y)} c_{xy}$, as desired.■

**Proof of Theorem 7:** Consider a balanced collection of observations $f$, and two preference relations $P$ and $P'$. Clearly,

$$\begin{aligned}
\sum_{(A,a)} f(A,a)(u_{\widehat{m}(P,A)} - u_{\widehat{a}(P)}) &\leq \sum_{(A,a)} f(A,a)(u_{\widehat{m}(P',A)} - u_{\widehat{a}(P')}) \Leftrightarrow \\
\sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{m}(P,A)=t}} f(A,a)u_t - \sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{a}(P)=t}} f(A,a)u_t &\leq \sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{m}(P',A)=t}} f(A,a)u_t - \sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{a}(P')=t}} f(A,a)u_t
\end{aligned}$$

Notice that $\sum_{t=1}^{k} \sum_{(A,a):\widehat{m}(P,A)=t} f(A,a)u_t$ is independent of the chosen alternatives. Since $f$ is balanced, for each $A$ such that $\widehat{m}(P,A) = t$, there is one and only one $A'$ with the same cardinality than $A$ such that $\widehat{m}(P',A') = t$. Since these two menus are observed the same number of times, it follows that $\sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{m}(P,A)=t}} f(A,a)u_t = \sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{m}(P',A)=t}} f(A,a)u_t$. Hence:

$$\begin{aligned}
\sum_{(A,a)} f(A,a)(u_{\widehat{m}(P,A)} - u_{\widehat{a}(P)}) &\leq \sum_{(A,a)} f(A,a)(u_{\widehat{m}(P',A)} - u_{\widehat{a}(P')}) \Leftrightarrow \\
\sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{a}(P)=t}} f(A,a)u_t &\geq \sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{a}(P')=t}} f(A,a)u_t \Leftrightarrow \\
\sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{a}(P)=t}} f(A,a)(u_t - u_k) &\geq \sum_{t=1}^{k} \sum_{\substack{(A,a):\\ \widehat{a}(P')=t}} f(A,a)(u_t - u_k) \Leftrightarrow \\
\sum_{s=1}^{k-1} \sum_{l=1}^{s} \sum_{\substack{(A,a):\\ \widehat{a}(P)=l}} f(A,a)(u_s - u_{s+1}) &\geq \sum_{s=1}^{k-1} \sum_{l=1}^{s} \sum_{\substack{(A,a):\\ \widehat{a}(P')=l}} f(A,a)(u_s - u_{s+1}).
\end{aligned}$$

The definition of a basic preference relation $P$ guarantees that, for any $s$ and any other preference relation $P'$, $\sum_{l=1}^{s} \sum_{(A,a):\widehat{a}(P)=l} f(A,a) \geq \sum_{l=1}^{s} \sum_{(A,a):\widehat{a}(P')=l} f(A,a)$. Finally, since $u_s - u_{s+1} > 0$ for every $s$, the 'if' claim follows. Clearly, when $P'$ is non-basic, one of the former inequalities is strict, and hence the 'only-if' claim follows.■

## References


[1] Afriat, S.N. (1973), "On a System of Inequalities in Demand Analysis: An Extension of the Classical Method," *International Economic Review*, 14:460–472.

[2] Apesteguia, J., M.A. Ballester, and R. Ferrer (2011), "On the Justice of Decision Rules," *Review of Economic Studies*, 78:1–16.

[3] Baldiga, K.A. and J.R. Green (2010), "Choice-based Measures of Conflict in Preferences," mimeo.

[4] Banker, R.D. and A. Maindiratta (1988), "Nonparametric Analysis of Technical and Allocative Efficiencies in Production," *Econometrica*, 56:1315–1332.

[5] Beatty, T and I. Crawford (2010), "How Demanding is the Revealed Preference approach to Demand?," *American Economic Review*, forthcoming.

[6] Bernheim, B.D. and A. Rangel (2009), "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics," *Quarterly Journal of Economics*, 124:51–104.

[7] Bossert, W. and Y. Sprumont (2003), "Non-Deteriorating Choice," *Mathematical Social Sciences*, 45:131–142.

[8] Bronars, S.G. (1987), "The Power of Nonparametric Tests of Preference Maximization," *Econometrica*, 55:693–698.

[9] Brusco, M.J., H.F. Kohn, and S. Stahl (2008), "Heuristic Implementation of Dynamic Programming for Matrix Permutation Problems in Combinatorial Data Analysis," *Psychometrika*, 73:503–522.

[10] Chalfant, J. and J.M. Alston (1988), "Accounting for Changes in Tastes," *Joumal of Political Economy*, 96:391–410.

[11] Chambers, C.P. and T. Hayashi (2009), "Choice and Individual Welfare," mimeo.

[12] Chaovalitwongse, W.A., C.A.S. Oliveira, B. Chiarini, P.M. Pardalos, and M.G.C. Resende (2010), "Revised GRASP with Path-Relinking for the Linear Ordering Problem," *Journal of Combinatorial Optimization*, forthcoming.

[13] Dean, M. and D. Martin (2010), "How Rational are your Choice Data?," mimeo.

[14] Echenique, F, S. Lee, and M. Shum (2010), "The Money Pump as a Measure of Revealed Preference Violations," mimeo.

[15] Famulari, M. (1995), "A Household-Based, Nonparametric Test of Demand Theory," *Review of Economics and Statistics*, 77:372–383.

[16] Fukui, Y. (1986), "A More Powerful Method for Triangularizing Input-Output Matrices and the Similarity of Production Structures," *Econometrica*, 54:1425–1433.

[17] Green, J.R. and D.A. Hojman (2009), "Choice, Rationality and Welfare Measurement," mimeo.

[18] Grötschel, M., M. Jünger, and G. Reinelt (1984), "A Cutting Plane Algorithm for the Linear Ordering Problem." *Operations Research*, 2:1195–1220.

[19] Houtman, M. and J.A. Maks (1985), "Determining All Maximal Data Subsets Consistent with Revealed Preference." *Kwantitatieve Methoden*, 19:89–104.

[20] Howe, E.C. (1991), "A More Powerful Method for Triangularizing Input-Output Matrices: A comment," *Econometrica*, 59, 521–523.

[21] Iyengar, S.S. and M.R. Lepper (2000), "When Choice is Demotivating: Can One Desire Too Much of a Good Thing?," *Journal of Personality and Social Psychology*, 79:995–1006.

[22] Korte, B. and W. Oberhofer (1970), "Triangularizing Input-Output Matrices and the Structure of Production," *European Economic Review*, 1:482–511.

[23] Köszegi, B. and M. Rabin (2007), "Mistakes in Choice-Based Welfare Analysis," *American Economic Review P&P*, 97:482–511.

[24] Manzini, P. and M. Mariotti (2007), "Sequentially Rationalizable Choice," *American Economic Review*, 97:477–481.

[25] Manzini, P. and M. Mariotti (2009), "Choice Based Welfare Economics for Boundedly Rational Agents" mimeo.

[26] Masatlioglu, Y. and D. Nakajima (2008), "Theory of Choice by Elimination," mimeo.

[27] Masatlioglu, Y., D. Nakajima and E.Y. Ozbay (2010), "Revealed Attention," *American Economic Review*, forthcoming.

[28] Masatlioglu, Y. and E.A. Ok (2005), "Rational Choice with Status Quo Bias," *Journal of Economic Theory*, 121:1–29.

[29] May, K.O. (1954), "Intransitivity, Utility, and the Aggregation of Preference Patterns," *Econometrica*, 22:1–13.

[30] Rubinstein, A. and Y. Salant (2011), "Eliciting Welfare Preferences from Behavioral Datasets," *Review of Economic Studies*, forthcoming.

[31] Salant, Y. and A. Rubinstein (2008), "$(A, f)$: Choice with Frames," *Review of Economic Studies*, 75:1287–1296.

[32] Swofford, J.L. and G.A. Whitney (1987), "Nonparametric Test of Utility Maximization and Weak Separability for Consumption, Leisure and Money," *Review of Economic and Statistics*, 69:458–464.

[33] Thaler, R. (1980), "Toward a positive theory of consumer behavior," *Journal of Economic Behavior and Organization*, 1:39–60.

[34] Tversky, A. and D. Kahneman (1981), "The Framing of Decisions and the Psychology of Choice," *Science*, 211:453–458.

[35] Tversky, A. and I. Simonson (1993), "Context-dependent Preferences," *Management Science*, 39:1179-1189.

[36] Varian, H.R. (1990), "Goodness-of-Fit in Optimizing Models," *Journal of Econometrics*, 46:125–140.