A Measure of Rationality and Welfare

Jose Apesteguia
Miguel Angel Ballester

May 2010

A MEASURE OF RATIONALITY AND WELFARE[*]

JOSE APESTEGUIA[†] AND MIGUEL A. BALLESTER[‡]

ABSTRACT. There is ample evidence to show that choice behavior often deviates from the classical principle of maximization. This evidence raises at least two important questions: (i) how severe the deviation is and (ii) which method is the best for extracting relevant information from the choices of the individual for the purposes of welfare analysis. In this paper we address these two questions by proposing a set of foundational conditions on which to build a proper measure of the rationality of individuals, and enable individual welfare analysis of potentially inconsistent subjects, all based on standard revealed preference data. In our first result, we show that there is a unique measure of rationality that satisfies all of the proposed axioms: the weighted-loss indices. In the second part of the paper, we study some relevant properties of weighted-loss indices.

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

1. INTRODUCTION

The maximization principle in the classical theory of choice has two principal features. The first is that it provides a simple and versatile account of individual behavior. In other words, the alternative chosen by the individual is the one that maximizes a well-behaved preference relation over the menu of available alternatives. It is difficult to conceive of a simpler and more operational model. Its second feature is that it suggests the maximized preference relation as a tool for individual welfare analysis. That is, the standard approach involves the policy-maker aiming to reproduce the decisions that, given the opportunity, 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 classical theory of choice. Some phenomena that have attracted a great deal of attention, both empirical and theoretical, and prove difficult, if not impossible, to accommodate within the classical maximization principle are framing effects, menu effects, the importance of reference points, cyclic choice patterns, choice overload effects, etc.[1] The violation in some instances of the principle of maximization 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 determine whether the classical maximization principle is a reasonable way to describe behavior. That is, the question is not whether or not individuals violate the maximization principle in a given situation, but 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 fundamental primitives. Furthermore, the possibility of performing meaningful comparisons of rationality will allow comparison 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).

As for Q.2, if the behavior of the individual is incompatible with the maximization principle, it immediately follows that there is no single preference relation that explains the behavior and thus, it is difficult, from an external perspective, to identify the good or bad alternatives for the individual. In this case, therefore, the possibility of conducting welfare analysis from revealed choices is not immediate.

---

[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 (2009).

In this paper we address these two questions by suggesting a set of foundational conditions on which to build a proper measure of the rationality of individuals, and enable individual welfare analysis using potentially inconsistent subjects. Our approach retains the following features of the classical theory of choice: (i) the maximization principle as the benchmark of reference, and (ii) the standard revealed preference data as the primitives of the model. In other words, behavior is judged on the basis of its closeness to the maximization of a preference relation, and all that has to be observed are the menus of alternatives and the corresponding choices. The presence of potentially inconsistent individuals, however, raises the issue of which preference relation provides the best approximation of the revealed choices, and how close the approximation can be. That is, when the decision-maker is inconsistent, no matter which preference relation is used to represent the decision-maker, there will inevitably be some decisions that cannot be explained as the maximization of that preference relation. It is crucial to notice, however, that the use of one preference relation or another to summarize the choices of the inconsistent decision-maker still matters. Different preference relations explain different choices, and lead to mistakes of a very different nature. In order to properly address this issue, we introduce the notion of an *inconsistency index*. An inconsistency index evaluates how well a preference relation explains observed behavior. The problem that now arises is that there is a wide range of different possible inconsistency indices. Hence, a clear choice is difficult a priori. Therefore, in order to obtain a sound basis for selecting a particular inconsistency index, we propose a number of desirable conditions that an inconsistency index I should satisfy. Then, in our first result, we show that there is a unique class of inconsistency indices satisfying all the proposed axioms. We call them weighted-loss indices.

A weighted-loss index measures the inconsistency between a preference relation and revealed choices 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 weighting the divergence by means of a collection of weights that assigns higher values to the alternatives that are higher in the order determined by the preference relation. The weights may represent, for example, the utility values associated to the different alternatives, according to their position in the ranking. That is, the severity of the inconsistency is measured not only in terms of the number of menus in which the individual deviates from the maximization principle, but also by how far the chosen alternative is from the maximal alternative in each instance.

Taking the weighted-loss index as the inconsistency index of reference, and given the revealed choices of the individual, the minimization of the index over the set of

possible preference relations gives the preference relation that best explains observed choices and the inconsistency level associated to it. The latter provides a measure of the rationality of the individual, while the former gives a preference relation that enables classical welfare analysis.

Given our characterization result, the second part of the paper focuses on some properties of weighted-loss indices for a domain of menus of alternatives satisfying a regularity condition. The observed domains include standard cases, such as the universal domain, which includes all possible menus of alternatives; the binary domain, which consists exclusively of menus presenting only two alternatives; etc. We first show that the preference relation minimizing the weighted-loss index is a very basic one: it simply places the mostly frequently chosen alternatives higher in the ranking. That is, the preference relation that best explains behavior is invariant with respect to the vector of weights and, hence, depends only on choice behavior and is extremely easy to compute. Therefore, the proposed model can easily deal with both issues: measuring the degree of rationality of an individual and conducting welfare analysis. We next address the comparison of degrees of rationality between different individuals or different models of choice. Based on the previous result, we know that the only relevant data for such comparisons is the vector that gives the number of menus from which each alternative is chosen. We then show that the welfare-loss index is a Schur-convex function over such vectors, which enables the use of well-known standard majorization techniques, often used in, for example, the literature on income inequalities. Finally, we show the feasibility of a choice pattern in which all the alternatives are chosen roughly the same number of times, and that provides a tight upper bound for the maximum attainable degree of irrationality.

We conclude this part of the introduction by illustrating our framework and results with two simple examples involving three alternatives: (i) a simplified version of the *attraction effect*, and (ii) a *pairwise cycle*. To this end, denote by (A, a) an observation where A represents the available menu of alternatives and $a \in A$ the alternative chosen from the menu A. Starting with the attraction effect, suppose that alternative a is regarded as better than alternative b and that c is undoubtedly dominated by b but not by a. By the attraction effect, the availability of alternative c induces the individual to choose the dominant alternative b rather than a in the menu $\{a, b, c\}$. Thus, an individual trapped by the attraction effect may exhibit the following behavior in a universal domain: $[(\{a, b, c\}, b), (\{a, b\}, a), (\{a, c\}, a), (\{b, c\}, b)]$.[2] We first obtain the optimal preference relation, that is, the one that minimizes the inconsistency value

---

[2]Singletons can clearly be ignored.

associated to a weighted-loss index. For this, we simply need to count the number of times the different alternatives are chosen. Since both a and b are chosen twice, while c is never chosen, the preference that minimizes the weighted-loss index is either $aPbPc$ or $bP'aP'c$. For any given vector of weights w with $w_1 > w_2 > w_3$, the inconsistency value is, in either case, $w_1 - w_2$. This follows because both preference relations P and P' fail to explain one observation where the maximal alternative is top in the preference relation but the chosen alternative is in the middle. That is, in the case of preference P the observation $(\{a, b, c\}, b)$ cannot be rationalized, since alternative b, which is ranked second, is chosen in the presence of alternative a, which is ranked first. Similarly, with P', observation $(\{a, b\}, a)$ cannot be rationalized, in this case because alternative a, which is ranked second, is chosen in the presence of b, which is ranked first. In both cases, the degree of inconsistency is $w_1 - w_2$.

In a *pairwise cycle*, the individual chooses a over b, b over c and c over a in pairwise comparisons and, let us say, a when all three alternatives are available. Thus, in a universal domain, the following is a pairwise cyclical choice pattern: $[(\{a, b, c\}, a),$ $(\{a, b\}, a), (\{a, c\}, c), (\{b, c\}, b)]$. Notice that a is chosen twice, while b and c are chosen only once. Consequently, the preference that minimizes the weighted-loss index is either $aPbPc$ or $aP'cP'b$. In both cases, the inconsistency value is $w_1 - w_3$. To see the latter, note that the preference P fails to explain observation $(\{a, c\}, c)$, where the maximal alternative is at the top of the preference relation but the chosen alternative is at the bottom, and hence the total inconsistency value according to the weighted-loss index is $w_1 - w_3$. In terms of the preference P' there are two observations, $(\{a, c\}, c)$ and $(\{b, c\}, b)$, where the choices do not match the corresponding maximal alternatives. The additive nature of the weighted-loss index results in a total inconsistency value of $(w_1 - w_2) + (w_2 - w_3) = w_1 - w_3$.

Finally, we can use the above analysis to compare the degrees of inconsistency in the two examples. It immediately emerges that, since for every vector of weights w, $w_1 - w_3 > w_1 - w_2$, the example involving the pairwise cycle has a greater degree of inconsistency than the one featuring the attraction effect.

1.1. **Related Literature.** This 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. The first to propose this type of test was Afriat (1973), that suggested to measure 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 an efficient algorithm to compute the maximal acyclic subset.

The second is the growing number of papers proposing models enabling 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 former is chosen over the latter in every single ancillary condition. It is clear that this method will typically provide an incomplete preference relation, and hence will be often uninformative. Chambers and Hayashi (2009) characterizes 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 and suggest some generalizations. Rubinstein and Salant (2009) also adopt the view that a decision-maker may be affected by frames, but deviate from the above models in several ways. Rubinstein and Salant require information regarding the cognitive process that distorts the unobservable welfare relation. Then, a welfare relation is consistent with a data set made up of framed preference relations, if any preference relation in the data set could have been generated by the cognitive process. Green and Hojman (2009) adopt a multiple-selves model. Given a set of standard revealed choices, Green and Hojman 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.

2. Notation

Let X be a finite set of k alternatives. An observation consists of a non-empty menu of alternatives $A \subseteq X$ and an element a that is chosen from A. $\mathcal{A} = [(A_i, a_i)_{i=1}^n]$ denotes a collection of n observations. We allow for the possibility that the same

menu A may be in $\mathcal{A}$ more than once, with the same or different associated chosen elements. That is, a collection of observations is simply a set of observations that allows for repetition of observations. The conjunction of two collections of observations $\mathcal{A} = [(A_i, a_i)_{i=1}^n]$ and $\mathcal{B} = [(B_j, b_j)_{j=1}^m]$ is the collection of observations $\mathcal{A} \oplus \mathcal{B} = [(A_1, a_1), \ldots, (A_n, a_n), (B_1, b_1), \ldots, (B_m, b_m)]$. Let $d > 0$ be an integer, then $d\mathcal{A}$ denotes the d-fold conjunction of the collection of observations $\mathcal{A}$.

A preference relation $P \subseteq X \times X$ is a strict linear order over X, that is, an asymmetric, transitive, and connected binary relation over X. Given a preference relation P and a menu of alternatives A, the maximal element in A under P is denoted by $m(P, A)$. The preference relation P determines a ranking of the alternatives in X from 1, which denotes the most preferred alternative, to k, which denotes the least preferred alternative. Given a preference relation P and an observation (A, a), we denote by $\widehat{m}(P, A)$ the position of the maximal element $m(P, A)$ in the preference relation P, and by $\widehat{a}(P)$ the position of the chosen element a in the preference relation P. We say that the preference relation P *rationalizes* the collection of observations $\mathcal{A} = [(A_i, a_i)_{i=1}^n]$ whenever the chosen element and the maximal element coincide in every single observation, that is, whenever $m(P, A_i) = a_i$ for all $1 \le i \le n$. For any two menus of alternatives, A and B, we write APB whenever every alternative in A is preferred, under P, to any alternative in B.

Given a preference relation P and a collection of observations $\mathcal{A}$, denote by $\lambda_{ts}(P, \mathcal{A})$, $1 \le t \le s \le k$, the number of observations (A, a) that are in $\mathcal{A}$, such that t is the rank of the maximal element and s is the rank of the chosen element, that is, $\widehat{m}(P, A) = t$ and $\widehat{a}(P) = s$. The vector of all the possible combinations is $\lambda(P, \mathcal{A}) = (\lambda_{11}(P, \mathcal{A}), \ldots, \lambda_{1k}(P, \mathcal{A}), \lambda_{22}(P, \mathcal{A}), \ldots, \lambda_{2k}(P, \mathcal{A}), \ldots, \lambda_{kk}(P, \mathcal{A}))$. Denote by $\delta_t(P, \mathcal{A})$ the number of observations (A, a) in $\mathcal{A}$ where t is the rank of the maximal element, that is, $\delta_t(P, \mathcal{A}) = \sum_{s=t}^k \lambda_{ts}(P, \mathcal{A})$. The vector of all possible such entries is denoted by $\delta(P, \mathcal{A}) = (\delta_1(P, \mathcal{A}), \ldots, \delta_k(P, \mathcal{A}))$. Finally, denote by $\rho_s(P, \mathcal{A})$ the number of observations (A, a) in $\mathcal{A}$ where s is the rank of the chosen alternative, that is, $\rho_s(P, \mathcal{A}) = \sum_{t=1}^s \lambda_{ts}(P, \mathcal{A})$, and denote the associated vector by $\rho(P, \mathcal{A}) = (\rho_1(P, \mathcal{A}), \ldots, \rho_k(P, \mathcal{A}))$. In sum, λ refers to the (mis)matches between the chosen and the maximal alternatives, δ to the maximal alternatives, and ρ to the chosen alternatives.

Let $\sigma : X \to X$ denote a permutation. Given a collection of observations $\mathcal{A}$, the permuted collection of observations is denoted by $\sigma(\mathcal{A}) = [(\sigma(A_i), \sigma(a_i))_{i=1}^n]$. Similarly, given a preference relation P, $\sigma(P)$ denotes the permuted preference relation, that is, $xPy \Leftrightarrow \sigma(x)\sigma(P)\sigma(y)$.

3. The Model

We now introduce a key concept into our analysis, an inconsistency index I. The mapping I evaluates the inconsistency incurred by a preference relation P in explaining a collection of observations $\mathcal{A}$.

Definition 1. An **inconsistency index** I assigns to each preference relation P and each collection of observations $\mathcal{A}$ a non-negative real number $I(P, \mathcal{A})$.

Given an inconsistency index I and a collection of observations $\mathcal{A}$, different preference relations may incur in different inconsistencies. Hence, we are interested in: (1) the minimum value of I across all preference relations, which we denote by $I^*(\mathcal{A})$; and (2) the preference relation that provides that minimal inconsistency level $I^*(\mathcal{A})$. The former measures the irrationality of an individual, while the latter provides a tool for the welfare analysis of possibly inconsistent individuals.

3.1. **Properties on Inconsistency Indices.** In principle, there is a wide range of possible different inconsistency indices and hence, a clear choice is problematic a priori. Therefore, to create a sound basis for the selection of a particular inconsistency index, we now suggest a number of conditions that an inconsistency index I should desirably satisfy. Then, in section 3.3, we show that, in fact, the five properties we now suggest uniquely identify a particular family of inconsistency indices.

Neutrality (NEU). For any permutation σ, for any preference relation P, and for any collection of observations $\mathcal{A}$, $I(P, \mathcal{A}) = I(\sigma(P), \sigma(\mathcal{A}))$.

Neutrality imposes that the inconsistency index should be independent of the names of the alternatives. That is, any relabeling should have no effect on the level of inconsistency. This is a standard axiom in various areas of research, such as the foundations of voting rules, for example.

Rationality (RAT). For any preference relation P, and for any collection of observations $\mathcal{A}$, $I(P, \mathcal{A}) = 0$ if and only if P rationalizes $\mathcal{A}$.

Since an inconsistency index assigns non-negative numbers, this condition 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 the preference relation, and any other case gives a strictly higher inconsistency level. The property also normalizes the minimal

inconsistency level by assigning a value of 0.

Separability (SEP). For any preference relation P and for any three collections of observations $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$, $I(P, \mathcal{A}) \geq I(P, \mathcal{B})$ if and only if $I(P, \mathcal{A} \oplus \mathcal{C}) \geq I(P, \mathcal{B} \oplus \mathcal{C})$.

Recall that a collection of observations is a set that allows for repetition of observations. Then if, under the preference relation P, the inconsistency index judges $\mathcal{A}$ to be more inconsistent than $\mathcal{B}$, then separability imposes that the addition of the same collection of observations $\mathcal{C}$ to both cases, $\mathcal{A}$ and $\mathcal{B}$, should not reverse the order of inconsistency. That is, two collections of observations maintain the same order of inconsistency when the same extra collection of observations is added to both sides.

Composition (COM). For any preference relation P, and any two menus A and B, such that APB, $I(P, [(A, x), (B \cup \{x\}, y)]) = I(P, [(A \cup B, y)])$.

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, given a preference P, and an observation (A, x), it may be the case that x is different from the maximal element in A, which would imply an inconsistency. Now suppose that we observe the menu $\{x\} \cup B$, where all of the alternatives in the set B are worse than x under P. Clearly, it follows that x is the maximal alternative in the menu $\{x\} \cup B$. It may be the case that the chosen alternative in this menu, y, is different from x, again implying a inconsistency in the preference relation P. 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.

Archimedean property (ARCH). For any preference relation P and for any four collections of observations $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$, if $I(P, \mathcal{A}) > I(P, \mathcal{B})$ then there is an integer $d > 0$ such that $I(P, d\mathcal{A} \oplus \mathcal{C}) \geq I(P, d\mathcal{B} \oplus \mathcal{D})$.

The Archimedean property imposes that there is no level of inconsistency that cannot be reversed, if sufficient data are added in the opposite direction. That is, if given the preference relation P, the inconsistency index judges the collection of observations $\mathcal{D}$ to be more inconsistent than the collection $\mathcal{C}$, the judgment can be reversed by adding a collection of observations $\mathcal{A}$ to $\mathcal{C}$ a sufficient number of times, and a collection of observations $\mathcal{B}$ to $\mathcal{D}$ the same number of times, such that the inconsistency

index judges $\mathcal{A}$ to be more inconsistent than $\mathcal{B}$.

3.2. **Weighted-loss Indices.** We now signify a particular class of inconsistency indices, which we call weighted-loss indices, that will play a key role in this paper. When judging the inconsistency of a collection of observations in terms of a preference relation P, a weighted-loss index additively considers every single observation in which there is a divergence between the maximal element and the chosen element. It then weights each divergence using a set of weightings that assigns higher values to the alternatives that come higher in the order determined by the preference relation. The weights may represent, for example, the utility values associated to the different alternatives, according to their position in the ranking. More precisely:

Definition 2. Denote by $\mathcal{W}$ the class of vectors of weights $w = (w_1, \ldots, w_k)$ such that $w_1 > w_2 > \cdots > w_k$. We say that an inconsistency index is a **weighted-loss index** associated to $w \in \mathcal{W}$ and denoted by I_w, if, for any preference relation P and for any collection of observations $\mathcal{A} = [(A_i, a_i)_{i=1}^n]$, $I_w(P, \mathcal{A}) = \sum_{i=1}^n (w_{\widehat{m}(P, A_i)} - w_{\widehat{a}_i(P)})$. Denote the class of weighted-loss indices by $\mathcal{I}_{\mathcal{W}}$.

For those observations (A_i, a_i) for which the chosen element is the maximal element in A_i, that is $m(P, A_i) = a_i$, the weighted-loss index I_w assigns an inconsistency value of 0. Any other observation receives a positive inconsistency value, which varies according to how much higher or lower the preference relation ranks the chosen element in comparison to the maximal element. The further apart these two elements are in the preference relation, the greater the inconsistency in the explanation of the observation by the preference relation. Then, the inconsistency incurred by the preference relation over the whole set of observations is the sum of the inconsistency values obtained in the individual observations. Notice, therefore, that in terms of the weighted-loss index, it is not only the number of inconsistencies that matters, but also their significance in the preference relation P.

The following immediate lemma shows two alternative equivalent formulations of weighted-loss indices.

Lemma 1. *For any $w \in \mathcal{W}$, any preference relation P, and any collection of observations $\mathcal{A} = [(A_i, a_i)_{i=1}^n]$, $I_w(P, \mathcal{A}) = \sum_{t=1}^k \sum_{s=t}^k \lambda_{ts}(P, \mathcal{A})(w_t - w_s) = \sum_{t=1}^k \delta_t(P, \mathcal{A}) w_t - \sum_{s=1}^k \rho_s(P, \mathcal{A}) w_s$.*

To grasp the first formulation in Lemma 1, $I_w(P, \mathcal{A}) = \sum_{t=1}^k \sum_{s=t}^k \lambda_{ts}(P, \mathcal{A})(w_t - w_s)$, note that, in a weighted-loss index, those observations that have both the same maximal element and the same chosen element can be regarded as indistinguishable. This shows

that in a weighted-loss index, the inconsistency of a collection of observations may be fully captured by the vector λ.

With regard to the second formulation in Lemma 1, $I_w(P, \mathcal{A}) \sum_{t=1}^{k} \delta_t(P, \mathcal{A}) w_t - \sum_{s=1}^{k} \rho_s(P, \mathcal{A}) w_s$, simply note that $\sum_{t=1}^{k} \sum_{s=t}^{k} \lambda_{ts}(P, \mathcal{A}) w_t = \sum_{t=1}^{k} \delta_t(P, \mathcal{A}) w_t$ and that $\sum_{t=1}^{k} \sum_{s=t}^{k} \lambda_{ts}(P, \mathcal{A}) w_s = \sum_{s=1}^{k} \sum_{t=1}^{s} \lambda_{ts}(P, \mathcal{A}) w_s = \rho_s(P, \mathcal{A}) w_s$. That is, a weighted-loss index can also be understood as the difference between two components, one depending only on the maximal alternatives and the other depending only on the chosen alternatives.

For the sake of illustration, in the remainder of this subsection, we compare the degrees of inconsistency assigned by weighted-loss indices to collections of observations that violate one of the following two classical properties in the choice theory literature: No Binary Cycles and Always Chosen.

No Binary Cycles: For all $a_1, \ldots, a_r \in X$ and any collection of observations $\mathcal{A}$, if $(\{a_j, a_{j+1}\}, a_j) \in \mathcal{A}$, $j = 1, \ldots, r - 1$, then $(\{a_1, a_r\}, a_r) \notin \mathcal{A}$.

Always Chosen: For all $a_1, \ldots, a_r \in X$ and any collection of observations $\mathcal{A}$, if $(\{a_1, a_j\}, a_1) \in \mathcal{A}$, $j = 2, \ldots, r$, then $(\{a_1, \ldots, a_r\}, a_j) \notin \mathcal{A}$ with $j \neq 1$.

The interpretation of the two properties is immediate. No Binary Cycles imposes consistent behavior in binary problems by forbidding a pairwise cycle of choice. According to Always Chosen, if an alternative a is chosen in pairwise choices involving it and all the other alternatives in a set, the only possible choice from the set is alternative a itself. Manzini and Mariotti (2007) show that, in a universal choice domain, choice behavior that cannot be rationalized by the maximization of a single preference relation violates either No Binary Cycles, or Always Chosen, or both. That is, irrational behavior is fully characterized by the violation of either or both of these properties.

Now, we use weighted-loss indices to evaluate the severity of the departure from rationality of a violation of No Binary Cycles versus a violation of Always Chosen. To this end, consider two collections of observations $\mathcal{A}$ and $\mathcal{B}$, where $\mathcal{A}$ is composed of a pairwise cycle involving all the alternatives in X, while $\mathcal{B}$ is composed of a violation of Always Chosen in the same menu of alternatives X. That is, $r = k$ in both cases.[3]

Proposition 1 shows that, from the point of view of weighted-loss indices, a decision-maker violating No Binary Cycles incurs in a higher degree of irrationality than one violating Always Chosen. Therefore, although inconsistent behavior necessarily implies

---

[3]The arguments that follow can be immediately extended to any violation of No Binary Cycles or Always Chosen involving any subset of alternatives from X.

a violation of Always Chosen or of No Binary Cycles, weighted-loss indices unambiguously discriminate between violating one property or another.

Proposition 1. *For any vector of weights $w \in \mathcal{W}$,*

- *$I_w^*(\mathcal{A}) = w_1 - w_k$, and*
- *$I_w^*(\mathcal{B}) = w_1 - w_2$*

The intuition for the first claim of the proposition is a generalization of the argument used in the pairwise cycle example in the introduction.[4] The preference relation $a_1 P a_2 P \cdots a_{k-1} P a_k$, for example, gives the minimum inconsistency value associated with $\mathcal{A}$, $w_1 - w_k$. In the case of $\mathcal{B}$, the collection of observations involving a violation of Always Chosen, it is easy to see that any preference relation where the top two alternatives are a_1 and the alternative chosen from menu X, provides the minimal inconsistency value, namely, $I_w^*(\mathcal{B}) = w_1 - w_2$.

3.3. **The Characterization Result.** Our first result shows that an inconsistency index satisfies the five properties proposed in subsection 3.1 above if and only if the inconsistency index is a weighted-loss index.

Theorem 1. *An inconsistency index I satisfies (NEU), (RAT), (SEP), (COM) and (ARCH) if and only if it is a weighted-loss index $I \in \mathcal{I}_\mathcal{W}$.*

The main steps of the proof of Theorem 1 are as follows. We focus on the 'only-if' part. First, we show that for any preference relation P, the degree of inconsistency of any collection of observations $\mathcal{A}$, is equivalent to the degree of inconsistency involved in two particular collections of observations $\mathcal{B}$ and $\mathcal{R}$, that is, $I(P, \mathcal{A}) = I(P, \mathcal{B} \oplus \mathcal{R})$. Collection $\mathcal{B}$ consists of binary menus of alternatives where the maximal alternative is not chosen, while collection $\mathcal{R}$ consists of menus of alternatives where the maximal alternative is in fact chosen.

The next step uses the previous one to show that an inconsistency function is fully characterized by the vector $\lambda(P, \mathcal{A})$. Consequently, we can define the binary relation $\succeq$ over $\mathbb{N}^{k(k+1)/2}$ in order to compare the associated λ-vectors of two preference relations P and P', and two collections of observations $\mathcal{A}$ and $\mathcal{A}'$. We then show that the triple $(\mathbb{N}^{k(k+1)/2}, \succeq, +)$ is a closed extensive structure, hence we can use known results in the foundations of measurement to conclude that there exists a real-valued function ϕ on $\mathbb{N}^{k(k+1)/2}$ representing $\succeq$ such that $\lambda(P, \mathcal{A}) \succeq \lambda(P', \mathcal{A}') \Leftrightarrow \phi(\lambda(P, \mathcal{A})) \geq \phi(\lambda(P', \mathcal{A}'))$ and furthermore, ϕ is additive, that is $\phi(\lambda(P, \mathcal{A}) + \lambda(P', \mathcal{A}')) = \phi(\lambda(P, \mathcal{A})) + \phi(\lambda(P', \mathcal{A}'))$.

It follows immediately that $I = \phi$ up to a scalar transformation, and $I(P, \mathcal{A}) = \sum_{t=1}^{k} \sum_{s=t}^{k} \theta_{ts} \lambda_{ts}(P, \mathcal{A})$, where θ_{ts} is the canonical value assumed by ϕ over the vector

---

[4]The formal proof of Proposition 1, and all the results to come, can be found in Appendix A.

with value 1 for the component ts and value 0 for all the other components. That is, θ_{ts} is the inconsistency value associated with an observation where the maximal element is ranked at t, but the chosen element is ranked at s.

Finally, we show that $\theta_{tt} = 0$ for every t, $\theta_{ts} > 0$ whenever $t < s$, and that $\theta_{th} = \theta_{ts} + \theta_{sh}$, and therefore, we define $w_t = \sum_{s=t}^{k-1} \theta_{s,s+1}$ and $w_k = 0$, and hence we can write the inconsistency index as $I(P, \mathcal{A}) = \sum_{t=1}^{k} \sum_{s=t}^{k} \lambda_{ts}(P, \mathcal{A})(w_t - w_s)$, as desired.

4. Discussion

Theorem 1 identifies weighted-loss indices as especially prominent inconsistency indices since these are the only ones that satisfy the set of properties here proposed. Therefore, we now take a more detailed look at various important aspects of the class of weighted-loss indices. We first characterize the preference relation minimizing weighted-loss indices. We argue that such a preference relation may be of importance in conducting welfare analysis with inconsistent individuals. We then provide a sound tool for comparing the degree of irrationality of two individuals. Finally, we close this section by characterizing the maximum degree of inconsistency that can actually be attained. We study these issues on a particular domain of collections of observations satisfying a regularity condition.

Definition 3. The collection of observations $\mathcal{A}$ is a **balanced collection of observations** if all the menus of alternatives in $\mathcal{A}$ with the same cardinality are observed the same number of times.

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.

4.1. **Individual Welfare.** If an individual is inconsistent it immediately implies that there is no single preference relation explaining her choices, and hence welfare analysis becomes problematic. It is difficult in such a case to elucidate what the individual would choose for herself, and hence it is problematic for the policy to choose on the individual's behalf. Our approach, in this respect, is to focus on the preference relation that minimizes the inconsistency with the data as measured by the class of weighted-loss indices, and take this preference relation as the best approximation of the individual's actual choices. The 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.

Intuitively, the more often an alternative is chosen, the higher it should be placed in the preference relation. This is certainly true in the rational case, where the maximization of a preference relation accounts for all choices. The following definition formally states the preference relations that place the more frequently chosen alternatives higher in the ranking.

Definition 4. Given a collection of observations $\mathcal{A} = [(A_i, a_i)_{i=1}^n]$, a **basic preference relation** $P^B(\mathcal{A})$ is any preference relation such that $|\{i : a_i = x\}| > |\{i : a_i = y\}| \Rightarrow xP^B(\mathcal{A})y$.

It turns out that, when the collection of observations is balanced, our intuition that the weighted-loss index is minimized with basic preference relations is correct. Since basic preference relations do not depend on the vector of weights w, this implies that the preference relations that minimize the weighted-loss index are invariant with respect to w.

Theorem 2. *For any vector of weights $w \in \mathcal{W}$ and any balanced collection of observations $\mathcal{A}$, any basic preference relation $P^B(\mathcal{A})$ minimizes the weighted-loss index I_w.*

The main steps in the proof of Theorem 2 are as follows. We first show that the weighted sum of the number of observations in $\mathcal{A}$ with maximal elements in the various positions t in the ranking $\sum_{t=1}^k \delta_t(P, \mathcal{A})w_t$ does not depend on the preference relation P. That is, for a given w, for any balanced collection of observations $\mathcal{A}$, and for any two preference relations P and P', $\sum_{t=1}^k \delta_t(P, \mathcal{A})w_t = \sum_{t=1}^k \delta_t(P', \mathcal{A})w_t = K(\mathcal{A}, w)$, where $K(\mathcal{A}, w)$ is a function that depends only on $\mathcal{A}$ and the vector of weights w. Hence, the vector δ referring to the maximal alternatives can be ignored.

Next, we use the above result together with Lemma 1 to obtain a useful expression for the comparison of the irrationality incurred by two preference relations P and P', explaining a given collection of observations $\mathcal{A}$, that depends only on the vector ρ measuring the number of times each alternative is chosen. Namely $I_w(P, \mathcal{A}) \leq I_w(P', \mathcal{A}) \Leftrightarrow \sum_{s=1}^{k-1} \left(\sum_{l=1}^s \rho_l(P, \mathcal{A}) \right)\theta_{s,s+1} \geq \sum_{s=1}^{k-1} \left(\sum_{l=1}^s \rho_l(P', \mathcal{A}) \right)\theta_{s,s+1}$.

Finally, the definition of a basic preference relation $P^B(\mathcal{A})$ guarantees precisely that $\sum_{l=1}^s \rho_s(P, \mathcal{A})$ is maximized at $P^B(\mathcal{A})$. Then, given the positivity of θ, $P^B(\mathcal{A})$ maximizes $\sum_{s=1}^{k-1} \left(\sum_{l=1}^s \rho_l(P, \mathcal{A}) \right)\theta_{s,s+1}$ and therefore, by the expression above, $I(P^B(\mathcal{A})) \leq I(P, \mathcal{A})$, as desired.

4.2. **Comparing the Level of Rationality.** We now address the question of comparing two collections of observations in terms of their rationality levels. This is an important question, both for the analysis of different theoretical models of bounded rationality, and for establishing meaningful interpersonal comparisons of rationality in empirical applications. The former provides a natural tool to give structure to the rapidly growing literature on alternative models of choice, while the latter may serve as the basis for a better understanding of individual behavior in certain environments, and may also provide the primitives for the development of potential new choice theories relying on the distribution of individuals by rationality levels.

Consider two balanced collections of observations with the same menus of alternatives. By Theorem 2, we know that for each of the balanced collections, the preference relation that minimizes the weighted-loss index is a basic preference relation, which is fully determined by the vector ρ reporting the number of times each alternative is chosen. Corollary 1 below, complements this result by showing that the decision as to which behavior is more rational is also exclusively determined by the vector ρ of chosen alternatives. Hence, again, the comparison between different collections of observations is invariant with respect to the vector of weights w.

Furthermore, Corollary 1 identifies a simple test for comparing collections of observations. It shows that, whenever the vector ρ associated to one collection of observations majorizes the vector ρ associated to another collection of observations, we can conclude that the former is more rational than the latter. That is, weighted-loss indices are Schur-convex functions for collections of observations with the same menus of alternatives. This leads to the use of techniques developed in the majorization literature becoming standard in the income inequality literature.

Corollary 1. *For any vector of weights $w \in \mathcal{W}$ and any two balanced collections of observations $\mathcal{A}$ and $\mathcal{A}'$ with the same menus of alternatives, $I_w^*(\mathcal{A}) \leq I_w^*(\mathcal{A}')$ if and only if $\sum_{s=1}^k \rho_s(P^B(\mathcal{A}), \mathcal{A})w_s \geq \sum_{s=1}^k \rho_s(P^B(\mathcal{A}'), \mathcal{A}')w_s$. In particular, if $\rho(P^B(\mathcal{A}), \mathcal{A})$ majorizes $\rho(P^B(\mathcal{A}'), \mathcal{A}')$ then $I_w^*(\mathcal{A}) \leq I_w^*(\mathcal{A}')$. That is, I_w^* is a Schur-convex function for collections of observations with the same menus.*

The proof of Corollary 1 follows from Theorem 2 by imposing the condition that the two collections of observations $\mathcal{A}$ and $\mathcal{A}'$ must have the same menus of alternatives.

4.3. **Maximum Degree of Inconsistency.** We now address the question of what maximal degree of irrationality can actually be reached. Corollary 1 above shows that if it were possible to construct a balanced collection of observations in which all the alternatives were chosen roughly the same number of times, it would provide the

highest level of inconsistency. Theorem 3 below shows that the construction of such a collection is feasible.

First, consider the following notation. Recall that k denotes the number of alternatives in the grand set X. Given a collection of n observations $\mathcal{A}$, denote by $a(n,k)$ and $b(n,k)$ the unique positive integers such that (1) $n = a(n,k)k + b(n,k)$, and (2) $b(n,k) < k$. Theorem 3 shows that it is possible to construct a collection of n observations in which each of the k alternatives is chosen roughly $a(n,k) \approx n/k$ times.

Theorem 3. *For any vector of weights $w \in \mathcal{W}$ and any balanced collection of n observations $\mathcal{A}$, $I_w^*(\mathcal{A}) \leq K(\mathcal{A}, w) - a(n,k) \sum_{s=1}^{k} w_s - \sum_{s=1}^{b(n,k)} w_s$, and the bound is tight.*

The structure of the proof is as follows. Given a balanced collection of n observations $\mathcal{A}$, we construct another balanced collection of n observations $\mathcal{B}$ with the same menus as $\mathcal{A}$, and such that $I_w^*(\mathcal{B}) = K(\mathcal{A}, w) - a(n,k) \sum_{s=1}^{k} w_s - \sum_{s=1}^{b(n,k)} w_s$. Thus constructed, $\mathcal{B}$ satisfies that $\rho(P^B(\mathcal{A}), \mathcal{A})$ majorizes $\rho(P^B(\mathcal{B}), \mathcal{B})$, and hence, the application of Corollary 1 concludes the proof.

5. Final Remarks

It is now widely accepted that individuals often deviate from the maximization principle when making choices. It is therefore crucial to assign a proper meaning to such deviations from the maximization principle. A deep understanding of such deviations 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 propose a particular measure of the rationality of individuals. Our approach is axiomatic, and keeps as close as possible to the classical framework. Thus, we need only observe standard revealed preference data, and we retain the maximization principle as the benchmark for rationality.

Our characterization exercise obtains that the family of weighted-loss indices is the unique inconsistency functions that satisfies a set of properties. Weighted-loss indices have several attractive characteristics. A weighted-loss index weights each observation in which the individual's revealed choice does not correspond to the maximization of a given preference relation by how far the preference relation ranks the chosen alternative from the maximal alternative. The greater the distance between these two alternatives in the order of preference, the greater the inconsistency incurred in explaining the observation by the preference relation. Then, the sum of the inconsistency values obtained in all such observations gives the total inconsistency level. We further show that weighted-loss indices are very manageable in certain general domains. In particular,

we show that identifying the optimal preference relation that minimizes the level of inconsistency involved in a set of observations is as simple as computing a basic preference relation, that is, one that places the more frequently chosen alternatives higher in the ranking. We argue that such a preference relation may serve as a sound tool for welfare economics, even with inconsistent decision-makers. Interpersonal comparison is also very simple, since weighted-loss indices respect the majorization relation, and hence are Schur-convex functions.

Appendix A. Proofs

Proof of Proposition 1: For the first claim, we first show that for any preference P, the inconsistency is equal to or greater than $w_1 - w_k$. Consider any preference P. Let a_p be the alternative ranked in position k in P. Clearly, observation $(\{a_p, a_{p+1}\}, a_p) \in \mathcal{A}$ (with the convention $a_{k+t} = a_t$) is not rationalized by P. Let l_1 be the position of alternative a_{p+1} in P. The inconsistency of this observation is $w_{l_1} - w_k$. If $l_1 = 1$, the inconsistency index is equal to or greater than $w_1 - w_k$, and we are done. If $l_1 > 1$, find the first alternative a_{p+r}, with $r \geq 1$, such that a_{p+r+1} is ranked higher than l_1. Since $l_1 > 1$, this alternative must exist. Denote the position of a_{p+r+1} in P by l_2. Clearly, observation $(\{a_{p+r}, a_{p+r+1}\}, a_{p+r}) \in \mathcal{A}$ is not rationalized by P. The inconsistency of this observation is greater than or equal to $w_{l_2} - w_{l_1}$. Repeat this process until the alternative ranked in the first position in P is reached. The inconsistency index of the collection of observations is thus greater than or equal to $w_1 - w_k$. Now, consider the preference $a_1 P a_2 P \dots P a_k$. This preference explains all observations except $(\{a_k, a_1\}, a_k) \in \mathcal{A}$, giving an inconsistency value of $w_1 - w_k$.

For the second claim, if a_1 is placed in position $l \neq 1$ in P, the inconsistency value associated with P is at least $w_1 - w_l$, since a_1 is chosen in the pairwise comparison with the alternative ranked in the first position in P. This is equal to or greater than $w_1 - w_2$. Now, place a_1 at the top of the ranking. The placing of a_1 at the top of the ranking rationalizes all the observations except the choice from menu X. If the alternative chosen in X is placed in position l, there is an inconsistency value of $w_1 - w_l$ that is equal to or larger than $w_1 - w_2$. Now, place the alternative chosen in X in second place and the result follows.∎

Proof of Theorem 1: It is easy to see that any weighted-loss index I_w satisfies the axioms. We prove the converse statement by way of four lemmata.

Lemma 2. *For any preference relation P and for any collection of observations $\mathcal{A}$, there exist two collections of observations $\mathcal{B}$ (not necessarily non-empty) and $\mathcal{R}$ (non-empty) such that $I(P, \mathcal{A}) = I(P, \mathcal{B} \oplus \mathcal{R})$ where for all $(B, b) \in \mathcal{B}$, $|B| = 2$ and $b \neq m(P, B)$ and for all $(R, r) \in \mathcal{R}$, $r = m(P, R)$.*

Proof of Lemma 2: Consider a preference relation P and a collection of observations $\mathcal{A}$. Let $\mathcal{A} = \mathcal{A}^1 \oplus \mathcal{A}^2 \oplus \mathcal{A}^3$, where for all $(A, a) \in \mathcal{A}^1$, $|A| = 2$ and $a \neq m(P, A)$; for all $(A, a) \in \mathcal{A}^2$, $a = m(P, A)$; and for all $(A, a) \in \mathcal{A}^3$, $|A| > 2$ and $a \neq m(P, A)$. That is, observations in $\mathcal{A}^1$ are binary menus of alternatives where the maximal alternative is not chosen. Observations in $\mathcal{A}^2$ are menus of alternatives where the maximal alternative is chosen. Observations in $\mathcal{A}^3$ are menus with at least 3 alternatives where the maximal alternative is not chosen.

First, we show that for any collection of observations $\mathcal{A}$ where $\mathcal{A}^3$ is empty, we can construct collections of observations $\mathcal{B}$ and $\mathcal{R}$, as defined in the statement of the lemma, such that $I(P, \mathcal{A}) = I(P, \mathcal{B} \oplus \mathcal{R})$. Suppose, then, that $\mathcal{A}^3$ is empty. If $\mathcal{A}^2$ is non-empty, simply take $\mathcal{B} = \mathcal{A}^1$ and $\mathcal{R} = \mathcal{A}^2$, and we are done. If $\mathcal{A}^2$ is empty, then $\mathcal{A}^1$ is non-empty. In this case, consider an observation $(A, a) \in \mathcal{A}^1$. Obviously, $\{m(P, A)\}$ and $\{a\}$ constitute a partition of A such that $m(P, A) P a$. By (COM), $I(P, [(A, a)]) = I(P, [(\{m(P, A)\}, m(P, A)), (A, a)])$. If $\mathcal{A}^1$ contains more observations than (A, a), by (SEP) we know that $I(P, \mathcal{A}^1) = I(P, \mathcal{A}^1 \oplus [(\{m(P, A)\}, m(P, A))])$. Then, take $\mathcal{B} = \mathcal{A}^1$ and $\mathcal{R} = [(\{m(P, A)\}, m(P, A))]$, and we are done.

Second, suppose that $\mathcal{A}^3$ is non-empty. Consider an observation $(A, a) \in \mathcal{A}^3$. If $|\{x \in A : x P a\}| > 1$, then partition A into the sets $B = \{x \in A : x P a\}$ and $A \setminus B$. By (COM), $I(P, [(A, a)]) = I(P, [(B, m(P, A)), (\{m(P, A)\} \cup A \setminus B, a)])$. We have decomposed the observation (A, a) into two observations such that the first is of the class $\mathcal{A}^2$ type and the second is either of the class $\mathcal{A}^1$ type or the class $\mathcal{A}^3$ type. Note that in the latter case, the menu of alternatives $\{m(P, A)\} \cup A \setminus B$ has strictly fewer alternatives than the original menu, A.

If $|\{x \in A : x P a\}| = 1$, then it must be that $\{x \in A : a P x\} \neq \emptyset$. Partition the menu of alternatives A into the sets $\{m(P, A), a\}$ and $\{x \in A : a P x\}$. By (COM), $I(P, [(A, a)]) = I(P, [(\{m(P, A), a\}, a), (\{a\} \cup \{x \in A : a P x\}, a)])$. Note that we have decomposed (A, a) into two observations such that the first is of the class $\mathcal{A}^1$ type and the second is of the class $\mathcal{A}^2$ type.

Given that X is finite, one can iterate the above argument until the class $\mathcal{A}^3$ is exhausted. Using (SEP), the new collection of observations has the same inconsistency

index I as the original one. Thus, for any collection of observations where $\mathcal{A}^3$ is non-empty, we can also construct collections of observations $\mathcal{B}$ (not necessarily non-empty) and $\mathcal{R}$ (non-empty) as desired, which concludes the proof of the lemma.$\square$

Lemma 3. *For any two preference relations P and P' and for any two collections of observations $\mathcal{A}$ and $\mathcal{A}'$, $\lambda(P, \mathcal{A}) = \lambda(P', \mathcal{A}') \Rightarrow I(P, \mathcal{A}) = I(P', \mathcal{A}')$.*

Proof of Lemma 3: Given Lemma 2, we can assume that $\mathcal{A} = \mathcal{B} \oplus \mathcal{R}$ and $\mathcal{A}' = \mathcal{B}' \oplus \mathcal{R}'$ where the collections of observations $\mathcal{B}$ and $\mathcal{B}'$ consist of binary menus of alternatives in which the maximal alternative is not chosen, the non-empty collections of observations $\mathcal{R}$ and $\mathcal{R}'$ consist of menus of alternatives in which the maximal alternative is chosen, and $I(P, \mathcal{A}) = I(P, \mathcal{B} \oplus \mathcal{R})$ and $I(P', \mathcal{A}') = I(P', \mathcal{B}' \oplus \mathcal{R}')$.

Note that the construction of $\mathcal{B}$ and $\mathcal{R}$ in the proof of Lemma 2 guarantees that $\lambda_{ts}(P, \mathcal{A}) = \lambda_{ts}(P, \mathcal{B} \oplus \mathcal{R})$ for all $t < s$ (note that the equality does not necessary hold for $t = s$). Hence, $\lambda(P, \mathcal{A}) = \lambda(P', \mathcal{A}')$ implies that $\mathcal{B}$ is empty if and only if $\mathcal{B}'$ is empty. If $\mathcal{B}$ and $\mathcal{B}'$ are empty, then $I(P, \mathcal{R}) = I(P', \mathcal{R}') = 0$, and the claim follows. If, on the contrary, $\mathcal{B}$ and $\mathcal{B}'$ are both non-empty, it follows that $\lambda(P, \mathcal{B}) = \lambda(P', \mathcal{B}')$.

Consider a permutation σ such that $\sigma(P) = P'$. The application of (NEU) guarantees that $I(P, \mathcal{B} \oplus \mathcal{R}) = I(\sigma(P), \sigma(\mathcal{B} \oplus \mathcal{R})) = I(P', \sigma(\mathcal{B}) \oplus \sigma(\mathcal{R}))$. Given that $\lambda(P, \mathcal{B}) = \lambda(P', \mathcal{B}')$, we can consider, without loss of generality, that $\sigma(\mathcal{B}) = \mathcal{B}'$ and then $I(P, \mathcal{A}) = I(P', \mathcal{B}' \oplus \sigma(\mathcal{R}))$. Since all the observations in $\sigma(\mathcal{R})$ and $\mathcal{R}'$ are rationalized by P', it follows that $I(P', \sigma(\mathcal{R})) = I(P', \mathcal{R}') = 0$. By (SEP), the addition of observations $\mathcal{B}'$ gives $I(P', \mathcal{B}' \oplus \sigma(\mathcal{R})) = I(P', \mathcal{B}' \oplus \mathcal{R}')$, and the claim follows.$\square$

Lemma 3 implies that I induces a real function (which we denote by I to simplify the notation) from the $\frac{k(k+1)}{2}$-cartesian product of non-negative integers, $\mathbb{N}^{k(k+1)/2}$. Consider the binary relation $\succeq$ over $\mathbb{N}^{k(k+1)/2}$ defined by $\alpha \succeq \beta$ if and only if there are preference relations P and P' and collections of observations $\mathcal{A}$ and $\mathcal{A}'$ such that $\lambda(P, \mathcal{A}) = \alpha, \lambda(P', \mathcal{A}') = \beta$ and $I(P, \mathcal{A}) \geq I(P', \mathcal{A}')$. The strict part of $\succeq$ is denoted by $\succ$. Also, denote by $+$ the usual addition operator on $\mathbb{N}^{k(k+1)/2}$.

We now need to introduce the notion of a *closed extensive structure*. We say that the triple $(\mathbb{N}^{k(k+1)/2}, \succeq, +)$, where $\mathbb{N}^{k(k+1)/2}$ is clearly a non-empty set, $\succeq$ is a binary relation on $\mathbb{N}^{k(k+1)/2}$ and $+$ is a closed binary operation on $\mathbb{N}^{k(k+1)/2}$, is a closed extensive structure whenever the following four properties hold:

(1) Complete Preorder: $\succeq$ is a complete preorder on $\mathbb{N}^{k(k+1)/2}$.
(2) Weak Associativity: For all $\alpha, \beta, \gamma \in \mathbb{N}^{k(k+1)/2}$, $\alpha + (\beta + \gamma) \sim (\alpha + \beta) + \gamma$.

(3) Monotonicity: For all $\alpha, \beta, \gamma \in \mathbb{N}^{k(k+1)/2}$, $\alpha \succeq \beta \Leftrightarrow \alpha + \gamma \succeq \beta + \gamma \Leftrightarrow \gamma + \alpha \succeq \gamma + \beta$.

(4) Archimedean: For all $\alpha, \beta, \gamma, \mu \in \mathbb{N}^{k(k+1)/2}$, if $\alpha \succ \beta$, then there exists a positive integer d such that $d\alpha + \gamma \succeq d\beta + \mu$ where $d\alpha$ is defined inductively as $1\alpha = \alpha$ and $(d + 1)\alpha = d\alpha + \alpha$.

Lemma 4. *The triple* $(\mathbb{N}^{k(k+1)/2}, \succeq, +)$ *is a closed extensive structure.*

Proof of Lemma 4: First, given that for all $\alpha \in \mathbb{N}^{k(k+1)/2}$ there exists a preference relation P and a collection of observations $\mathcal{A}$ such that $\lambda(P, \mathcal{A}) = \alpha$, and given that $I : \mathbb{N}^{k(k+1)/2} \to \mathbb{R}_+$ is a real function, $\succeq$ is a complete preorder on $\mathbb{N}^{k(k+1)/2}$.

Second, notice that the usual addition operator $+$ on $\mathbb{N}^{k(k+1)/2}$ is an associative binary operation and therefore weakly associative.

Third, consider $\alpha, \beta, \gamma \in \mathbb{N}^{k(k+1)/2}$ and let $\alpha \succeq \beta$. By definition, there exist preference relations P and P' and collections of observations $\mathcal{A}$ and $\mathcal{A}'$ such that $\lambda(P, \mathcal{A}) = \alpha, \lambda(P', \mathcal{A}') = \beta$ and $I(P, \mathcal{A}) \geq I(P', \mathcal{A}')$. By (NEU), we can assume, w.l.o.g., that $P = P'$. Furthermore, we can construct a collection of observations $\mathcal{A}''$ such that $\lambda(P, \mathcal{A}'') = \gamma$. Using (SEP) and the fact that $I(P, \mathcal{A}) \geq I(P, \mathcal{A}')$, we obtain that $I(P, \mathcal{A} \oplus \mathcal{A}'') \geq I(P, \mathcal{A}' \oplus \mathcal{A}'')$, or, equivalently, $\alpha + \gamma \succeq \beta + \gamma$. Similarly, it can be proved that $\alpha + \gamma \succeq \beta + \gamma$ implies $\alpha \succeq \beta$, and therefore Monotonicity is satisfied.

Finally, consider $\alpha, \beta \in \mathbb{N}^{k(k+1)/2}$ such that $\alpha \succ \beta$ and consider any $\gamma, \mu \in N^{k(k+1)/2}$. By the definition of $\succeq$, and by (NEU) and (ARCH), there exists a strictly positive integer d such that $d\alpha + \gamma \succeq d\beta + \mu$ and thus the triple $(\mathbb{N}^{k(k+1)/2}, \succeq, +)$ satisfies Archimedean.$\square$

By Theorem 1 in Krantz, Duncan Luce, Suppes, and Tversky (1971, p. 74), Lemma 4 guarantees that there exists a real-valued function ϕ on $\mathbb{N}^{k(k+1)/2}$ such that for all $\alpha, \beta \in \mathbb{N}^{k(k+1)/2}$:

(1) $\alpha \succeq \beta \Leftrightarrow \phi(\alpha) \geq \phi(\beta)$.

(2) $\phi(\alpha + \beta) = \phi(\alpha) + \phi(\beta)$.

Furthermore, another function ϕ' satisfies conditions (1) and (2) if and only if there exists a strictly positive real value ψ such that $\phi' = \psi\phi$. Therefore, given the definition of $\succeq$, it follows immediately that $I = \phi$ up to a scalar transformation, and $I(P, \mathcal{A}) = \sum_{t=1}^{k} \sum_{s=t}^{k} \theta_{ts} \lambda_{ts}(P, \mathcal{A})$, where θ_{ts} is the canonical value of ϕ over the vector whose value is 1 at component ts and 0 elsewhere.

The following lemma states some properties of the parameters θ_{ts}.

Lemma 5.

    (1) *For all $1 \leq t \leq k$, $\theta_{tt} = 0$.*
    (2) *For all $1 \leq t < s \leq k$, $\theta_{ts} > 0$.*
    (3) *For all $1 \leq t \leq s \leq h \leq k$, $\theta_{th} = \theta_{ts} + \theta_{sh}$.*

Proof of Lemma 5: The first two claims follow immediately from the fact that I satisfies (RAT). For the third claim, consider any preference relation P, and alternatives a, b and c, such that the number of alternatives strictly preferred to a, b and c according to P is $t - 1$, $s - 1$ and $h - 1$, respectively. Consider the observation $(\{a, b, c\}, c)$. Clearly, $\{a, b\}$ and $\{c\}$ constitute a partition of $\{a, b, c\}$ where $\{a, b\}P\{c\}$. Therefore, by (COM), $\theta_{th} = I(P, [(\{a, b, c\}, c)]) = I(P, [(\{a, b\}, b), (\{b, c\}, c)]) = I(P, [(\{a, b\}, b)]) + I(P, [(\{b, c\}, c)]) = \theta_{ts} + \theta_{sh}.\square$

To conclude the proof of the theorem, define $w_t = \sum_{s=t}^{k-1} \theta_{s,s+1}$ and $w_k = 0$. Then, for any collection of observations $\mathcal{A}$, and for any preference relation P, $I(P, \mathcal{A}) = \sum_{t=1}^{k} \sum_{s=t}^{k} \lambda_{ts}(P, \mathcal{A})(w_t - w_s)$, and I is a weighted-loss index associated to w, as desired.∎

Proof of Theorem 2: We prove the claim of the theorem by way of two lemmata.

Lemma 6. *For any balanced collection of observations $\mathcal{A}$, and for any two preference relations P and P', $\sum_{t=1}^{k} \delta_t(P, \mathcal{A})w_t = \sum_{t=1}^{k} \delta_t(P', \mathcal{A})w_t = K(\mathcal{A}, w)$, where $K(\mathcal{A}, w)$ is a function that depends only on $\mathcal{A}$ and w.*

Proof of Lemma 6: Consider a balanced collection of observations $\mathcal{A}$, and two preference relations P and P'. Select a permutation σ, such that $P' = \sigma(P)$. Then, it must be that $\sum_{t=1}^{k} \delta_t(P, \mathcal{A})w_t = \sum_{t=1}^{k} \delta_t(\sigma(P), \sigma(\mathcal{A}))w_t = \sum_{t=1}^{k} \delta_t(P', \sigma(\mathcal{A}))w_t$. Since $\mathcal{A}$ is balanced, the menus observed in $\sigma(\mathcal{A})$ are exactly the same as in $\mathcal{A}$. Given that $\sum_{t=1}^{k} \delta_t(P', \sigma(\mathcal{A}))w_t$ does not depend on the chosen elements in $\sigma(\mathcal{A})$, but only on the menus observed, then $\sum_{t=1}^{k} \delta_t(P', \sigma(\mathcal{A}))w_t = \sum_{t=1}^{k} \delta_t(P', \mathcal{A})w_t$. Hence, $\sum_{t=1}^{k} \delta_t(P, \mathcal{A})w_t = \sum_{t=1}^{k} \delta_t(P', \mathcal{A})w_t.\square$

Lemma 7. *For any balanced collection of observations $\mathcal{A}$, and for any two preference relations P and P', $I_w(P, \mathcal{A}) \leq I_w(P', \mathcal{A}) \Leftrightarrow \sum_{s=1}^{k-1} \left(\sum_{l=1}^{s} \rho_l(P, \mathcal{A}) \right) \theta_{s,s+1} \geq \sum_{s=1}^{k-1} \left(\sum_{l=1}^{s} \rho_l(P', \mathcal{A}) \right) \theta_{s,s+1}.$*

Proof of Lemma 7: Consider a collection of observations $\mathcal{A}$, and two preference relations P and P'. Given Lemma 6, we can conclude that

$$I_w(P,\mathcal{A}) \leq I_w(P',\mathcal{A}) \quad \Leftrightarrow \quad \sum_{t=1}^{k} \delta_t(P,\mathcal{A})w_t - \sum_{s=1}^{k} \rho_s(P,\mathcal{A})w_s \leq \sum_{t=1}^{k} \delta_t(P',\mathcal{A})w_t - \sum_{s=1}^{k} \rho_s(P',\mathcal{A})w_s$$

$$\Leftrightarrow \quad \sum_{s=1}^{k} \rho_s(P,\mathcal{A})w_s \geq \sum_{s=1}^{k} \rho_s(P',\mathcal{A})w_s.$$

By way of the inverse construction of w_s used in the proof of Theorem 1, that is, $\theta_{s,s+1} = w_s - w_{s+1}$, we can conclude that

$$\sum_{s=1}^{k} \rho_s(P,\mathcal{A})w_s \geq \sum_{s=1}^{k} \rho_s(P',\mathcal{A})w_s \Leftrightarrow \sum_{s=1}^{k-1} \Big(\sum_{l=1}^{s} \rho_l(P,\mathcal{A})\Big)\theta_{s,s+1} \geq \sum_{s=1}^{k-1} \Big(\sum_{l=1}^{s} \rho_l(P',\mathcal{A})\Big)\theta_{s,s+1},$$

and thus,

$$I_w(P,\mathcal{A}) \leq I_w(P',\mathcal{A}) \Leftrightarrow \sum_{s=1}^{k-1} \Big(\sum_{l=1}^{s} \rho_l(P,\mathcal{A})\Big)\theta_{s,s+1} \geq \sum_{s=1}^{k-1} \Big(\sum_{l=1}^{s} \rho_l(P',\mathcal{A})\Big)\theta_{s,s+1},$$

, which concludes the proof of the lemma.$\square$

Now, the definition of a basic preference relation $P^B(\mathcal{A})$ guarantees that, for any s, and for any preference relation P, $\sum_{l=1}^{s} \rho_s(P^B(\mathcal{A}),\mathcal{A}) \geq \sum_{l=1}^{s} \rho_s(P,\mathcal{A})$. Since $\theta_{s,s+1} > 0$ for every s, $\sum_{s=1}^{k-1} \big(\sum_{l=1}^{s} \rho_l(P^B(\mathcal{A}),\mathcal{A})\big)\theta_{s,s+1} \geq \sum_{s=1}^{k-1} \big(\sum_{l=1}^{s} \rho_l(P,\mathcal{A})\big)\theta_{s,s+1}$ and therefore, by Lemma 7, $I(P^B(\mathcal{A})) \leq (P,\mathcal{A}).\blacksquare$

Proof of Corollary 1: The proof follows immediately by extending Lemmata 6 and 7 to the case where the collections of observations $\mathcal{A}$ and $\mathcal{A}'$, have the same menus of alternatives. Then, the fact that $\rho(P^B(\mathcal{A}),\mathcal{A})$ majorizes $\rho(P^B(\mathcal{A}'),\mathcal{A}')$ simply means that $\sum_{l=1}^{s} \rho_s(P^B(\mathcal{A}),\mathcal{A}) \geq \sum_{l=1}^{s} \rho_s(P^B(\mathcal{A}'),\mathcal{A}')$. Then $\sum_{s=1}^{k-1} \big(\sum_{l=1}^{s} \rho_l(P^B(\mathcal{A}),\mathcal{A})\big)\theta_{s,s+1} \geq \sum_{s=1}^{k-1} \big(\sum_{l=1}^{s} \rho_l(P^B(\mathcal{A}'),\mathcal{A}')\big)\theta_{s,s+1}$, and consequently $I_w^*(\mathcal{A}) \leq I_w^*(\mathcal{A}').\blacksquare$

Proof of Theorem 3: Denote the alternatives in X by $1, 2, \ldots, k$. Consider the permutation σ given by $\sigma(h) = h + 1$ whenever $h < k$ and $\sigma(k) = 1$. For any positive integer q denote by σ^q the composition of σ with itself q times, with the convention $\sigma^0 = \sigma^k$ equal to the identity mapping. Also, given any positive integer q and the number of alternatives in X, k, denote by $a(q,k)$ and $b(q,k)$ the unique positive integers such that (1) $q = a(q,k)k + b(q,k)$, and (2) $b(q,k) < k$.

Consider the menu A_1 observed in $\mathcal{A}$. Consider all the menus $A_i \in \mathcal{A}$ such that there exists $0 \leq r \leq k$ with $A_i = \sigma^r(A_1)$. Denote this collection by Ψ_1. Pick any menu B_1

in Ψ_1 such that $1 \in B_1$. The existence of such a menu is guaranteed by the fact that $\mathcal{A}$ is balanced. Clearly, for every $0 < r \leq k$, $k \in \sigma^{r-1}(B_1)$ whenever $b(r,k) = 0$ and $b(r,k) \in \sigma^{r-1}(B_1)$ whenever $b(r,k) > 0$. Hence, we can design k or $b(r,k)$, respectively, to be the chosen alternative in menus that take the form $\sigma^{r-1}(B_1)$, $0 < r \leq k$.

Suppose that we have defined collections of menus $\Psi_1, \Psi_2, \ldots, \Psi_{v-1}$ and the corresponding chosen alternatives. If $n > |\Psi_1| + |\Psi_2| + \cdots + |\Psi_{v-1}|$, select one menu A_v observed in $\mathcal{A}$, but not present in $\Psi_1, \Psi_2, \ldots, \Psi_{v-1}$. Consider all the menus $A_i \in \mathcal{A}$ such that there exists $0 \leq r \leq k$ with $A_i = \sigma^r(A_v)$. Denote this collection by Ψ_v. Let $b(|\Psi_1| + |\Psi_2| + \cdots + |\Psi_{v-1}|, k) = b_{v-1}$. Pick any menu B_v in Ψ_v such that $b_{v-1} + 1 \in B_v$. Clearly, for every $0 < r \leq k$, $k \in \sigma^{r-1}(B_v)$ whenever $b(b_{v-1} + r, k) = 0$ and $b(b_{v-1} + r, k) \in \sigma^{r-1}(B_v)$ whenever $b(b_{v-1} + r, k) > 0$. Hence we can design k or $b(b_{v-1} + r, k)$, respectively, to be the chosen alternative in menus that take the form $\sigma^{r-1}(B_v)$, $0 < r \leq k$.

This process inductively defines a collection of observations $\mathcal{B}$ with the same menus as $\mathcal{A}$. By Theorem 2, $I_w^*(\mathcal{B}) = K(\mathcal{B}, w) - \sum_{s=1}^k \rho_s(P^B(\mathcal{B}), \mathcal{B})$. Given that $\mathcal{B}$ has the same menus as $\mathcal{A}$, it follows that $K(\mathcal{B}, w) = K(\mathcal{A}, w)$. Furthermore, given the construction of $\mathcal{B}$, it follows that $1, 2, \ldots, b(n,k)$ are chosen in exactly $a(n,k)+1$ menus, while $b(n,k) + 1, \ldots, k$ are chosen in exactly $a(n,k)$ menus. Hence, $\rho_s(P^B(\mathcal{B}), \mathcal{B}) = (a(n,k) + 1, a(n,k) + 1, \ldots, a(n,k) + 1, a(n,k), a(n,k), \ldots, a(n,k))$. Thus,

$$I_w^*(\mathcal{B}) = K(\mathcal{A}, w) - \sum_{s=1}^{b(n,k)} (a(n,k) + 1)w_s - \sum_{s=b(n,k)+1}^{k} a(n,k)w_s$$

$$= K(\mathcal{A}, w) - a(n,k) \sum_{s=1}^{k} w_s - \sum_{s=1}^{b(n,k)} w_s.$$

It follows immediately that $\rho(P^B(\mathcal{A}), \mathcal{A})$ majorizes $\rho(P^B(\mathcal{B}), \mathcal{B})$. Thus, the application of Corollary 1 concludes the proof.∎

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

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

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

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

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

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

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

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

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

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

[12] 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.

[13] Krantz, D.H., R. Duncan Luce, P. Suppes, and A. Tversky (1971), *Foundations of Measurement. Additive and Polynomial Representations.* Academic Press.

[14] Manzini, P. and M. Mariotti (2007), "Sequentially Rationalizable Choice," *American Economic Review*, 97:1824–1839.

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

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

[17] Masatlioglu, Y., D. Nakajima and E.Y. Ozbay (2009), "Revealed Attention," mimeo.

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

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

[20] Rubinstein, A. and Y. Salant (2009), "Eliciting Welfare Preferences from Behavioral Datasets," mimeo.

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

[22] 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.

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

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

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

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