A MEASURE OF RATIONALITY AND WELFARE*

JOSE APESTEGUIA† AND MIGUEL A. BALLESTER‡

Abstract. Evidence showing that individual behavior often deviates from the classical principle of preference maximization has raised at least two important questions: (i) How serious are the deviations? and (ii) What is the best method for extracting relevant information from choice behavior for the purposes of welfare analysis? This paper addresses these questions by proposing an instrument to identify the preference relation closest to the revealed choices and evaluate the inconsistencies in terms of the associated welfare loss. We call this instrument the swaps index.

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

1. Introduction

The standard model of individual behavior is based on the maximization principle, whereby the alternative chosen by the individual is the one that maximizes a well-behaved preference relation over the menu of available alternatives. This has two key advantages. The first is that it provides a simple, versatile, and powerful account of individual behavior. The second is that it suggests the maximized preference relation as a tool for individual welfare analysis. That is, the standard approach to welfare economics involves the policy-maker reproducing the decisions that the individual would have made freely, if given the chance.

Research in recent years, however, has produced increasing amounts of evidence documenting deviations from the standard model of individual behavior. Some phenomena

Date: December, 2013.

* We thank Larbi Alaoui, Jorge Alcalde-Unzu, Stéphane Bonhomme, Guillermo Caruana, Christopher Chambers, Tugce Cuhadaroglu, Eddie Dekel, Jerry Green, Philipp Kircher, Paul Milgrom, Mauro Papi, John K.-H. Quah, Collin Raymond, Larry Samuelson, and Karl Schlag for helpful comments. Financial support by the Spanish Commission of Science (ECO2012-34202, ECO2011-25295, ECO2010-09555-E, ECO2008-04756–Grupo Consolidado C) and FEDER is gratefully acknowledged.

† ICREA, Universitat Pompeu Fabra and Barcelona GSE. E-mail: `jose.apesteguia@upf.edu`.

‡ Universitat Autonoma de Barcelona and Barcelona GSE. E-mail: `miguelangel.ballester@uab.es`.

that have attracted a great deal of empirical and theoretical attention, and which 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 serious are the deviations from the classical theory?

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

Properly addressing Q.1 would enable us to evaluate how accurately the classical theory of choice describes individual behavior. This would shift the focus from whether or not individuals violate the maximization principle to how close their behavior is to this benchmark. Moreover, the ability to assess the distance between actual behavior and behavior consistent with the maximization of a preference relation would provide a unique means to gain a deeper understanding of actual decision-making. Furthermore, the possibility of performing meaningful comparisons of rationality would enable 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 (see section 6 for some recent examples).

Dealing with Q.2, meanwhile, would enable us to distinguish, from an external perspective, alternatives that are good for the individual from those that are bad, even when the individual's behavior is not fully consistent with the maximization principle. This, of course, is of prime relevance since welfare analysis is a core area of economic research.

Although these two questions are intimately related, the literature has treated them separately. This paper provides the first unified treatment of the measurement of rationality and welfare. Relying on standard revealed preference data, we propose an instrument to identify the preference relation that is closest to the revealed choices, by evaluating inconsistencies in the data in welfare terms. We call this instrument the *swaps index*.

The swaps index evaluates the inconsistency in every observation unexplained by a preference relation in terms of the number of available alternatives in the menu that rank higher in the preference relation than the chosen one. That is, it counts the number of alternatives in each menu that must be swapped with the chosen alternative

[1]We review the relevant literature in section 6.

in order to rationalize the individual's choices. Thus, the swaps index is given by the preference relation that minimizes the total number of swaps in all the observations, weighted by their relative occurrence in the data. As in the classical approach, the swaps index uses the revealed choices to suggest a welfare ranking, interpreted as the best approximation to the choices of the individual, and further complements it with a measure of its accuracy: the inconsistency value. It contributes to the measurement of rationality in a singular fashion by evaluating inconsistent behavior in terms of welfare loss, thereby addressing the welfare implications of irrationality.

We then study the swaps index in detail. We begin, in section 3.2 with a simple example to illustrate the contrast between the treatment given by the swaps index to the measurement of rationality and proposals put forward in the literature. The example presents two scenarios that the swaps index treats as diametrically different, but that approaches in the literature may treat counter-intuitively.

In section 3.3 we compare the swaps index approach to welfare analysis with other proposals found in the literature. This exercise illustrates that the preference relation resulting from the swaps index ranks every pair of alternatives in light of the complete set of choice data, not just those observations featuring both alternatives. In principle, therefore, we propose a welfare criterion that provides the policy-maker with a tool to endogenize all the consequences of ranking one alternative over another.

In section 3.4 we apply the swaps index to three prominent cases, and study the preference relation that it identifies. In the first case, the random utility models, the decision-maker has an unambiguous true preference relation that is subject to mistakes, and hence may generate inconsistent revealed choices. We show that the swaps index always identifies the true underlying preference. Then we study the endowment effect, in which decision-makers typically value an alternative more highly when they own it. Here we show that the swaps index identifies the rational preference relation, the one that is independent of the reference point. Finally we study $\beta - \delta$ time preferences and establish that in the space of time-consistent preferences, under certain regularity conditions, the swaps index identifies the long-run preference relation, the one governed by the δ.

We then aim to gain a deeper understanding of the swaps index, by providing its complete axiomatic characterization. In section 4 we propose seven properties that should be satisfied by any inconsistency index relying exclusively on the endogenous information arising from the choice data, and show that they completely characterize the swaps index. Remarkably, this exercise makes the swaps index the first axiomatically-founded inconsistency measure in the literature.

In section 5, we study three novel generalizations of the swaps index, which, in different ways, use different kinds of information exogenous to the revealed choices, which may sometimes be available to the analyst and prove useful in the measurement of rationality and welfare. Importantly, the three proposals have a structural analogy with the swaps index, in that they identify the preference relation that is closest to observed behavior by additively evaluating the inconsistency of the data, measured in welfare terms. The first of the generalizations, which we call the *non-neutral swaps index* makes use of information on the nature of the alternatives, such as their monetary value. Based on this information, the non-neutral swaps index assigns different weights to the various alternatives in the upper contour sets. The second generalization, which we call the *positional swaps index*, is appropriate when information is available on the cardinal utility values of the alternatives based on their position in the ranking. It then weights an inconsistent choice by the sum of the utility value of the forgone alternatives. This can be interpreted as the total utility loss for the inconsistent choice in that observation. The last of our proposed indices, the *general weighted index*, represents a broad generalization of the previous ones. General weighted indices are flexible enough to evaluate the inconsistency of choice by weighting each inconsistent observation by the possible underlying values of the alternatives, the values of the various menus of alternatives, and using specific priors on the plausibility of the different welfare rankings. In section 5 we provide the complete characterizations of these three cases, which, importantly, proceed by relaxing some of the properties characterizing the swaps index.

In section 5 we also study two classical indices within our framework, the Varian and Houtman-Maks indices, and illustrate their structural commonality with the swaps index in that they are part of the general weighted index. In so doing, we offer axiomatic characterizations of these versions of Varian and Houtman-Maks. Finally, we also argue that the recent money pump index of Echenique, Lee and Shum (2011) is fundamentally different from the swaps index, since it is not a special case of the general weighted index.

The organization of the rest of the paper is as follows. Section 6 offers a brief review of the relevant literature and section 7 concludes the paper. All the proofs are contained in the appendix.

2. Framework

Let X be a finite set of k alternatives. Denote by $\mathcal{O}$ the set of all possible pairs (A, a), where $A \subseteq X$ and $a \in A$. We refer to such pairs as *observations*. Individual behavior is summarized by the relative number of times each observation (A, a) occurs in the data. Then, a *collection of observations* f assigns to each observation (A, a) a positive real value denoted by $f(A, a)$, with $\sum_{(A,a)} f(A, a) = 1$, interpreted as the relative frequency with which the individual confronts menu A and chooses alternative a. Denote by $\mathcal{F}$ the set of all such possible collections of observations.

Both the set of alternatives X and the collection of observations f are part of the primitives in our exercise, and hence are taken as given. As is customary in the literature, the grand set of alternatives X is a full description of all the relevant feasible alternatives. The collection f describes all the revealed choices of the individual, which constitute the empirical data to be evaluated in terms of rationality and welfare. The collection f allows for the possibility of accounting different observations with different frequencies. This is natural in empirical applications, where exogenous variations require the decision-maker to confront the menus of alternatives in uneven proportions. Consider, for example, the case of consumption data analysis, where the menus faced by the consumer are dictated by prices and wealth, which do not necessarily change uniformly. Furthermore, for the purposes of the present exercise, it is important to take into account the relative frequencies of choices, which, as shown in section 3.4, may, in some settings, be associated with intensities of preference, and provide crucial information on the underlying welfare ranking.

Another key ingredient in our framework are preference relations. 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 every single observation present in the data can be explained by the maximization of the same preference relation. Denote by $m(P, A)$ the maximal element in A according to P. Then, formally, we say that f is rationalizable if there exists a preference relation P such that $m(P, A) = a$ for every (A, a) with $f(A, a) > 0$. We will often use P^f to denote a preference relation rationalizing the rationalizable collection f. 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 removed from rationalizability, a collection of observations is.

3. The Swaps Index

3.1. **Definition.** Consider a given preference relation P and an observation (A, a) that is inconsistent with the maximization of 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 by the individual. It is natural, therefore, to entertain that the inconsistency of observation (A, a) with respect to P entails considering the number of alternatives in A that rank higher than the chosen one, namely $|\{x \in A : xPa\}|$. These are the alternatives that must be swapped with the chosen one in order to make the choice of a consistent with the maximization of P. It then follows that, weighting every single observation by its relative occurrence in the data, the inconsistency of f with respect to P would be $\sum_{(A,a)} f(A, a)|\{x \in A : xPa\}|$. The swaps index I_S adopts this criterion and finds the preference relation P_S that minimizes the weighted sum of swaps. That is:

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

The swaps index has several attractive characteristics that make it unique. Notably, it enables the joint treatment of inconsistency and welfare analysis. It identifies the preference relation closest to the revealed data, measuring its inconsistency in terms of the associated welfare loss. It discriminates between different degrees of inconsistency in the various choices, relying exclusively on the information contained in the choice data, and additively considers every single inconsistent observation weighted by its relative occurrence in the data. Moreover, one can easily show that almost all collections of observations f have a unique optimal preference relation P_S.[2] We now turn to the study of various important issues relating to the swaps index.[3]

3.2. **The Measurement of Rationality: A Comparison.** We illustrate the swaps index treatment of rationality assessment by way of a simple example, and contrast this with other proposals in the literature. The example shows that the swaps index discriminates sharply between different situations that may be treated counter-intuitively by the inconsistency measures offered in the literature.

Consider the set of alternatives $X = \{1, \ldots, k\}$, with $k > 2$. Suppose that the collection f is completely consistent with the preference relation P, ranking the alternatives

[2]Formally, P_S is a mapping from $\mathcal{F}$ to $\mathcal{P}$; in order to avoid excess notation we write P_S instead of $P_S(f)$.

[3]In addition, in Appendix A we deal with the computational complexity of obtaining P_S in practice.

as $1P2P\dots Pk$. Now assume two different scenarios, both involving the consistent evidence f with a high frequency, say $(1-\alpha)$, and one inconsistent observation with a low frequency α. In scenario I we observe that the individual chooses option k from menu X, while in scenario II we observe that the individual chooses option 2 from menu X. Clearly, the two scenarios are inconsistent, since we are assuming that there is sound evidence indicating that the individual should have chosen option 1 in both scenarios. This raises the question of how inconsistent these decisions are. The swaps index sees the two scenarios as representing markedly different situations. Scenario I shows high inconsistency, since the individual chooses the worst possible alternative, alternative k, ignoring all those remaining, which have, in fact, been shown to be better than the selected alternative k. Scenario II, also shows some inconsistency with the maximization principle, but, from the swaps index perspective, this inconsistency is orders of magnitude lower, since it involves choosing the second best available option, that is, option 2. Rationalization of the individual's behavior in scenario II requires ignoring only of alternative 1, while the case of scenario I requires ignoring of every single alternative in X except the chosen one, k. Hence, the swaps index discriminates between the two different scenarios in an intuitive way, offering an unambiguous answer based on the welfare implications of the inconsistent choice.

We now turn to the treatment given to the two scenarios under the classical proposals in the literature. To the best of our knowledge, the first method to measure inconsistency of behavior was proposed by Afriat (1973). In a consumer setting, Afriat suggests measuring the degree 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 all budget sets shrink, thus eliminating some revealed information, and thereby possibly removing some inconsistencies from the data. Then, the degree of inconsistency of a collection of observations proposed by Afriat is associated with the minimal wealth adjustment needed to make all the data consistent with the maximization principle. Therefore, Afriat's judgement of scenarios I and II depends crucially on some external structure, such as the monetary values of the alternatives. The latter, of course, need not necessarily agree with the welfare ranking, and hence may lead to conclusions contradicting the intuitive view that scenario I shows greater inconsistency. For example, if the monetary value of option k is higher than that of option 2, Afriat would judge scenario II more inconsistent than scenario I, since it would require a larger wealth adjustment to remove it from the data.

Varian (1990) extends Afriat to contemplate different relative wealth adjustments in the different observations, and then consider the aggregated relative wealth adjustments that would be required to prevent all the inconsistencies from being revealed. In our example, since both scenarios involve a single inconsistent observation with the same weight, Afriat and Varian coincide.

Houtman and Maks (1985) propose considering the minimal subset of observations that needs to be removed from the data in order to make the remainder rationalizable. The size of this minimal subset to be discarded suggests itself as a measure of inconsistency. In our example, since the sizes of the inconsistencies are identical in both scenarios, Houtman and Maks do not discriminate between them. Dean and Martin (2012) suggest an extension of Houtman and Maks, the HM-e index, which weights the binary comparisons of the alternatives by their monetary values. Hence, the HM-e index depends, like the Afriat index, on this kind of exogenous information.

Finally, rationality has also been measured by counting the number of times in the data a consistency property, such as IIA or GARP, is violated (see, e.g., Swofford and Whitney, 1987; Famulari, 1995). It turns out that the conclusions of this criterion in our example depend on the specific nature of the consistent part of the collection of observations, whereby the two scenarios may be treated alike, or scenario II may even be regarded as more inconsistent. Recently, Echenique, Lee and Shum (2011) make use of the monetary structure of budget sets to suggest a new measure, the money pump index, which evaluates not only the number of times GARP is violated, but also the severity of each violation. Their proposal is to weight every cycle in the data by the amount of money that could be extracted from the consumer. They then consider the total wealth lost in all the revealed cycles. The money pump index judges the two scenarios both on the specific nature of the rationalizable collection f and, like the Afriat index, on exogenous information on the monetary values of the alternatives.

3.3. **The Measurement of Welfare: A Comparison.** Let us illustrate our approach to welfare analysis by contrasting it with two pioneering proposals in the literature, Bernheim and Rangel (2009) and Green and Hojman (2009). Interestingly, although these two papers tackle the problem from different angles, they 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, x is ranked above y in the welfare ranking $\overline{P}$ if y is never chosen when x is available. Bernheim and Rangel show that whenever every menu A in X is present in the data, $\overline{P}$ is acyclic, and hence consistent with the maximization principle.

We now examine the relationship between $\overline{P}$ and the optimal preference relation of the swaps index P_S. It turns out to be the case that the two welfare relations are fundamentally different. That P_S is not contained in $\overline{P}$ follows immediately since P_S is a linear order, while $\overline{P}$ is 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 taking into account only those menus of alternatives where both x and y are available, P_S takes all the data into consideration. Hence, P_S and $\overline{P}$ may rank two alternatives in opposite directions. A simple example illustrates this point.

Consider a collection f where: $f(\{x,y\},x) = f(\{y,z\},y) = \frac{1-2\epsilon}{2}$ and $f(\{x,y,z\},y) = f(\{x,z\},z) = \epsilon$, where ϵ is small. 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 swaps index considers the whole collection f. Data $f(\{x,y\},x) = f(\{y,z\},y) = \frac{1-2\epsilon}{2}$ signify a strong argument for the preference ranking x over y and y over z, and consequently x over z. This preference implies that the data $f(\{x,y,z\},y) = f(\{x,z\},z) = \epsilon$ are not accounted for, but rationalizes the more frequent evidence of $(\{x,y\},x)$ and $(\{y,z\},y)$. In fact, for every sufficiently small ϵ such a preference is the optimal preference relation P_S for the swaps index, and hence $\overline{P}$ and P_S may follow different directions.

3.4. **Identifying the Underlying Preference Relation.** Here we study three settings, very diverse in nature, and study the preference relation that the swaps index uncovers from the possibly inconsistent data. In a nutshell, we show that the swaps index identifies (i) in the random utility models, the true underlying preference, the one that is not subject to mistakes, (ii) in the endowment effect, the rational preference, the one that is not subject to the distortion of the reference point, and (iii) in the $\beta-\delta$ model, among the time-consistent preferences, the long-run preference governed by the δ parameter.

3.4.1. *Random Utility Models.* Consider a situation where the decision-maker has in mind a preference relation over the alternatives, but when the time comes to select her preferred option, she mistakenly chooses a suboptimal alternative. Mistakes may arise from lack of attention, errors of calculation, misunderstanding of the choice situation, or a 'trembling hand' when about to select the desired alternative, etc. This is the case in the highly influential and widely used random utility models.[4]

Consider a utility function $u : X \to \mathbb{R}_{++}$, that assigns to each alternative a cardinal utility value.[5] The utility function u represents the true preference relation of the individual P^u, that is xP^uy if and only if $u(x) > u(y)$. However, the true valuation of the alternatives $u(x)$ is subject to random shocks $\epsilon(x)$, resulting in the final valuation of the alternatives as $U(x) = u(x) + \epsilon(x)$. That is, the valuation $U(x)$ depends on the true utility function $u(x)$, but also on a random mistake variable $\epsilon(x)$, which, as is customary in the literature, we assume to be continuous i.i.d. Different random utility models assume different distributions. For example, when $\epsilon(x)$ is an Extreme Value Type I random variable, we obtain the prominent Luce model.

We now formalize the way in which a random utility model generates a collection of observations. First, let us put the case that the individual faces the menus of alternatives with a certain probability distribution ρ, where $\rho(A)$ denotes the probability of confronting $A \subseteq X$. Then, each observation is generated by a draw that is independent both of the menu and of the shocks in the alternatives contained in the menu, with the individual choosing the alternative that maximizes U. We can now define the collection of observations generated by the random utility model associated with the parameters (ρ, u, ϵ), denoted by $f_{\rho,u,\epsilon}$, as $f_{\rho,u,\epsilon}(A, a) = \rho(A)Pr[a = \arg\max_{x\in A} U(x)]$, for every $(A, a) \in \mathcal{O}$, where $Pr[a = \arg\max_{x\in A} U(x)]$ denotes the probability by which alternative a is the maximal alternative in A according to U.[6]

Our next result establishes that the swaps index identifies the true underlying preference P^u for every collection of observations generated by a random utility model. It is particularly interesting that when all the menus are present in the data, then the swaps index uniquely identifies the preference P^u.[7]

Theorem 1. *For every collection of observations generated by a random utility model, $f_{\rho,u,\epsilon}$, the preference P^u is an argument that minimizes the swaps index. Moreover,*

[4]Classic references are Luce (1959) and McFadden (1974). More recent developments are Gul, Natenzon and Pesendorfer (2012) and Manzini and Mariotti (2013).

[5]For simplicity of exposition, assume that $u(x) \neq u(y)$ for every $x, y \in X$, $x \neq y$.

[6]Notice that, since $\epsilon(x)$ is continuously distributed, the probability of ties is zero and hence $Pr[a = \arg\max_{x\in A} U(x)]$ is well-defined.

[7]We obtain an equivalent result for the tremble model of mistakes (see Harless and Camerer, 1994), where these take the form of constant probability shocks, as in the trembling hand equilibrium of game theory. We could provide the details upon request.

whenever $\rho(A) > 0$ holds for every menu A, the preference P^u is the unique argument that minimizes the swaps index.

3.4.2. *Endowment Effect.* There is a large literature supporting the view that a decision-maker typically values an alternative more highly when she owns it, than otherwise; this is the so-called endowment effect. The endowment effect has profound implications for the understanding of the valuation of objects, contingent on property rights, and it has been shown to play an important role in a variety of settings like the housing market, the stock market, etc.[8] Here we adapt the revealed preference model of Masatlioglu and Ok (2005, 2013) to our setting, and show that, under certain regularity conditions, the swaps index identifies the underlying *rational* preference relation.

An extended menu A^r is composed of a menu A and a reference point $r \in A \cup \{\diamond\}$. The case of $r = \diamond \notin X$ represents the case where the decision-maker chooses without a reference point, as in the standard rational model. The case $r \in A$ represents the case in which at the time of contemplating menu A, the decision-maker has in mind alternative $r \in A$, which may bias the choice from A, since alternative r get an "utility boost". Consider then a function $u : X \to \mathbb{R}$ and a parameter $\phi \in \mathbb{R}_+$ such that, for every extended menu A^r, the selected alternative is the argument in A that maximizes $u(x) + \phi\delta_{r=x}$. The indicator function $\delta_{r=x}$ gives a value of 1 whenever the reference point r is the alternative x, and 0 otherwise.[9] Consequently, option r gets the "utility boost" ϕ if and only if r is the reference point. When there is no reference point, i.e. $r = \diamond \notin X$, the individual simply maximizes the standard utility function $u(x)$. Denote by P^u the ordinal preference relation represented by $u(x)$.

Let us denote by $\rho(A^r)$ the probability of confronting the extended menu A^r. The collection of observations generated by the endowment effect model associated with the parameters (ρ, u, ϕ), denoted by $f_{\rho,u,\phi}$, is $f_{\rho,u,\phi}(A, a) = \sum_{r \in A\cup\{\diamond\}} \rho(A^r)\delta_{a=\arg\max_A u(x)+\phi\delta_{r=x}}$.

We say that a collection of observations $f_{\rho,u,\phi}$ is *reference-regular*, if for every menu A, and for every pair of alternatives $x, y \in A$, $u(x) > u(y)$ implies that $\rho(A^x) \geq \rho(A^y)$. Intuitively, this represents the case where stronger reference points are found more often as reference points in the data. Our next result establishes that whenever the

[8]See Thaler (1980) and Kahneman, Knetsch and Thaler (1990). See also Genesove and Mayer (2001) and Barberis and Xiong (2009).

[9]Let us assume that u and ϕ are such that for every extended menu the set of selected alternatives is a singleton.

collection of observations is reference-regular, the swaps index identifies the preference P^u.

Theorem 2. *For every reference-regular collection of observations $f_{\rho,u,\phi}$, the preference P^u is an argument that minimizes the swaps index. Moreover, whenever $\rho(A^\diamond) > 0$ holds for every menu A, the preference P^u is the unique argument that minimizes the swaps index.*

3.4.3. *$\beta - \delta$ Preferences.* The standard model of time preferences uses an exponential discount function, implying that the preferences of the decision-maker over time are time-consistent. However, evidence shows that people tend to use greater discount rates for the short-run than for the long-run, which has lead to the adoption in many situations of a hyperbolic discount function. The $\beta - \delta$ model here studied is the most influential model using a hyperbolic discount function.[10]

Let $O = \{o_1, \ldots, o_k\}$ be a set of monetary outcomes, with $o_1 < o_2 < \cdots < o_k$. Let $u : O \to [0, 1]$ be a utility function over outcomes. For tractability, we assume that O is sufficiently large and utilities are distributed uniformly.[11] An alternative $x = (o, t)$ is a pair describing an outcome $o \in O$ dated in a certain moment in time $t \in \{0, 1, \ldots, T\}$. Therefore, $X = O \times \{0, 1, \ldots, T\}$.[12]

In the $\beta - \delta$ model, the valuation of (o, t) is $U_{\beta,\delta}(o, t) = u(o)$ whenever $t = 0$ and $U_{\beta,\delta}(o, t) = \beta\delta^t u(o)$ whenever $t > 0$, with $\beta, \delta \in [0, 1]$.[13] Denote by $P^{U_{\beta,\delta}}$ the preference relation over X represented by such utility function and by $\rho(A)$ the probability of confronting menu $A \subseteq X$. Then, the collection of observations generated by the $\beta - \delta$ model associated with the parameters $(\rho, U_{\beta,\delta})$, denoted by $f_{\rho,U_{\beta,\delta}}$, is $f_{\rho,U_{\beta,\delta}}(A, a) = \rho(A)$ whenever $a = \arg\max_{x \in A} U_{\beta,\delta}(x)$, and $f_{\rho,U_{\beta,\delta}}(A, a) = 0$ otherwise.

It is convenient to notice that the $\beta - \delta$ model departs from the standard exponential model only in the discount function. That is, completeness and transitivity still hold, and hence it follows immmediately that for every collection of observations $f_{\rho,U_{\beta,\delta}}$, the preference $P^{U_{\beta,\delta}}$ is an argument that minimizes the swaps index. Moreover, whenever $\rho(A) > 0$ holds for every menu A, the preference $P^{U_{\beta,\delta}}$ is the unique argument that

[10]See Strotz (1956) and Laibson (1997). See also O'Donoghue and Rabin (1999) and Blow and Crawford (2013).

[11]Formally, this implies that $|\{i : u(o_i) \in [a, b]\}|/k$ can be approximated by $(b - a)$.

[12]This is the time preference setting of Fishburn and Rubinstein (1982) and more recently of Ok and Masatlioglu (2007).

[13]We assume that $U_{\beta,\delta}(x) \neq U_{\beta,\delta}(y)$ for every $x, y \in X$.

minimizes the swaps index. Indeed, in any case, the inconsistency associated to the collection of observations is zero, and therefore the swaps index always identifes the underlying preference relation $P^{U_{\beta,\delta}}$ with no associated error.

In occasions, however, we are interested in the best preference relation within a subset of the set of linear orders. In particular, in the present context, we wonder about the *time-consistent preference* that best explains the data. That is, we aim to judge a possibly time-inconsistent collection of observations, from the perspective of the standard exponential time preference model. The evaluation of alternative (o,t) in the standard discounted utility model is $V_\alpha(o,t) = \alpha^t u(o)$ for every t, with $\alpha \in [0,1]$.[14] Let P^{V_α} denote the preference relation over X represented by such utility function, and by $\mathcal{P}^S$ the set of preference relations over X that admit such stationary representation. Clearly, $\mathcal{P}^S \subset \mathcal{P}$. The purpose now is to identify the preference in $\mathcal{P}^S$ that best approximates the data $f_{\rho,U_{\beta\delta}}$.

In the second part of the following result we assume that there is no bias on the different timings of the alternatives. That is, we say that the collection of observations is *time-regular* if the frequency with which each time appears in the data is equal and independent of the outcomes.

Theorem 3. *For every collection of observations generated by a $\beta - \delta$ model, $f_{\rho,U_{\beta,\delta}}$, there exists an $\alpha \in [\beta\delta, \delta]$ such that the stationary preference relation given by α minimizes the swaps index in $\mathcal{P}^S$. Moreover, there is a $\hat{T}$ such that whenever $f_{\rho,U_{\beta,\delta}}$ is time-regular and $T > \hat{T}$, the stationary preference relation given by δ is the unique minimizer.*

Theorem 3 establishes that, under a regularity condition, the unique preference relation identified by the swaps index is the long-run preference of the $\beta - \delta$ model, the one governed by δ. The proof of Theorem 3 shows that the threshold $\hat{T}$ can be very low in practice. It is shown that when $f_{\rho,U_{\beta,\delta}}$ is a time-regular collection composed by the binary menus $\hat{T} = 3$.

4. A Characterization of the Swaps Index

In this section we propose seven conditions that shape the treatment that an inconsistency index I may give to different sorts of collections of observations. We then show

[14] Again, assume that $V_\alpha(x) \neq V_\alpha(y)$ for every $x, y \in X$.

that the swaps index is characterized by this set of properties. As we will argue in section 5, the first four properties are the minimal set of properties that any inconsistency function must possess, while the appeal of the last three properties may depend on the possible availability of additional external information on the nature of the menus, the nature of the alternatives, or both. For the time being, when no information other than the revealed choices is assumed to be available, we contend that all seven properties are desirable and that any inconsistency function I should ideally satisfy them.

Continuity (CONT). I is a continuous function. That is, for every sequence $\{f_n\} \subseteq \mathcal{F}$, if $f_n \to f$, then $I(f_n) \to I(f)$.

This is the standard definition of continuity, and its justification is in turn the standard one. Namely, it is desirable that the inconsistency value does not change abruptly when there is a small variation in the data.

Rationality (RAT). For every $f \in \mathcal{F}$, $I(f) = 0$ if and only if f is rationalizable.

Rationality imposes that a collection of observations is perfectly consistent if and only if the collection is rationalizable. In line with the maximization principle, Rationality establishes that the minimal inconsistency level of 0 is reached only when every single choice in the collection of observations can be explained by maximizing the same preference relation.

Concavity (CONC). I is a concave function. That is, for every $f, g \in \mathcal{F}$ and every $\alpha \in [0,1]$, $I(\alpha f + (1-\alpha)g) \geq \alpha I(f) + (1-\alpha)I(g)$.

To illustrate that this is a desirable property in our context, take any two collections of observations f and g, and suppose that these are rationalizable when taken separately. Clearly, a convex combination of f and g does not need to be rationalizable and hence, the collection of observations $\alpha f + (1-\alpha)g$ can only take the same or a higher inconsistency value than the combination of the inconsistency values of the two collections separately. The same idea applies when either f or g or both are not rationalizable. The combination of f and g can only generate the same or more frictions with the maximization principle, and hence should yield the same or a higher inconsistency value.

In practical data analysis, for ease of procedure, it is often desirable for the function of study to be linear. In our context, linearity would imply that for every $\alpha \in [0,1]$, $I(\alpha f + (1-\alpha)g) = \alpha I(f) + (1-\alpha)I(g)$. Hence, we could compute the inconsistency value of a combination of two collections of observations by simply combining the corresponding inconsistency values. However, this property is too strong to be imposed on an inconsistency function, as the above discussion of Concavity shows. Note that linearity and the property of Rationality are, in fact, logically inconsistent. This is because Rationality implies that the inconsistency is 0 for the rationalizable collections of observations. Since any collection of observations can be expressed as the combination of rationalizable collections, linearity would imply that the inconsistency of any collection is 0, but this contradicts Rationality.

We now introduce a property, Piecewise Linearity, that solves this difficulty by minimally departing from the notion of linearity. Piecewise Linearity imposes linearity not on the entire domain $\mathcal{F}$, but only within subdomains of $\mathcal{F}$ that are organized on the basis of rationalizable collections of observations. Let $\mathcal{R}$ be the set of rationalizable collections of observations that assign the same relative frequency to each possible menu of alternatives $A \subseteq X$. Notice that every collection $r \in \mathcal{R}$ is rationalized by a unique preference relation P^r.[15] Piecewise Linearity creates as many subdomains as elements in $\mathcal{R}$, or equivalently, as preferences on X. The different collections of observations are then organized in such subdomains. To emphasize, Piecewise Linearity implies that every collection of observations is associated with a rationalizable collection, and thus is judged from the perspective of a preference relation. This is most natural when one aims to understand departures from the classical notion of rationality, as in our case. Formally, a cover $\mathcal{C} = \{\mathcal{C}_j\}$ of $\mathcal{F}$ is *rationally founded* if each subdomain $\mathcal{C}_j$ contains one and only one element r_j of $\mathcal{R}$.[16]

Piecewise Linearity (PWL). I is a piecewise linear function with respect to a rationally founded cover $\mathcal{C}$. That is, for every j, every $f, g \in \mathcal{C}_j$ and every $\alpha \in [0, 1]$, $I(\alpha f + (1-\alpha)g) = \alpha I(f) + (1-\alpha)I(g)$.

Continuity, Rationality, Concavity and Piecewise Linearity are four general properties that should be satisfied by every inconsistency index aiming to measure the distance between actual behavior and the maximization of a preference relation. This will become apparent in the next section. We now introduce three more properties that are attractive when no information beyond the revealed data is available. To do

[15]The purpose here is to create a bijection between $\mathcal{P}$ and a set of rationalizable collections of observations. The set $\mathcal{R}$ is just an intuitive way of creating this bijection, that comes without loss of generality.

[16]To recall, for $\mathcal{C}$ to be a cover of $\mathcal{F}$ implies that $\cup_j \mathcal{C}_j = \mathcal{F}$.

so, let us first consider the following notation. We denote by $\mathbf{1}_{(A,x)}$ the collection of observations placing all the mass on (A, x).

Ordinal Consistency (OC). For every $(A, x) \in \mathcal{O}$ and every $r, r' \in \mathcal{R}$ with $r(\{a, b\}, a) = r'(\{a, b\}, a)$ for all $a, b \in A$, $I(\alpha\mathbf{1}_{(A,x)} + (1-\alpha)r) = I(\alpha\mathbf{1}_{(A,x)} + (1-\alpha)r')$ for any sufficiently small $\alpha > 0$.

Consider the combinations of two rationalizable collections of observations, r and r', with $\mathbf{1}_{(A,x)}$. Under certain conditions, Ordinal Consistency imposes that the inconsistency levels associated with the two resulting combinations should be equal. Intuitively, whenever the two rationalizable collections are sufficiently prevalent in the respective combinations, and their associated preference relations, P^r and $P^{r'}$, coincide exactly in the valuation of every single possible pair of alternatives in A, then the resulting inconsistencies should be the same.

The purpose of the requirement that the rationalizable collections of observations should be sufficiently prevalent, or equivalently, that α should be sufficiently small, is to guarantee that $\mathbf{1}_{(A,x)}$ is judged, in both cases, from the perspective of the corresponding rationalizable collections. This ensures that possible inconsistencies in the combined collections are evaluated from the viewpoint of the corresponding preference relations P^r and $P^{r'}$, that coincide in the ranking of all the alternatives in A. Otherwise, if the weight of $\mathbf{1}_{(A,x)}$ is large, the inconsistencies associated with the combined collections $\alpha\mathbf{1}_{(A,x)} + (1-\alpha)r$ and $\alpha\mathbf{1}_{(A,x)} + (1-\alpha)r'$ may be assessed from the standpoint of preference relations other than P^r and $P^{r'}$, in which case the inconsistencies of the combined collections may be very different in nature. In the latter case, it would make little sense to impose that the inconsistency index I should treat the two cases alike.

Disjoint Composition (DC). For every $(A_1, x), (A_2, x) \in \mathcal{O}$ with $A_1 \cap A_2 = \{x\}$ and every $r \in \mathcal{R}$, $I(\alpha\mathbf{1}_{(A_1 \cup A_2, x)} + (1-\alpha)r) = I(\alpha\mathbf{1}_{(A_1,x)} + \alpha\mathbf{1}_{(A_2,x)} + (1-2\alpha)r)$ for any sufficiently small $\alpha > 0$.

Disjoint Composition establishes that, under very special circumstances, two collections of observations can be merged without affecting the inconsistency level. Take the collections $\mathbf{1}_{(A_1,x)}$ and $\mathbf{1}_{(A_2,x)}$ where menus A_1 and A_2 share the same chosen alternative x and nothing else. Suppose that these two collections are combined with a rationalizable collection r, provided that the rationalizable collection r is sufficiently prevalent.[17] Then, Disjoint Composition implies that the collections $\mathbf{1}_{(A_1,x)}$ and $\mathbf{1}_{(A_2,x)}$ can be merged into $\mathbf{1}_{(A_1 \cup A_2,x)}$, while respecting the prevalence of r, and with no consequences for the the inconsistency value.

In order to introduce our last property, let us consider the following definition. Given a permutation σ 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 σ and every $f \in \mathcal{F}$, $I(f) = I(\sigma(f))$.

Neutrality imposes that the inconsistency index should be independent of the names of the alternatives. That is, any relabeling of the alternatives should have no effect on the level of inconsistency. Theorem 4 states the characterization result.

Theorem 4. *An inconsistency index I satisfies CONT, RAT, CONC, PWL, OC, DC and NEU if and only if it is a positive scalar transformation of the swaps index.*

Remark 1. We now comment on the intuition of the main steps involved in the proof of the sufficiency part of Theorem 4.

(i) By way of Piecewise Linearity, we start with a rationally founded cover of all collections of observations over which the index is piecewise linear. In steps 1 to 3 in the proof of Theorem 4, we show, using Rationality, Continuity and Piecewise Linearity, how we can extend each of the classes of the cover to the convex hull of its closure to obtain another rationally founded cover, over which the index is piecewise linear. Moreover, we show that the rationalizable collections of observations are not in the boundary of the new classes.

(ii) In step 4 we assign to each preference P and each observation (A,a), a weight $w(P,A,a)$. To enable this, we argue that any collection $\mathbf{1}_{(A,a)}$ can be combined with any rationalizable collection r such that the resulting collection of observations belongs to the same subdomain of r, provided that the rationalizable collection r is sufficiently prevalent. That is, since r is not in the boundary of its class, there is a value α^r such that for every $\alpha \leq \alpha^r$, and every (A,a), the collection of observations $\alpha\mathbf{1}_{(A,a)} + (1-\alpha)r$ is in the subdomain of r. Then, we can define $w(P^r,A,a) = \frac{I(\alpha^r\mathbf{1}_{(A,a)}+(1-\alpha^r)r)}{\alpha^r}$. That is, we set the weight of the inconsistency associated with the choice of a from menu

[17] The intuition for the requirement that the rationalizable collection of observations r should be sufficiently prevalent in the combined collections is analogous to that of Ordinal Consistency, discussed above. This also applies to those of the following properties that impose this condition.

A from the viewpoint of P^r, $w(P^r, A, a)$, as the normalized value of the inconsistency generated by $\mathbf{1}_{(A,a)}$ when combined with the rationalizable collection r at the level α^r. This guarantees that the resulting combination belongs to the subdomain of r, and hence can be judged from the perspective of P^r.

(iii) In step 5, linearity allows us to compute the inconsistency of the collection f belonging to the subdomain of r, as $I(f) = \sum_{(A,a)} f(A,a)w(P^r, A, a)$. That is, we can separate the inconsistency of f on the basis of the different observations (A, a) and weight each of these observations by the constructed weights $w(P^r, A, a)$ and its relative occurrence in the data $f(A, a)$.

(iv) In step 6, by Concavity we can show that the value of f is indeed minimal across all the preference relations, i.e., $I(f) = \sum_{(A,a)} f(A,a)w(P^r, A, a) \leq \sum_{(A,a)} f(A,a)w(P', A, a)$ for any other preference P'. We also show there that $w(P^r, A, a) = 0$ if and only if a is maximal in A according to P^r. Since all the steps in this part of the proof use only the four mentioned properties and the index we have represented so far takes the form $I(f) = \min_P \sum_{(A,a)} f(A,a)w(P, A, a)$, this gives the intuition of the characterization of the general weighted index I_G of Proposition 3, to be studied later in section 5.3.

(v) In steps 7 to 9 we show the extra implications of Ordinal Consistency, Disjoint Composition and Neutrality, and in the final step 10 we combine all these to show that the index must be a scalar transformation of the swaps index. Intuitively, applying Ordinal Consistency and Disjoint Composition, we show that for every (A, a) and every pair of preferences P and P' such that coincide in their ranking of all the alternatives in A, $w(P, A, a) = \sum_{y \in A} w(P, \{a, y\}, a) = \sum_{y \in A} w(P', \{a, y\}, a) = w(P', A, a)$. That is, as long as the preference relations give the same upper contour set, the weight does not depend on the particular preference relation and can, in fact, be additively decomposed into the binary sets encompassing the choice from A and the elements in the upper contour set. Finally, Neutrality shows that $w(P, \{x, y\}, x) = w(P', \{z, t\}, z)$, provided that the comparison of x and y by P is the same as the comparison of z and t by P'. Taking these results together, we then show that we can write $I(f) = \min_P \sum_{(A,a)} f(A,a)w(P^r, A, a) = K \min_P \sum_{(A,a)} f(A,a)|\{x \in A : xP^r a\}|$, with $K > 0$, which shows that I is a positive scalar transformation of the swaps index.

Remark 2. Piecewise Linearity organizes the collections of observations on the basis of the rationalizable collections $\mathcal{R}$, and imposes linearity within the different subdomains. We have argued that this is a natural approach when one aims to understand the distance of the revealed choices of an individual with respect to the preference maximization model. Here, we note that exactly the same logic can be applied to measure their distance with respect to other models of reference. In settings involving uncertainty, for example, one may want to measure the distance with respect to the subset of $\mathcal{R}$ that satisfies independence; or, in the case of time preferences, with respect to the subset of $\mathcal{R}$ that satisfies stationarity, as we have done in section 3.4.3, etc. Furthermore, the same approach can be adopted for other non-rationalizable models that might be taken as the benchmark of comparison, such as reference-dependent models, etc. In this case, the choice datasets would be organized on the basis of the different models of reference-dependence. Thus, we believe that our characterization exercise is very flexible and may be helpful in settings other than the one studied here. It can be taken as providing a framework for the axiomatic development of inconsistency indices measuring the distance between behavior and other relevant models of choice.

5. Extensions of the Swaps Index

One of the features of the swaps index is its exclusive reliance on the endogenous information contained in the revealed choices. This makes the swaps index particularly interesting and amenable for use in applications. On occasions, however, the analyst may have more information and may want to use it to assess the consistency of choice, and identify the optimal welfare ranking. For this reason, in this section we propose three extensions, the non-neutral swaps index, the positional swaps index and the general weighted index, which vary in the use they make of information external to the revealed preferences. Importantly, the three generalizations follow readily by relaxing some of the characterizing properties of the swaps index. The general weighted index, characterized by the first four axioms of Continuity, Rationality, Concavity and Piecewise Linearity, emerges as the broadest class of indices in our approach. In addition, we also study the classical Varian and Houtman-Maks indices within our framework, and demonstrate their structural commonality with the swaps index, in that they are part of the general weighted index. All these characterizations contribute to a deeper understanding of the indices. Finally, we also argue that the recent money pump index of Echenique, Lee and Shum (2011) is fundamentally different from the swaps index, since it is not a special case of the general weighted index.

5.1. **Non-Neutral Swaps Index.** Suppose that the analyst has at her disposal additional information on the nature of the alternatives, and that she wants to use it in the assessment of rationality and welfare. In this case, it seems logical to argue that it is not just the number of alternatives in the upper contour set that should be relevant, but also their nature. For example, the analyst may consider a metric among the alternatives, say their monetary values or an aggregation of their attributes, and judge each alternative in the upper contour set by its distance from the chosen alternative. Formally, this would imply the following inconsistency index, which we call the non-neutral swaps index,

$$I_{NNS}(f) = \min_{P} \sum_{(A,a)} f(A,a) \sum_{x \in A : xPa} w_{x,a}$$

where $w_{x,a} \in \mathbb{R}_{++}$ denotes the weight of the ordered pair of alternatives x and a.

Note that with regard to the characterizing properties of Theorem 4, NEU immediately loses its appeal, since now one wishes to treat different pairs of alternatives differently, using the exogenous information that is available on them. It turns out that the remaining six properties characterize the non-neutral swaps index.

Proposition 1. *An inconsistency index I satisfies CONT, RAT, CONC, PWL, OC and DC if and only if it is a non-neutral swaps index.*

5.2. **Positional Swaps Index.** We now suggest another novel extension of the swaps index, one that uses not only the endogenous information arising from the revealed preference, but also some exogenous *cardinal* information on the value of the alternatives. We call it the positional swaps index. Suppose that the analyst has information on the cardinal utility values of the different alternatives, based on their position in the ranking.[18] Then, the positional swaps index evaluates an inconsistent observation according to the utility values of the foregone alternatives with respect to that of the chosen alternative. This can be interpreted as the total utility loss due to the inconsistent choice in that observation. Then, the positional swaps index is given by the preference relation that minimizes the sum of utility losses. Denote by $\widehat{x}(P)$ the ranking of alternative x in P. An inconsistency index is a positional swaps index if

$$I_{PS}(f) = \min_{P} \sum_{(A,a)} f(A,a) \sum_{x \in A : xPa} w_{\widehat{x}(P),\widehat{a}(P)}$$

[18]For example, the cardinal utility values may be modeled as 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 expected cardinal values of the alternatives based on their ranking, as in Apesteguia, Ballester and Ferrer (2011).

where $w_{i,j} \in \mathbb{R}_{++}$ denotes the weight associated with the positions i and j.

The main distinctive feature of the positional swaps index is that it incorporates information on the evaluation of alternatives based not on their nature but on their position in the ranking. Hence, NEU regains its appeal. At the same time, the incorporation of information on the ranking of the alternatives immediately implies the non-fulfilment of OC, since it completely disregards this type of information. As stated in the following proposition, the elimination of OC from the system of properties characterizing the swaps index characterizes the positional swaps index.

Proposition 2. *An inconsistency index I satisfies CONT, RAT, CONC, PWL, DC and NEU if and only if it is a positional swaps index.*

5.3. **General Weighted Index.** We now present a broad generalization of the swaps index that may incorporate the sort of information reflected by the non-neutral swaps index, the positional swaps index, and other kinds of information, such as priors on the plausibility of the different welfare rankings, etc. We call this index the general weighted index. The basic purpose of general weighted indices is to consider every possible inconsistency between an observation and a preference relation through a weight that may depend on the nature of the menu of alternatives, the nature of the chosen alternative, and the nature of the preference relation. Then, for a given collection of observations f, the inconsistency index takes the form of the minimum total inconsistency across all preference relations:

$$I_G(f) = \min_P \sum_{(A,a)} f(A,a)w(P,A,a)$$

where $w(P,A,a) = 0$ if $a = m(P,A)$ and $w(P,A,a) \in \mathbb{R}_{++}$ otherwise.

As argued in Remark 1, it turns out that general weighted indices are characterized by the first four axioms used in Theorem 4.

Proposition 3. *An inconsistency index I satisfies CONT, RAT, CONC and PWL if and only if it is a general weighted index.*

5.4. **Varian and Houtman-Maks.** Two popular measures of the consistency of behavior are due to Varian (1990) and Houtman and Maks (1985). In this section we bring these two measures into our framework, and show that they belong to the class of general weighted indices.

As advanced in section 3.2, Varian's inconsistency measure is the minimal sum of all the necessary wealth adjustments needed to remove any violations of consistency. In our setting this can be presented as

$$I_V(f) = \min_P \sum_{(A,a)} f(A,a) \max_{x \in A: xPa} w_x^A$$

where $w_x^A \in \mathbb{R}_{++}$ denotes the weight of alternative x in menu A.[19] We now argue why I_V is a proper representation of the original Varian index. First, the weights w_x^A play the role of the exogenous structure implied by the monetary system in Varian's original setting. Specifically, given a menu A, w_x^A captures the necessary adjustment required to remove option x from A. This is analogous to Varian's approach in the consumer setting, where the wealth adjustment required to eliminate one alternative depends on the budget set. Second, given a preference relation P, I_V by focusing on the maximum weight w_x^A across all the options in the upper contour set of (A,a), gives the maximum adjustment required to eliminate the inconsistent choice of a from the menu A. This parallels the original approach of Varian, where the required wealth adjustment is given by the alternative in the upper contour set that is farthest from the original budget line. Finally, I_V is defined on the basis of the preference relation minimizing the inconsistency value $\sum_{(A,a)} f(A,a) \max_{x \in A: xPa} w_x^A$. This captures the search for the minimal aggregated wealth adjustment required to make the remaining data consistent with the maximization of a preference relation.

We show below that I_V satisfies the first four properties of the swaps index. However, the characterization of I_V requires of further structure, related to the search for the maximum weight in a given upper contour set. To provide this further structure, let us first consider the following definition. Take any $r \in \mathcal{R}$ and any $(A,x) \in \mathcal{O}$, and denote by $\mathcal{R}^r_{(A,x)}$ all rationalizable collections of observations $r_{yz} \in \mathcal{R}$ such that the highest-ranking alternative in X according to $P^{r_{yz}}$ is $y \in A \setminus \{x\}$ with $r(\{x,y\},x) = 0$, and the second- highest-ranking alternative in X is $z \in A$. That is, the collection $\mathcal{R}^r_{(A,x)}$ encompasses all the rationalizable collections r_{yz} that assign the highest position in the ranking $P^{r_{yz}}$ to an alternative, y, which is in the upper contour set of (A,x) according to the preference relation P^r, while placing immediately below y an alternative z contained in A.

[19]For notational convenience, let $\max_{x \in \emptyset} w_x^A = 0$.

Varian's Consistency (VC). For every $(A,x) \in \mathcal{O}$ and every $r \in \mathcal{R}$, $I(\alpha \mathbf{1}_{(A,x)} + (1-\alpha)r) = \max_{r_{yz} \in \mathcal{R}^r_{(A,x)}} I(\alpha \mathbf{1}_{(A,z)} + (1-\alpha) r_{yz})$ for any sufficiently small $\alpha > 0$.[20]

Varian's Consistency imposes that the inconsistency that $\mathbf{1}_{(A,x)}$ generates in $\alpha\mathbf{1}_{(A,x)} + (1-\alpha)r$ can be related to collections of observations in which the upper contour set of the only inconsistency involves a single alternative y, that is ranked higher than x according to P^r. Varian's Consistency may then make sense under the interpretation that, for each alternative, we can think of the magnitude of the shock that must occur for this alternative to be neither mentally nor physically available. In this sense, the inconsistency of an observation might be related to the highest required magnitude of shock involving the alternatives in the upper contour set.

Varian's Consistency implies Ordinal Consistency. Ordinal Consistency establishes the equality between the inconsistency associated with the combination of an observation (A,x) with two sufficiently prevalent rationalizable collections r and r', whenever these treat all the alternatives in A in exactly the same way. Clearly, the latter requirement implies that the collections $\mathcal{R}^r_{(A,x)}$ and $\mathcal{R}^{r'}_{(A,x)}$ are exactly the same. Hence, Varian's Consistency implies Ordinal Consistency. However, it might be the case that the maximum inconsistency of (A,x) in relation to $\mathcal{R}^r_{(A,x)}$ coincides with the maximum inconsistency of (A,x) in relation to $\mathcal{R}^{r'}_{(A,x)}$, even if $\mathcal{R}^r_{(A,x)}$ and $\mathcal{R}^{r'}_{(A,x)}$ are not the same. It is in this sense that Varian's Consistency may lead to stronger conclusions than Ordinal Consistency.

The following result establishes the characterization of Varian's index I_V.

Proposition 4. *An inconsistency index I satisfies CONT, RAT, CONC, PWL and VC if and only if it is a Varian index.*

We now turn to the analysis of Houtman and Maks (1985). Recall from section 3.2 that Houtman and Maks' proposal has to do with the size of the minimal subset of observations that needs to be eliminated from the data, in order to make the remainder rationalizable. It follows immediately that, in our setting, the Houtman-Maks index, which we denote by I_{HM}, is but a special case of the Varian index when $w^A_x = 1$ for every A and every $x \in A$. Consequently, the characterization of I_{HM} must build on that of I_V, and impose some additional structure. First, notice that I_{HM} does not discriminate between the alternatives, and hence any relabeling of the alternatives

[20] Again, for notational convenience, let $\max_{r \in \emptyset} I(\cdot) = 0$.

should have no effect on the level of inconsistency. This implies that Neutrality regains its appeal. I_{HM}, however, requires further structure:

Houtman-Maks' Composition (HMC). For every $(A_1, x), (A_2, x) \in \mathcal{O}$ with $A_1 \cap A_2 = \{x\}$ and every $r \in \mathcal{R}$, $I(\alpha \mathbf{1}_{(A_1 \cup A_2, x)} + (1-\alpha)r) = \max\{I(\alpha \mathbf{1}_{(A_1,x)} + (1-\alpha)r), I(\alpha \mathbf{1}_{(A_2,x)} + (1-\alpha)r)\}$ for any sufficiently small $\alpha > 0$.

Houtman-Maks' Composition establishes that, under the same conditions of Disjoint Composition, two collections of observations can be merged into one, in which case the resulting collection of observations adopt the maximum inconsistency value of the two collections. We can now establish the characterization result of I_{HM}.

Proposition 5. *An inconsistency index I satisfies CONT, RAT, CONC, PWL, VC, HMC and NEU if and only if it is a scalar transformation of the Houtman-Maks index.*

5.5. **Non-General Weighted Indices.** The third approach described in the literature for measuring the rationality of a collection of revealed data involves counting the number of times a consistency property, say GARP, is violated (see, e.g., Swofford and Whitney, 1987; Famulari, 1995). As argued in section 3.2, the money pump index of Echenique, Lee and Shum (2011) weights every violation of GARP by the amount of money that could be extracted from the consumer. We now show that this class of measures do not belong to the class of general weighted indices.[21] The intuition is very simple. The class of general weighted indices assesses the inconsistency of the data from the perspective of the closest preference relation to it, whereas the above-mentioned indices measure the extent of *internal consistency* of the collection of observations. We illustrate this point in what follows by way of an example that compares two situations with the same number of cycles, but are treated differently by the general weighted indices. Similar examples can be constructed for the case in which the severity of each cycle is taken into account, as in Echenique, Lee and Shum (2011).[22]

Let $X = \{x, y, a_1, a_2, a_3, a_4\}$ and consider the following two scenarios, which report two rationalizable collections of observations that involve menus (i) in which alternatives x and y are never simultaneously present, (ii) which reveal that a_i is preferred to a_j whenever $i < j$, and (iii) that both x and y are better than any a_i. In addition,

[21] Afriat's index does not belong to the class of general weighted indices, either; the reason being that it replaces additivity with the maximum across all observations.

[22] These examples would involve the distortion of the relative frequency of the observations involving the cycles.

scenario 1 accounts for observations $(\{x, y, a_i\}, x)$, $i = 1, \ldots, 4$, and $(\{x, y, a_5\}, y)$, each of which occurs with a frequency ϵ, while scenario 2 additionally reports observations $(\{x, y, a_1\}, x), (\{x, y, a_2\}, x), (\{x, y, a_4\}, y)$ and $(\{x, y, a_5\}, y)$ each occurring with frequency ϵ. Clearly, the relative frequencies of cycles in the two scenarios are exactly the same.[23] In both cases it is understood that ϵ is relatively small, so that the rationalizable datasets are sufficiently prevalent in both scenarios.

We now show how any weighted general index necessarily discriminates between the two scenarios. Denote by f_1 and f_2 the collections of observations defined from scenarios 1 and 2, respectively. Given the prevalence of the rationalizable evidence in both scenarios, there are only two candidates to be considered as optimal preference relations: $xPyPa_1P \ldots Pa_5$ and $yP'xP'a_1P' \ldots P'a_5$. Therefore, the inconsistency of the first scenario is $I_G(f_1) = \epsilon \cdot \min\{w(P, \{x, y, a_5\}, y), \sum_{i=1}^{4} w(P', \{x, y, a_i\}, x)\}$ and that of the second is $I_G(f_2) = \epsilon \cdot \min\{w(P, \{x, y, a_4\}, y) + w(P, \{x, y, a_5\}, y), w(P', \{x, y, a_1\}, x) + w(P', \{x, y, a_2\}, x)\}$.

Let us proceed by contradiction and entertain the possibility of there being a general weighted index treating the two scenarios as equivalent, namely $I_G(f_1) = I_G(f_2)$. Clearly, if $I_G(f_1) = \epsilon \cdot \sum_{i=1}^{4} w(P', \{x, y, a_i\}, x)$, then $I_G(f_2) \leq \epsilon \cdot (w(P', \{x, y, a_1\}, x) + w(P', \{x, y, a_2\}, x)) < I_G(f_1)$, which is a contradiction. Then, it must be the case that $I_G(f_1) = \epsilon \cdot w(P, \{x, y, a_5\}, y)$. In a similar fashion, one can show that $I_G(f_2) = \epsilon \cdot (w(P', \{x, y, a_1\}, x) + w(P', \{x, y, a_2\}, x))$, which implies that $w(P, \{x, y, a_5\}, y) = w(P', \{x, y, a_1\}, x) + w(P', \{x, y, a_2\}, x)$. Now, since this argument can be reproduced for any permutation of the alternatives a_i, $i = 1, \ldots, 4$, it follows that $w(P, \{x, y, a_i\}, y) = 2w(P', \{x, y, a_i\}, x)$. Finally, note that we can also switch the roles of x and y concluding that $w(P', \{x, y, a_i\}, y) = 2w(P, \{x, y, a_i\}, x)$, and hence, $w(P, \{x, y, a_i\}, y) = 0$, which is a contradiction. This shows that every general weighted index discriminates between the two scenarios.

6. Related Literature

Although we have referred here and there to the related literature, it might be more useful to provide a brief summary all in one place. This should begin with the large empirical literature documenting deviations from the classical model of individual behavior. It is by now well-established that individual behavior is often dependent on the framing of the choice situation (Tversky and Kahneman, 1981), exhibits cyclic

[23]Therefore, one can equivalently consider the number of cycles.

choice patterns (May, 1954), is influenced by reference points (Thaler, 1980), and is susceptible to various sorts of menu manipulations (Iyengar and Lepper, 2000). Reacting to this evidence, later theoretical models adopt a revealed preference approach and expand the classical notion of rationality to incorporate, in various ways, stylized accounts of these behavioral phenomena. Some prominent recent examples are Bossert and Sprumont (2003, 2009), Masatlioglu and Ok (2005), Manzini and Mariotti (2007, 2012), Xu and Zhou (2007), Salant and Rubinstein (2008), Masatlioglu and Nakajima (2012), Green and Hojman (2009), Ok, Ortoleva and Riella (2012), and Masatlioglu, Nakajima and Ozbay (2012).

Continuing, in section 3.2 we reviewed the literature on revealed preference tests of the maximization principle, and studied the three approaches: Afriat (1973) and Varian (1990); Houtman and Maks (1985) and Dean and Martin (2012); and Swofford and Whitney (1987) and Echenique, Lee and Shum (2011). Halevy, Persitz and Zrill (2012) extend the approach of Afriat and Varian by complementing Varian's inconsistency index with an index measuring the misspecification with a set of utility functions. In addition, recent empirical applications of some of these measures have provided valuable information on the relationship between rationality and various demographics (see Choi, Kariv, Müller and Silverman, 2013; Dean and Martin, 2012; and Echenique, Lee and Shum, 2011). Finally, Beatty and Crawford (2011) axiomatically characterize a relative measure of the success of a revealed preference theory.

Lastly, there is a growing number of papers dealing with individual welfare analysis, when the individual's behavior is inconsistent. Bernheim and Rangel (2009) add to the standard choice data the notion of ancillary conditions, or frames. Ancillary conditions are assumed to be observable and potentially to affect individual choice, but they are irrelevant in terms of the welfare associated with the chosen alternative. Bernheim and Rangel suggest a welfare preference relation that ranks an alternative as welfare-superior to another only if the latter is never chosen when the former is available. Chambers and Hayashi (2012) characterize an extension of Bernheim and Rangel's model to probabilistic settings. Manzini and Mariotti (2009) offer a critical assessment of Bernheim and Rangel. Rubinstein and Salant (2012) 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. Masatlioglu, Nakajima and Ozbay (2012) suggest a welfare preference based on their limited-attention model of decision-making. Green and Hojman's (2009) proposal is to identify a list of conflicting selves, aggregate them to 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 measure 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 will explain a given set of choice data, and suggest this as a welfare measure.[24]

7. Conclusions

In this paper we propose a novel tool for the unified treatment of the measurement of rationality and welfare, namely, the swaps index. The swaps index identifies the closest preference relation to the revealed choices, i.e. the welfare ranking, and measures its associated inconsistency by enumerating the total number of available alternatives that rank above the chosen ones. The swaps index is unique in measuring rationality in terms of welfare considerations. In addition, it is the first tool in the literature with an axiomatic foundational analysis. With respect to welfare analysis, the swaps index evaluates the welfare ranking of any two alternatives by considering the whole collection of observations, and hence internalizes all the consequences of ranking one alternative above another. Moreover, it associates an error term, the inconsistency value, with the welfare ranking.

The swaps index relies exclusively on the endogenous information arising from the choice data. We offer generalizations which share the main features of the swaps index, while also being sensitive to various sorts of exogenous information that may be available to the analyst, such as information on the nature of the alternatives, on their cardinal utility values, etc. Our characterization of such indices provides the first axiomatic foundations for the measurement of rationality.

We would like to conclude by pointing to possible avenues of future research. On the technical side, it would be illuminating to extend our setting to include the possibility of indifferences between alternatives and the consideration of uncountable sets. On the empirical side, it would be important to investigate the practical differences between the swaps index and some of the indices we have discussed throughout the paper. On

[24]There are also papers describing methods for ranking objects such as teams or journals given a tournament matrix describing the information on the paired results of the objects (see Rubinstein, 1980; Palacios-Huerta and Volij, 2004).

the methodological side, the next natural step would involve the axiomatic development of rationality and welfare measures based on the revealed preference data, but also on some other relevant behavioral data.

Appendix A. Computational Considerations

Here, we deal with the question of identifying, in practice, the closest preference relation to the choice data, and its associated inconsistency level. Given that we have imposed no restriction whatsoever on the nature of the collections of observations, it is hardly surprising that the task of finding the optimal preference relation can sometimes prove computationally complex. Fortunately, we are able to show that it is possible to obtain good solutions by drawing upon existing techniques for addressing computational problems formally equivalent to ours.

Computational considerations are common in the application of the various inconsistency indices provided by the literature. Importantly, Dean and Martin (2012) establish that the problem studied by Houtman and Maks is equivalent to a well-known problem in the computer science literature, namely, the minimum set covering problem (MSCP). Smeulders, Cherchye, De Rock and Spieksma (2012) relate Varian and Houtman and Maks to the independent set problem (ISP). Then, one can draw from a wide range of algorithms developed by the operations research literature to solve these problems, for the purpose of computing the desired index in practice.

Exactly the same strategy can be adopted for the swaps index. Consider another well-known problem in the computer science literature, the linear ordering problem (LOP). The LOP has been related to a variety of problems, including various economic problems, particularly the triangularization of input-output matrices for the study of the hierarchical structures of the productive sectors in an economy.[25] Formally, the LOP problem over the set of vertices Y, and directed weighted edges connecting all vertices x and y in Y with cost c_{xy}, consists of finding the linear order over the set of vertices Y that minimizes the total aggregated cost. That is, if we denote by Π the set of all mappings from Y to $\{1, 2, \ldots, |Y|\}$, the LOP involves solving $\arg\min_{\pi \in \Pi} \sum_{\pi(x)<\pi(y)} c_{xy}$. As the following result shows, the LOP and the problem of computing the optimal preference relation for the swaps index are equivalent.

Proposition 6.

[25]See Korte and Oberhofer (1970) and Fukui (1986).

(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 swaps index.*

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

Intuitively, what are linear orders in the LOP are preference relations in our setting; while the cost of having one alternative rank above another is the inconsistency that arises from revealed data. Note that the evaluation of the inconsistency associated with having one alternative x that ranks above another alternative y is very simple. It is merely the sum of all values $f(A, y)$ across all menus A where x is present. Then, the computation of the optimal preference relation requires consideration of all the inconsistency values associated with having one alternative that ranks above another, exactly as in the LOP.

Proposition 6 enables the techniques offered by the literature for the solution of the LOP to be used directly in the computation of the optimal preference relation for the swaps index. These techniques involve an ample array of algorithms for finding the globally optimal solution.[26] Alternatively, the literature also offers methods, which, while not computing the globally optimal solution, are much lighter in computational intensity, and provide good approximations.[27]

Appendix B. Proofs

Proof of Theorem 1: Consider a collection of observations generated by a random utility model, $f_{\rho,u,\epsilon}$. Then, we can write $\sum_{(A,a)} f_{\rho,u,\epsilon}(A,a)|\{x \in A : xP^u a\}| = \sum_{A\in\mathcal{X}} \sum_{a\in A} f_{\rho,u,\epsilon}(A,a)|\{x \in A : xP^u a\}| = \sum_{A\in\mathcal{X}} \rho(A) \sum_{a\in A} Pr[a = \arg\max_{x\in A} U(x)]$ $|\{x \in A : xP^u a\}| = \sum_{A\in\mathcal{X}} \rho(A) \sum_{i=1}^{|A|} Pr[a_i^A = \arg\max_{x\in A} U(x)](i-1)$, where a_i^A represents the i-th best alternative in A according to P^u. Consider any other preference relation P'. It follows that $\sum_{(A,a)} f_{\rho,u,\epsilon}(A,a)|\{x \in A : xP'a\}| = \sum_{A\in\mathcal{X}} \rho(A) \sum_{i=1}^{|A|} Pr[a_i^A = \arg\max_{x\in A} U(x)])(\sigma^A(i) - 1)$, where σ^A is a permutation on $\{1, \ldots, |A|\}$ that transforms the rank of every alternative in A according to P^u in the rank of the same alternative according to P'.

[26]See, e.g., Grötschel, Jünger, and Reinelt (1984); see also Chaovalitwongse et al (2011) for a good introduction to the LOP, a review of the relevant algorithmic literature, and the analysis of one such algorithm.

[27]See Brusco, Kohn and Stahl (2008) for a good general introduction and relevant references.

We now show that the values $Pr[a_i^A = \arg\max_{x\in A} U(x)]$ are decreasing in i. To see this, consider two alternatives a_i^A and a_j^A with $i < j$. Given the menu A, consider a realization of the error terms such that U is maximized at a_j^A. That is, $u(a_j^A)+\epsilon(a_j^A) > u(a_i^A)+\epsilon(a_i^A)$ and $u(a_j^A)+\epsilon(a_j^A) > u(a_k^A)+\epsilon(a_k^A)$ for any other alternative $a_k^A \in A \setminus \{a_i^A, a_j^A\}$. Now, consider an alternative realization of the errors, where j receives the shock $\epsilon(a_i^A)$, i receives the shock $\epsilon(a_j^A)$ and k receives the same shock $\epsilon(a_k^A)$. Clearly, now the utility of a_i^A is $u(a_i^A) + \epsilon(a_j^A) > u(a_j^A) + \epsilon(a_j^A)$. Then, for all $a_k^A \in A\setminus\{a_i^A, a_j^A\}$, it follows that $u(a_i^A)+\epsilon(a_j^A) > u(a_k^A)+\epsilon(a_k^A)$. Also, since $u(a_i^A) > u(a_j^A)$ we have that $u(a_i^A) + \epsilon(a_j^A) > u(a_j^A) + \epsilon(a_i^A)$. Then, the i.i.d. nature of the errors guarantees that $Pr[a_i^A = \arg\max_{x\in A} U(x)] \geq Pr[a_j^A = \arg\max_{x\in A} U(x)]$, as desired. Therefore, for every menu $A \subseteq X$ we have $\rho(A)\sum_{i=1}^{|A|} Pr[a_i^A = \arg\max_{x\in A} U(x)](i-1) \leq \rho(A)\sum_{i=1}^{|A|} Pr[a_i^A = \arg\max_{x\in A} U(x)](\sigma^A(i)-1)$. This shows that P^u is an argument minimizing the swaps index.

To prove the second part of the theorem, notice first that, for any P' other than P^u there exists at least one pair of alternatives $x, y \in X$ such that $xP^u y$ and $yP'x$. Since $u(x) > u(y)$, and the mistakes are i.i.d, it follows that $Pr[U(x) > U(y)] > Pr[U(y) > U(x)]$. Since $\rho(\{x,y\}) > 0$, it follows from the conclusion of the previous paragraph that $\sum_{(A,a)} f(A,a)|x \in A : xP^u a| < \sum_{(A,a)} f(A,a)|x \in A : xP'a|$, as desired.■

Proof of Theorem 2: Consider a regular collection of observations $f_{\rho,u,\phi}$. Notice that from any preference relation P, we can reach P^u in a sequence of single u-improving consecutive flips, i.e., there exists a sequence of preference relations $P_0 = P, P_1, \ldots, P_m = P^u$ such that P_i and P_{i+1} differ only in ranking two consecutive alternatives and P_{i+1} ranks these two alternatives as P^u while P_i does not. To do so, we can for instance flip consecutively the top alternative in P^u from its initial position in P to the top, then proceed with the second best alternative in P^u, etc. We now prove, by using such a sequence, that the inconsistency associated to P_i is always larger than the inconsistency associated to P_{i+1}, which leads to the result.

Let $y, z \in X$ such that $yP_i z$ and $zP_{i+1}y$. Hence, $u(z) > u(y)$. Define $\omega_1 = \{(A,y) : \{y,z\} \subseteq A\}$, $\omega_2 = \{(A,z) : \{y,z\} \subseteq A\}$ and $\omega_3 = \mathcal{O} \setminus (\omega_1 \cup \omega_2)$. Clearly, $\sum_{(A,a)} f_{\rho,u,\phi}(A,a)|\{x \in A : xP_{i+1}a\}| \leq \sum_{(A,a)} f_{\rho,u,\phi}(A,a)|\{x \in A : xP_i a\}|$ if and only if $\sum_{j=1}^{3}\sum_{(A,a)\in\omega_j} f_{\rho,u,\phi}(A,a)|\{x \in A : xP_{i+1}a\}| \leq \sum_{j=1}^{3}\sum_{(A,a)\in\omega_j} f_{\rho,u,\phi}(A,a)|\{x \in A : xP_i a\}|$.

Notice that for any $(A,a) \in \omega_3$, it must be $|\{x \in A : xP_i a\}| = |\{x \in A : xP_{i+1}a\}|$. This is because either y or z are not in A, or, whenever both are in A, the selected alternative a is below or above both alternatives at the same time. Hence, the former inequality can be expressed as: $\sum_{j=1}^{2}\sum_{(A,a)\in\omega_j} f_{\rho,u,\phi}(A,a)|\{x\in A : xP_{i+1}a\}| \leq \sum_{j=1}^{2}\sum_{(A,a)\in\omega_j} f_{\rho,u,\phi}(A,a)|\{x\in A : xP_i a\}|$.

For any observation $(A,a)\in\omega_1$, we have $|\{x\in A : xP_{i+1}y\}| = |\{x\in A : xP_i y\}| + 1$ and hence $f_{\rho,u,\phi}(A,y)\,(|\{x\in A : xP_{i+1}y\}| - |\{x\in A : xP_i y\}|) = f_{\rho,u,\phi}(A,y)$. Similarly, for any observation $(A,a)\in\omega_2$, $|\{x\in A : xP_{i+1}z\}| = |\{x\in A : xP_i z\}| - 1$ and hence, $f_{\rho,u,\phi}(A,z)(|\{x\in A : xP_i z\}| - |\{x\in A : xP_{i+1}z\}|) = f_{\rho,u,\phi}(A,z)$. Hence, we can write the previous inequality as: $\sum_{(A,a)\in\omega_1} f_{\rho,u,\phi}(A,a) \leq \sum_{(A,a)\in\omega_2} f_{\rho,u,\phi}(A,a)$.

Now notice that whenever $(A,y)\in\omega_1$, the frequency $f_{\rho,u,\phi}(A,y)$ is either 0 or $\rho(A^y)$, since y is not maximal in A according to u. By the regularity assumption, $\rho(A^z)\geq\rho(A^y)$. Now, whenever the frequency $f_{\rho,u,\phi}(A,y)$ is equal to $\rho(A^y)$ then the frequency $f_{\rho,u,\phi}(A,z)$ must be at least $\rho(A^z)$, proving that for such sets, $f_{\rho,u,\phi}(A,y)\leq f_{\rho,u,\phi}(A,z)$. Hence, $\sum_{(A,a)\in\omega_1} f_{\rho,u,\phi}(A,a) \leq \sum_{(A,a)\in\omega_2} f_{\rho,u,\phi}(A,a)$ as desired. We have proved that P_{i+1} gives lower swaps than P_i, and by transitivity P^u has minimal swaps.

Finally, notice that $\rho(\{y,z\}^\diamond) > 0$ implies $f_{\rho,u,\phi}(\{y,z\},y) < f_{\rho,u,\phi}(\{y,z\},z)$ and hence the previous inequality must be strict. This makes P^u to be uniquely identified.■

Proof of Theorem 3: Let u be a utility function over monetary outcomes, and consider β,δ and the associated collection of observations $f_{\rho,U_{\beta,\delta}}$. Notice that whenever $\beta = 1$, the choices correspond completely to the preference P^{V_δ} and there are no mistakes. From now on, we therefore assume $\beta < 1$.

We first characterize the types of mistakes that may arise when using a stationary preference P^{V_α}. Consider an observation $(A,(o,t))$ where (o,t) is not the maximizer of P^{V_α} in A and take an alternative (o',t') in A such that $(o',t')P^{V_\alpha}(o,t)$. We show that $t\neq t'$. To see this, simply notice that whenever $t=t'$, the choice of (o,t) leads to $u(o) > u(o')$ and the fact that $(o',t')P^{V_\alpha}(o,t)$ leads to $u(o) < u(o')$, a contradiction. We divide all the possible mistakes in the following four categories.

(1) $t' > t > 0$: Since the individual has chosen (o,t) over (o',t') it must be $U_{\beta,\delta}(o,t) > U_{\beta,\delta}(o',t')$ or equivalently, $\beta\delta^t u(o) > \beta\delta^{t'}u(o')$. Since $(o',t')P^{V_\alpha}(o,t)$, we have $\alpha^{t'}u(o') > \alpha^t u(o)$. We can express the two inequalities as $\delta^{t'-t} < \frac{u(o)}{u(o')} < \alpha^{t'-t}$. Notice that this case can arise if and only if $\alpha > \delta$.

(2) $t > t' > 0$: From the choice we have $\beta\delta^t u(o) > \beta\delta^{t'}u(o')$, while from the preference we have $\alpha^{t'}u(o') > \alpha^t u(o)$. We can express these two inequalities as: $\alpha^{t-t'} < \frac{u(o')}{u(o)} < \delta^{t-t'}$. Notice that this case occurs if and only if $\alpha < \delta$.

(3) $t > t' = 0$: This implies both $\beta\delta^t u(o) > u(o')$ and $u(o') > \alpha^t u(o)$, or equivalently: $\alpha^t < \frac{u(o')}{u(o)} < \beta\delta^t$. Notice that this case can only occur whenever $\alpha^t < \beta\delta^t$. For $\alpha < \beta\delta$, this condition holds for every t. For $\alpha \in [\beta\delta, \delta)$, this will be true only for sufficiently large integers, and denote by ψ the first integer that meets such condition.[28] For $\alpha \geq \delta$ the condition is never satisfied.

(4) $t' > t = 0$: This implies both $u(o) > \beta\delta^{t'} u(o')$ and $\alpha^{t'} u(o') > u(o)$, or equivalently: $\beta\delta^{t'} < \frac{u(o)}{u(o')} < \alpha^{t'}$, and notice that this case can only occur whenever $\beta\delta^{t'} < \alpha^{t'}$. For $\alpha \leq \beta\delta$, this condition never holds. For $\alpha \in (\beta\delta, \delta)$, this will be true only for integers sufficiently below ψ. Finally, for $\alpha \geq \delta$ the condition always holds.

We prove that for every $\alpha < \beta\delta$, the stationary preference $P^{V_{\beta\delta}}$ has less swaps than P^{V_α}. Notice that only mistakes of type-(2) and type-(3) apply to $P^{V_{\beta\delta}}$. Since both $\alpha^{t-t'} < (\beta\delta)^{t-t'}$ and $\alpha^t < (\beta\delta)^t$, any type-(2) and type-(3) mistake for $P^{V_{\beta\delta}}$ is also a mistake for P^{V_α}. We now prove that for every stationary preference P^{V_α} with $\alpha > \delta$, the stationary preference P^{V_δ} has less swaps. Notice that only mistakes of type-(4) apply to P^{V_δ}. Since $\delta^t < \alpha^t$, any type-(4) mistake for P^{V_δ} is also a mistake for P^{V_α}. This proves the first part of the theorem.

For the second part of the theorem, consider a time-regular collection $f_{\rho,U_{\beta\delta}}$. To provide a computation of the bound that gives a sense of the requirements in the theorem, we consider binary menus. In the general case, the threshold would depart from this bound. Given the previous arguments, we can assume that $\alpha \in [\beta\delta, \delta]$. Notice that type-(1) mistakes cannot happen for such values of α and hence, we focus on the other three types.

We first consider the mass of mistakes of type-(2), $S_2(\alpha)$. Fix a pair of times t and t'. The probability of a binary menu having the alternatives dated in such times is $\frac{1}{(T+1)^2}$. Given t and t', the probability of having outcomes satisfying the inequality $\alpha^{t-t'} < \frac{u(o')}{u(o)} < \delta^{t-t'}$ is $\frac{\delta^{t-t'}-\alpha^{t-t'}}{2}$. This follows from the assumption on the utilities being uniformly distributed, since in this case the probability of the ratio being below 1 is $1/2$, and conditional on this the distribution is also uniform. Hence, the mass of type-(2) mistakes for P^{V_α} is $S_2(\alpha) = \frac{1}{2(T+1)^2}\sum_{t'=1}^{T-1}\sum_{t=t'+1}^{T}(\delta^{(t-t')} - \alpha^{(t-t')}) = \frac{1}{2(T+1)^2}\sum_{i=1}^{T-1}(T-i)(\delta^i - \alpha^i)$.

We now consider the mass of mistakes of type-(3), $S_3(\alpha)$. Given t and t', the probability of having outcomes that satisfy $\alpha^t < \frac{u(o')}{u(o)} < \beta\delta^t$ is $\frac{\beta\delta^t-\alpha^t}{2}$ whenever $t \geq \psi$ (if

[28]This integer obviously depends on the parameters α, β, and δ, but for notational convenience we simply write ψ.

$\psi \leq T$), and 0 otherwise. Consequently: $S_3(\alpha) = \frac{1}{2(T+1)^2}\sum_{i=\psi}^{T}(\beta\delta^i - \alpha^i)$, provided that $\psi \leq T$, and zero otherwise.

Finally, we consider the mass of type-(4) mistakes, $S_4(\alpha)$. Given t and t', the probability of having outcomes that satisfy $\beta\delta^{t'} < \frac{u(o)}{u(o')} < \alpha^{t'}$ is $\frac{\alpha^{t'}-\beta\delta^{t'}}{2}$ whenever $t < \psi$ (whenever ψ does not exist, i.e., $\alpha = \delta$, for notational purposes just take $\psi = T+1$), and 0 otherwise. Then: $S_4(\alpha) = \frac{1}{2(T+1)^2}\sum_{i=1}^{\psi-1}(\alpha^i - \beta\delta^i)$.

Consider the total mistakes $S(\alpha) = S_2(\alpha) + S_3(\alpha) + S_4(\alpha)$. We study the derivative of $S(\alpha)$ with respect to α, within the interior of the set of α-values that share the same integer ψ. We have:

$$\frac{\partial S(\alpha)}{\partial \alpha} = \frac{1}{2(T+1)^2}(-\sum_{i=1}^{T-1} i(T-i)\alpha^{i-1} - \sum_{i=\psi}^{T} i\alpha^{i-1} + \sum_{i=1}^{\psi-1} i\alpha^{i-1}).$$

Since the second term is always negative and the third is always positive, this derivative is bounded above by the case in which $\psi > T$, and hence:

$$\frac{\partial S(\alpha)}{\partial \alpha} \leq \frac{1}{2(T+1)^2}(-\sum_{i=1}^{T-1} i(T-i)\alpha^{i-1} + \sum_{i=1}^{T} i\alpha^{i-1}) = \frac{1}{2(T+1)^2}(\sum_{i=1}^{T} i[1-(T-i)]\alpha^{i-1}).$$

Now notice that the coefficients associated to α^{i-1} are increasing in i and they are all negative except for the case $i = T$. Hence, the derivative is bounded above by $\frac{1}{2(T+1)^2}\alpha^{T-1}(\sum_{i=1}^{T} i[1-(T-i)])$, which is simply $\alpha^{T-1}\frac{T(4-T)}{12(T+1)}$. Clearly, for any value of T greater or equal than 4, we have $\frac{\partial S(\alpha)}{\partial \alpha} \leq 0$. This implies that mistakes always decrease with α and hence, the stationary preference relation that minimizes the number of swaps is P^{V_δ}, as desired.■

Proof of Theorem 4: That a swaps index satisfies the axioms is immediate. We prove sufficiency of the axioms in 10 steps.

Step 1. PWL guarantees that there is a rationally founded cover $\mathcal{C} = \{\mathcal{C}_j\}$ of $\mathcal{F}$ for which the index is piecewise linear. We show that the index must also be piecewise linear over the cover $\bar{\mathcal{C}} = \{\bar{\mathcal{C}}_j\}$, where each component $\bar{\mathcal{C}}_j$ is the convex hull of the closure of $\mathcal{C}_j$. The application of CONT and linearity over each component $\mathcal{C}_j$ guarantees linearity over its closure. Since each element in the convex hull can be obtained as a linear combination of elements in the closure of $\mathcal{C}_j$, the repeated application of linearity shows that the index is also linear over the convex hull of the closure of $\mathcal{C}_j$.

Step 2. Here, we prove that $\bar{\mathcal{C}}$ is also rationally founded. Since for every j, $\mathcal{C}_j \subseteq \bar{\mathcal{C}}_j$, it follows that there is at least one $r_j \in \bar{\mathcal{C}}_j$. We now prove that for every $l \neq j$, $r_l \notin \bar{\mathcal{C}}_j$. Assume, by contradiction, that $r_l \in \bar{\mathcal{C}}_j$. Then, PWL and RAT guarantee that, for every

$\alpha \in (0,1)$, it must be that $I(\alpha r_j + (1-\alpha)r_l) = \alpha I(r_j) + (1-\alpha)I(r_l) = 0$. However, since $r_j \neq r_l$ and they put positive mass on every menu, there must exist at least one menu B such that the respectively unique chosen alternatives in r_j and r_l differ. Then, the collection of observations $\alpha r_j + (1-\alpha)r_l$ puts positive mass on two different observations from at least one menu B, and hence, it is not rationalizable. Then RAT implies $I(\alpha r_j + (1-\alpha)r_l) \neq 0$, which is a contradiction.

Step 3. We now prove that, for each r_j, there exists $\alpha^{r_j} \in (0,1]$ such that, for every observation (A,a), and for every $\alpha \in [0,\alpha^{r_j}]$, it is the case that $\alpha \mathbf{1}_{(A,a)} + (1-\alpha)r_j \in \bar{\mathcal{C}}_j$. Proceeding by contradiction, let us assume that there exists r_j, (A,a) and a sequence of real values $\{\alpha_n\}$, with $\alpha_n \to 0$, such that for each n, $\alpha_n \mathbf{1}_{(A,a)} + (1-\alpha_n)r_j$ does not belong to $\bar{\mathcal{C}}_j$. Hence, there exists a subsequence of such real numbers $\{\alpha_{n_i}\}$ with $\alpha_{n_i} \to 0$, for which all collections $\alpha_{n_i}\mathbf{1}_{(A,a)} + (1-\alpha_{n_i})r_j$ belong to a different common class in the cover, say $\bar{\mathcal{C}}_l \neq \bar{\mathcal{C}}_j$. Since $\bar{\mathcal{C}}_l$ is closed, the limit of such a subsequence of collections, r_j, also belongs to $\bar{\mathcal{C}}_l$, thus contradicting step 2.

Step 4. We now associate, to every preference P and every observation (A,a), a real-valued weight $w(P,A,a)$. Clearly, there exists a bijection between $\mathcal{P}$ and the set of rationalizable collections of observations putting equal mass on each menu $\mathcal{R}$. We then denote by P_j the preference that generates $r_j \in \bar{\mathcal{C}}_j$. Hence, for each P_j, step 3 guarantees that there exists $\alpha^{r_j} \in (0,1]$ such that for every observation (A,a) and for every $\alpha \in [0,\alpha^{r_j}]$, it is the case that $\alpha \mathbf{1}_{(A,a)} + (1-\alpha)r_j \in \bar{\mathcal{C}}_j$. Define, for every preference P_j and observation (A,a), the weight

$$w(P_j,A,a) = \frac{I(\alpha^{r_j}\mathbf{1}_{(A,a)} + (1-\alpha^{r_j})r_j)}{\alpha^{r_j}}.$$

Step 5. Since $\bar{\mathcal{C}}$ is a cover of $\mathcal{F}$, for every f there exists at least one component $\bar{\mathcal{C}}_j$ such that $f \in \bar{\mathcal{C}}_j$. In this step, we prove that $I(f) = \sum_{(A,a)} f(A,a)w(P_j,A,a)$. By PWL and RAT,

$$I(f) = \frac{\alpha^{r_j}I(f)}{\alpha^{r_j}} = \frac{\alpha^{r_j}I(f) + (1-\alpha^{r_j})I(r_j)}{\alpha^{r_j}} = \frac{I(\alpha^{r_j}f + (1-\alpha^{r_j})r_j)}{\alpha^{r_j}}.$$

Now, notice that $\alpha^{r_j}f + (1-\alpha^{r_j})r_j = \alpha^{r_j}(\sum_{(A,a)} f(A,a)\mathbf{1}_{(A,a)}) + (1-\alpha^{r_j})r_j = \sum_{(A,a)} f(A,a)(\alpha^{r_j}\mathbf{1}_{(A,a)}) + (1-\alpha^{r_j})r_j = \sum_{(A,a)} f(A,a)[\alpha^{r_j}\mathbf{1}_{(A,a)} + (1-\alpha^{r_j})r_j]$. By construction of α^{r_j}, all collections $\alpha^{r_j}\mathbf{1}_{(A,a)} + (1-\alpha^{r_j})r_j$ belong to $\bar{\mathcal{C}}_j$ and by convexity of $\bar{\mathcal{C}}_j$, all convex combinations of such collections must also lie in $\bar{\mathcal{C}}_j$. We can thus apply linearity repeatedly within $\bar{\mathcal{C}}_j$ to obtain

$$I(f) = \frac{I(\alpha^{r_j} f + (1-\alpha^{r_j})r_j)}{\alpha^{r_j}} = \frac{\sum_{(A,a)} f(A,a) I(\alpha^{r_j}\mathbf{1}_{(A,a)} + (1-\alpha^{r_j})r_j)}{\alpha^{r_j}} = \sum_{(A,a)} f(A,a) w(P_j, A, a).$$

Step 6. Here, we prove that, for every $f \in \mathcal{F}$, $I(f) = \min_P \sum_{(A,a)} f(A,a)w(P,A,a)$ with $w(P_t, A, a) = 0$ if and only if $aP_t x$ for all $x \in A \setminus \{a\}$. To see the first part, and given step 5, we only need to prove that $I(f) \leq \sum_{(A,a)} f(A,a)w(P_t,A,a)$ for every preference P_t. By CONC and RAT,

$$I(f) = \frac{\alpha^{r_t} I(f)}{\alpha^{r_t}} = \frac{\alpha^{r_t} I(f) + (1-\alpha^{r_t})I(r_t)}{\alpha^{r_t}} \leq \frac{I(\alpha^{r_t} f + (1-\alpha^{r_t})r_t)}{\alpha^{r_t}}.$$

We know that for every observation $\mathbf{1}_{(A,a)}$, the collection $\alpha^{r_t}\mathbf{1}_{(A,a)} + (1-\alpha^{r_t})r_t$ belongs to $\bar{\mathcal{C}}_t$. Convexity of this class guarantees that $\alpha^{r_t} f + (1-\alpha^{r_t})r_t$ also belongs to $\bar{\mathcal{C}}_t$. By step 5, we also know that $I(\alpha^{r_t} f + (1-\alpha^{r_t})r_t) = \alpha^{r_t} \sum_{(A,a)} f(A,a)w(P_t,A,a)$. Hence, $I(f) \leq \sum_{(A,a)} f(A,a)w(P_t,A,a)$.

We conclude by showing that $w(P_t, A, a) = 0$ if and only if $aP_t x$ for all $x \in A \setminus \{a\}$. Notice that whenever $aP_t x$ for all $x \in A \setminus \{a\}$, it is the case that $r_t(A,a) > 0$ and the collection $\alpha^{r_t}\mathbf{1}_{(A,a)} + (1-\alpha^{r_t})r_t$ is also rationalizable by P_t. Hence, by RAT, $w(P_t, A, a) = I(\alpha^{r_t}\mathbf{1}_{(A,a)} + (1-\alpha^{r_t})r_t) = 0$. Whenever $aP_t x$ does not hold for all $x \in A \setminus \{a\}$, we have $r_t(A,x) > 0$ for some $x \neq a$, and the collection $\alpha^{r_t}\mathbf{1}_{(A,a)} + (1-\alpha^{r_t})r_t$ cannot be rationalized by any preference. Hence, by RAT, $w(P_t, A, a) = \frac{I(\alpha^{r_t}\mathbf{1}_{(A,a)} + (1-\alpha^{r_t})r_t)}{\alpha^{r_t}} > 0$.

Step 7. We now prove that for every observation (A, x), and for every pair of preferences P_j, P_l such that $yP_j z \Leftrightarrow yP_l z$ for every $y, z \in A$, it is the case that $w(P_j, A, x) = w(P_l, A, x)$. Consider the rationalizable collections $r_j, r_l \in \mathcal{R}$ associated with preferences P_j, P_l. By OC there is a sufficiently small $\alpha > 0$ such that $I(\alpha\mathbf{1}_{(A,x)} + (1-\alpha)r_j) = I(\alpha\mathbf{1}_{(A,x)} + (1-\alpha)r_l)$. Let $\hat{\alpha} = \min\{\alpha, \alpha^{r_j}, \alpha^{r_l}\}$. The previous equality holds for $\hat{\alpha}$, and also $\hat{\alpha}\mathbf{1}_{(A,x)} + (1-\hat{\alpha})r_j \in \bar{\mathcal{C}}_j$ and $\hat{\alpha}\mathbf{1}_{(A,x)} + (1-\hat{\alpha})r_l \in \bar{\mathcal{C}}_l$. Linearity together with step 5 guarantees that $w(P_j, A, x) = \frac{I(\alpha^{r_j}\mathbf{1}_{(A,x)} + (1-\alpha^{r_j})r_j)}{\alpha^{r_j}} = \frac{I(\hat{\alpha}\mathbf{1}_{(A,x)} + (1-\hat{\alpha})r_j)}{\hat{\alpha}} = \frac{I(\hat{\alpha}\mathbf{1}_{(A,x)} + (1-\hat{\alpha})r_l)}{\hat{\alpha}} = \frac{I(\alpha^{r_l}\mathbf{1}_{(A,x)} + (1-\alpha^{r_l})r_l)}{\alpha^{r_l}} = w(P_l, A, x)$, as desired.

Step 8. Here we prove that for every (A, x) and P_j, $w(P_j, A, x) = \sum_{y \in A} w(P_j, \{x, y\}, x)$. To do so, we prove that, for any two menus, A_1, A_2, such that $A_1 \cap A_2 = \{x\}$ and $A_1 \cup A_2 = A$, it is the case that $w(P_j, A, x) = w(P_j, A_1, x) + w(P_j, A_2, x)$. The recursive application of this idea, given the finiteness of X, concludes the step. To see this equality, consider the r_j associated with P_j. By DC, there exists a sufficiently small α such that $I(\alpha\mathbf{1}_{(A,x)} + (1-\alpha)r_j) = I(\alpha\mathbf{1}_{(A_1,x)} + \alpha\mathbf{1}_{(A_2,x)} + (1-2\alpha)r_j)$. Let $\alpha^* = \min\{\alpha, \frac{\alpha^{r_j}}{2}\}$. The previous equality holds for α^*, and also the collections $\alpha^*\mathbf{1}_{(A,x)} + (1-\alpha^*)r_j$ and $\alpha^*\mathbf{1}_{(A_1,x)} + \alpha^*\mathbf{1}_{(A_2,x)} + (1-2\alpha^*)r_j$ both belong to $\bar{\mathcal{C}}_j$. Linearity together with step 5 guarantees that $w(P_j, A, x) = \frac{I(\alpha^{r_j}\mathbf{1}_{(A,x)}+(1-\alpha^{r_j})r_j)}{\alpha^{r_j}} = \frac{I(\alpha^*\mathbf{1}_{(A,x)}+(1-\alpha^*)r_j)}{\alpha^*} = \frac{I(\alpha^*\mathbf{1}_{(A_1,x)}+\alpha^*\mathbf{1}_{(A_2,x)}+(1-2\alpha^*)r_j)}{\alpha^*} = \frac{I(\alpha^{r_j}\mathbf{1}_{(A_1,x)}+(1-\alpha^{r_j})r_j)}{\alpha^{r_j}} + \frac{I(\alpha^{r_j}\mathbf{1}_{(A_2,x)}+(1-\alpha^{r_j})r_j)}{\alpha^{r_j}} = w(P_j, A_1, x) + w(P_j, A_2, x)$.

Step 9. Here we prove that for every $x, y, z, t \in X$ and P_j, P_l such that the ranking of x (respectively, of y) in P_j is the same as the ranking of z (respectively, of t) in P_l, $w(P_j, \{x,y\}, y) = w(P_l, \{z,t\}, t)$. Consider the bijection $\sigma : X \to X$ given by the ranking of the alternatives in the preferences. That is, to the alternative ranked at s in P_j, we assign the alternative ranked at s in P_l. Notice, in particular, that we have $\sigma(x) = z$ and $\sigma(y) = t$ and also, $\sigma(r_j) = r_l$. Hence, we also have $\sigma(\alpha\mathbf{1}_{(\{x,y\},y)} + (1-\alpha)r_j) = \alpha\mathbf{1}_{(\{z,t\},t)} + (1-\alpha)r_l$ for every $\alpha \in [0,1]$. Let $\breve{\alpha} = \min\{\alpha^{r_j}, \alpha^{r_l}\}$. Then we have $\breve{\alpha}\mathbf{1}_{(\{x,y\},y)} + (1-\breve{\alpha})r_j \in \bar{\mathcal{C}}_j$ and $\breve{\alpha}\mathbf{1}_{(\{z,t\},t)} + (1-\breve{\alpha})r_l \in \bar{\mathcal{C}}_l$. By NEU, $I(\breve{\alpha}\mathbf{1}_{(\{x,y\},y)} + (1-\breve{\alpha})r_j) = I(\breve{\alpha}\mathbf{1}_{(\{z,t\},t)} + (1-\breve{\alpha})r_l)$. By using the decomposition obtained in step 5, this is equivalent to $\breve{\alpha}w(P_j, \{x,y\}, y) = \breve{\alpha}w(P_l, \{z,t\}, t)$, as desired.

Step 10. We end up by proving that I is a positive scalar transformation of the swaps index. By step 8, and the fact proved in step 6 that $w(P, \{x,a\}, a) = 0$ whenever aPx, we have that $\sum_{(A,a)} f(A,a)w(P,A,a) = \sum_{(A,a)} f(A,a)(\sum_{x\in A:xPa} w(P,\{x,a\},a))$. By step 7, we have that $\sum_{(A,a)} f(A,a)w(P,A,a) = \sum_{(A,a)} f(A,a)|\{x \in A : xPa\}|K(P)$ where $K(P)$ is a strictly positive real number possibly dependent on P. Finally, by step 9, this value $K(P)$ is constant across different preferences, and hence we have shown that $\sum_{(A,a)} f(A,a)w(P,A,a) = K\sum_{(A,a)} f(A,a)|\{x \in A : xPa\}|$. Therefore, by step 6 we have $I(f) = \min_P K\sum_{(A,a)} f(A,a)|\{x \in A : xPa\}|$, which shows that I is a positive scalar transformation of the swaps index.■

Proof of Proposition 1: It is easy to see that any non-neutral swaps index satisfies the axioms. Notice that we can prove the converse statement by using steps 1 to 8 in the proof of Theorem 4. By step 8 and the fact that the weight $w(P,\{x,a\},a) = 0$ whenever aPx, we have that $\sum_{(A,a)} f(A,a)w(P,A,a) = \sum_{(A,a)} f(A,a)\sum_{x\in A:xPa} w(P,\{x,a\},a)$. By step 7, the weight $w(P,\{x,a\},a)$ is independent of P, whenever xPa, and hence we can write $\sum_{(A,a)} f(A,a)\sum_{x\in A:xPa} w_{x,a}$. This, together with step 6, proves that the index is a non-neutral swaps index. ■

Proof of Proposition 2: It is easy to see that any positional swaps index satisfies the axioms. Notice that we can prove the converse statement by using steps 1 to 6 and steps 8 and 9 in the proof of Theorem 4. By step 8 and the fact that the weight $w(P, \{x, a\}, a) = 0$ whenever aPx, we have that $\sum_{(A,a)} f(A, a)w(P, A, a) = \sum_{(A,a)} f(A, a) \sum_{x\in A:xPa} w(P, \{x, a\}, a)$. Now, by step 9, $w(P, \{x, a\}, a)$ only depends on the rank of alternatives x and a in P. This, together with step 6, shows that the index is a positional swaps index. ■

Proof of Proposition 4: It is easy to see that any Varian index satisfies the axioms. Notice that we can use steps 1 to 7 in the proof of Theorem 4 to prove the converse statement. We now define, for every menu A and every alternative $y \in A$, the weight w_y^A. To do so, select any alternative $z \in A\setminus\{y\}$, and set $w_y^A(z) = \frac{I(\alpha^{r_{yz}}\mathbf{1}_{(A,z)}+(1-\alpha^{r_{yz}})r_{yz})}{\alpha^{r_{yz}}}$, where $\alpha^{r_{yz}}$ is defined as in step 4 of Theorem 4. We now show that this value does not depend on the selected z. Consider then another $z' \in A \setminus \{y\}$. Since $r_{yz}, r_{yz'} \in R^{r_{yz'}}_{(A,z')} = R^{r_{yz}}_{(A,z)}$, VC and PWL guarantee that $\alpha w_y^A(z) = I(\alpha\mathbf{1}_{(A,z)} + (1-\alpha)r_{yz}) = I(\alpha\mathbf{1}_{(A,z')} + (1-\alpha)r_{yz'}) = \alpha w_y^A(z')$ for any sufficiently small α.

Now, consider any preference P, its associated rationalizable collection $r \in \mathcal{R}$ and any observation (A, x). If $r(A, x) > 0$, then, by step 6, $w(P, A, x) = 0 = \max_{x\in\emptyset} w_x^A$. Assume then that $r(A, x) = 0$ and hence $\mathcal{R}^r_{(A,x)}$ is non-empty. By VC, there exists $\bar{\alpha}$ such that $I(\bar{\alpha}\mathbf{1}_{(A,x)} + (1-\bar{\alpha})r) = \max_{r_{yz}\in\mathcal{R}^r_{(A,x)}} I(\bar{\alpha}\mathbf{1}_{(A,z)} + (1-\bar{\alpha})r_{yz})$. Consider $\hat{\alpha} = \min_{r_{yz}\in\mathcal{R}^r_{(A,x)}} \alpha^{r_{yz}}$. Take $\alpha = \min\{\alpha^r, \bar{\alpha}, \hat{\alpha}\}$, where α^r is again defined as in Theorem 4. By step 6 in Theorem 4, VC, and the above construction of weights, we have that $w(P, A, x) = \frac{I(\alpha\mathbf{1}_{(A,x)}+(1-\alpha)r)}{\alpha} = \frac{\max_{r_{yz}\in\mathcal{R}^r_{(A,x)}} I(\alpha\mathbf{1}_{(A,z)}+(1-\alpha)r_{yz})}{\alpha} = \max_{x\in A:xPa} w_x^A$, as desired.■

Proof of Proposition 5: It is easy to see that the Houtman-Maks index satisfies the axioms. Notice that we can prove the converse statement by starting from the result of Proposition 4. We now prove that, for any menu A and any $x \in A$, we have $w_x^A = w_x^{\{x,y\}}$ for any $y \in A \setminus \{x\}$. This is trivial if $|A| = 2$. If $|A| > 2$, let $A_1 = \{x, y\}$ and $A_2 = A\setminus\{y\}$. Clearly, $A_1 \cap A_2 = \{x\}$. By HMC and RAT, for a sufficiently small α, it must be the case that $w_x^A = \frac{I(\alpha\mathbf{1}_{(A,x)}+(1-\alpha)r_{yx})}{\alpha} = \max\{\frac{I(\alpha\mathbf{1}_{(A_1,x)}+(1-\alpha)r_{yx})}{\alpha}, \frac{I(\alpha\mathbf{1}_{(A_2,x)}+(1-\alpha)r_{yx})}{\alpha}\} = \frac{I(\alpha\mathbf{1}_{(A_1,x)}+(1-\alpha)r_{yx})}{\alpha} = w_x^{\{x,y\}}$. A direct application of NEU guarantees that $w_x^{\{x,y\}} = w_z^{\{z,t\}}$ for every $x, y, z, t \in X$. Hence, the index is a scalar transformation of the Houtman-Maks index.■

Proof of Proposition 6: For the first part, consider the collection of observations f and define, for every pair of alternatives x and y in X, the weight $c_{xy} = \sum_{(A,y):x\in A} f(A,y)$. It follows that $\sum_{\pi(x)<\pi(y)} c_{xy} = \sum_{\pi(x)<\pi(y)} \sum_{(A,y):x\in A} f(A,y) = \sum_{(A,y)} f(A,y)|\{x \in A : \pi(x) < \pi(y)\}|$, and hence, by solving the LOP, we obtain the optimal preference for the swaps index. To see the second part, consider the LOP given by weights c_{xy}, with $x, y \in X$. Let c be the sum of all weights c_{xy}. Define the collection of observations f given by $f(\{x,y\},y) = \frac{c_{xy}}{c}$ and 0 otherwise. Since f is defined only 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.■

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] Barberis, N and W. Xiong (2009), "What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation," *Journal of Finance*, 64:751–784.

[5] Beatty, T.K.M and I.A. Crawford (2011), "How Demanding Is the Revealed Preference Approach to Demand?," *American Economic Review*, 101:2782–2795.

[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] Blow, L. and I. Crawford (2013), "Nonparametric Analysis of Time-Inconsistent Preferences: The Case of Hyperbolic Discounting," mimeo.

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

[9] Bossert, W. and Y. Sprumont (2009), "Non-Deteriorating Choice," *Economica*, 76:337–363.

[10] 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.

[11] Chambers, C.P. and T. Hayashi (2012), "Choice and Individual Welfare," *Journal of Economic Theory* 147(5):1818–1849.

[12] Chaovalitwongse, W.A., C.A.S. Oliveira, B. Chiarini, P.M. Pardalos, and M.G.C. Resende (2011), "Revised GRASP with Path-Relinking for the Linear Ordering Problem," *Journal of Combinatorial Optimization*, 22:572–593.

[13] Choi, S., S. Kariv, W. Müller, and D. Silverman (2013), "Who Is (More) Rational?" *American Economic Review*, forthcoming.

[14] Dean, M. and D. Martin (2012), "Testing for Rationality with Consumption Data," mimeo.

[15] Echenique, F, S. Lee, and M. Shum (2011), "The Money Pump as a Measure of Revealed Preference Violations," *Journal of Political Economy*, 119(6):1201–1223.

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

[17] Fishburn, P. and A. Rubinstein (1982), "Time preference," *International Econonomic Review*, 23:677–694.

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

[19] Genesove, D. and C. Mayer (2001), "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116:1233-1260.

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

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

[22] Gul, F., P. Natenzon and W. Pesendorfer (2012), "Random Choice as Behavioral Optimization," mimeo.

[23] Halevy, Y., D. Persitz and L. Zrill (2012), "Parametric Recoverability of Preferences," mimeo.

[24] Harless, D. and C.F. Camerer (1994), "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica* 62:1251-1289.

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

[26] 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.

[27] Kahneman, D., Knetsch, J., and R. Thaler (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98:1325–1348.

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

[29] Laibson, D. (1997), "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112:443–77.

[30] Luce, R.D. (1959), *Individual Choice Behavior: A Theoretical Analysis*. Wiley: New York.

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

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

[33] Manzini, P. and M. Mariotti (2012), "Categorize Then Choose: Boundedly Rational Choice and Welfare," *Journal of the European Economic Association*, 10(5):1141–1165.

[34] Manzini, P. and M. Mariotti (2013), "Stochastic Choice and Consideration Sets," *Econometrica*.

[35] Masatlioglu, Y. and D. Nakajima (2012), "Choice by Iterative Search," *Theoretical Economics*, forthcoming.

[36] Masatlioglu, Y., D. Nakajima and E.Y. Ozbay (2012), "Revealed Attention," *American Economic Review*, 102(5):2183–2205.

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

[38] Masatlioglu, Y. and E.A. Ok (2013), "A Canonical Model of Choice with Initial Endowments," *Review of Economic Studies*, forthcoming.

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

[40] McFadden D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," P. Zarembka (Ed.), *Frontiers of Econometrics*, Academic Press.

[41] O'Donoghue, T. and M. Rabin (1999), "Doing It Now or Later," *American Economic Review*, 89(1):103–124.

[42] Ok, E.A. and Y. Masatlioglu (2007), "A Theory of (Relative) Discounting," *Journal of Economic Theory*, 137:214–245.

[43] Ok, E.A., P. Ortoleva and G. Riella (2012), "Revealed (P)reference Theory," mimeo.

[44] Palacios-Huerta, I. and O. Voli (2004), "The Measurement of Intellectual Influence," *Econometrica*, 72(3):963–977.

[45] Rubinstein, A. (1980), "Ranking the Participants in a Tournament," *SIAM Journal of Applied Mathematics*, 38:108–111.

[46] Rubinstein, A. and Y. Salant (2012), "Eliciting Welfare Preferences from Behavioral Datasets," *Review of Economic Studies*, 79:375–387.

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

[48] Smeulders, B., L. Cherchye, B. De Rock and F.C.R. Spieksma (2012), "Goodness of Fit Measures for Revealed Preference Tests: Complexity Results and Algorithms," mimeo.

[49] Strotz, R.H. (1956), "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23:165–180.

[50] 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.

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

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

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

[54] Xu, Y. and L. Zhou (2007), "Rationalizability of Choice Functions by Game Trees," *Journal of Economic Theory*, 134:548–556.