A Measure of Rationality and Welfare

Jose Apesteguia

Catalan Institute for Research and Advanced Studies, Universitat Pompeu Fabra, and Barcelona Graduate School of Economics

Miguel A. Ballester

Universitat Autonoma de Barcelona and Barcelona Graduate School of Economics

Evidence showing that individual behavior often deviates from the classical principle of preference maximization has raised at least two important questions: (1) How serious are the deviations? (2) What is the best way to analyze choice behavior in order to extract information for the purpose of welfare analysis? This paper addresses these questions by proposing a new way to identify the preference relation that is closest, in terms of welfare loss, to the revealed choice.

I. Introduction

The standard model of individual behavior is based on the maximization principle, whereby the individual chooses the alternative that maximizes a preference over the menu of available alternatives. This has two key advantages. The first is that it provides a simple, versatile, and powerful

We thank Larbi Alaoui, Jorge Alcalde-Unzu, Stéphane Bonhomme, Guillermo Caruana, Christopher Chambers, Tugce Cuhadaroglu, Eddie Dekel, Jerry Green, Philipp Kircher, Paola Manzini, Marco Mariotti, Paul Milgrom, Mauro Papi, John K.-H. Quah, Collin Raymond, Larry Samuelson, Karl Schlag, Jesse Shapiro, and three referees for very helpful comments and Florens Odendahl for outstanding research assistance. Financial support by the Spanish Commission of Science (ECO2014-53051, ECO2014-56154, ECO2012-34202, ECO2011-25295, ECO2010-09555-E, ECO2008-04756-Grupo Consolidado C) and FEDER is gratefully acknowledged. Data are provided as supplementary material online.

Electronically published October 30, 2015
[*Journal of Political Economy*, 2015, vol. 123, no. 6]

account of individual behavior. The second is that it suggests the maximized preference as a tool for individual welfare analysis.

Research in recent years, however, has produced increasing amounts of evidence documenting deviations from the standard model of individual behavior.[1] The violation in some instances of the maximization principle raises at least two important questions: (1) How serious are the deviations from the classical theory? (2) What is the best way to analyze individual choice behavior in order to extract information for the purpose of welfare analysis?

The successful answering of question 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 closely their behavior approaches this benchmark. Addressing question 2, meanwhile, should help us to distinguish 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 useful for performing welfare analysis.

Although these two questions are intimately related, the literature has treated them separately. This paper provides the first joint approach to measuring rationality and welfare. Relying on standard revealed preference data, we propose the *swaps index*, which measures the welfare loss of the inconsistent choices with respect to the preference relation that comes closest to the revealed choices, the *swaps preference*. The swaps index evaluates the inconsistency of an observation with respect to a preference relation in terms of the number of alternatives in the menu that rank above the chosen one. That is, it counts the number of alternatives that must be swapped with the chosen alternative in order for the preference relation to rationalize the individual's choices. Then, the swaps index considers the preference relation that minimizes the total number of swaps in all the observations, weighted by their relative occurrence in the data.

To the best of our knowledge, the literature on rationality indices starts with the Afriat (1973) proposal for a consumer setting, which is to measure the amount of adjustment required in each budget constraint

[1] Some phenomena that have attracted a great deal of empirical and theoretical attention and that 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. For experimental papers, see May (1954), Thaler (1980), Tversky and Kahneman (1981), and Iyengar and Lepper (2000). Some theoretical papers reacting to this evidence are Kalai, Rubinstein, and Spiegler (2002), Bossert and Sprumont (2003), Masatlioglu and Ok (2005, 2014), Manzini and Mariotti (2007, 2012), Xu and Zhou (2007), Salant and Rubinstein (2008), Green and Hojman (2009), Masatlioglu, Nakajima, and Ozbay (2012), and Ok, Ortoleva, and Riella (2012).

to avoid any violation of the maximization principle. Varian (1990) extends the Afriat proposal to contemplate a vector of wealth adjustments, with different adjustments in the different observations.[2] An alternative proposal by Houtman and Maks (1985) is to compute the maximal subset of the data that is consistent with the maximization principle.[3] Yet a third approach, put forward by Swofford and Whitney (1987) and Famulari (1995), entails counting the number of violations of a consistency property detected in the data. Echenique, Lee, and Shum (2011) make use of the monetary structure of budget sets to suggest a version of this notion, the money pump index, which considers the total wealth lost in all revealed cycles. The swaps index contributes to the measurement of rationality in a singular fashion by evaluating inconsistent behavior directly in terms of welfare loss. It is also the first axiomatically based measure to appear in the literature. In Section III.A, we illustrate the contrast between the swaps index treatment of rationality measurement and these alternative proposals.

There is a growing number of papers analyzing individual welfare when the individual's behavior is inconsistent. Bernheim and Rangel (2009) add to the standard choice data the notion of ancillary conditions, which are assumed to be observable and potentially to affect individual choice but 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.[4] The proposal of Green and Hojman (2009) is to identify a list of conflicting selves, aggregate them to induce the revealed choices, and then perform individual welfare analysis using the aggregation rule. Nishimura (2014) builds a transitive welfare ranking on the basis of a nontransitive preference relation.[5] The swaps index uses the revealed choices, as in the classical approach, to suggest a novel welfare ranking, the swaps preference, interpreted as the best approximation to the choices of the individual and complemented with a measure of its accuracy: the incon-

[2] 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.

[3] Dean and Martin (2012) suggest an extension that weights the binary comparisons of the alternatives by their monetary values. Choi et al. (2014) apply the measures of Afriat and Houtman and Maks to provide valuable information on the relationship between rationality and various demographics.

[4] Chambers and Hayashi (2012) extend Bernheim and Rangel's model to probabilistic settings.

[5] Other approaches include Masatlioglu et al. (2012), Rubinstein and Salant (2012), and Baldiga and Green (2013). There are also papers describing methods for ranking objects such as teams or journals, based on a given tournament matrix describing the paired results of the objects (see Rubinstein 1980; Palacios-Huerta and Volij 2004).

sistency value. In Section III.B, we illustrate by way of examples other differences between our proposal and these other approaches.

In Section IV, we study the capacity of the swaps index to recover the true preference relation from collections of observations that, for a variety of reasons, may contain mistakes and hence potentially reveal inconsistent choices. We show that this is in fact the case for a wide array of stochastic choice models.

In Section V, we propose seven desirable properties of any inconsistency index relying only on endogenous information arising from the choice data and show that they characterize the swaps index. Then, in Section VI, we characterize several generalizations of the swaps index, together with versions of the classical Varian and Houtman-Maks indices within our framework. Section VII applies the swaps index to the experimental data of Harbaugh, Krause, and Berry (2001).

In the online appendix we discuss the relaxation of three assumptions made in the setup. We first show that it is immediate to make the swaps index capable of considering classes of preference relations with further structure, such as those admitting an expected utility representation. We then show how to extend the swaps index to include the treatment of indifferences. Third, we argue that it is possible to construct a natural version of the swaps index ready for application in settings with infinite sets of alternatives.

II. Framework and Definition of the Swaps Index

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. We denote by $\mathcal{F}$ the set of all possible collections of observations. The collection f allows us to entertain 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.

Another key feature of our framework is 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 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 $f(A, a) > 0$ implies $m(P, A) = a$.[6] 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, which we denote by P^r.[7] Clearly, not every collection 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.

We are now in a position to formally introduce our approach. Consider a given preference relation P and an observation (A, a) that is inconsistent with the maximization of P. This implies that there are a number of alternatives in A that, despite being preferred to the chosen alternative a according to P, are nevertheless ignored by the individual. We can therefore entertain that the inconsistency of observation (A, a) with respect to P entails consideration of 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. If every single observation is weighted by its relative occurrence in the data, the inconsistency of f with respect to P can be measured by $\sum_{(A,a)} f(A, a)\, |\{x \in A : xPa\}|$. The swaps index I_S adopts this criterion and finds the preference relations P_S that minimize the weighted sum of swaps. We refer to P_S as the swaps preference relations. Formally,

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

In summary, the swaps index enables the joint treatment of inconsistency and welfare analysis. 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. It identifies the preference relations closest to the revealed data, the swaps preferences, measuring their inconsistency in terms of the associated welfare loss. In Appendix B we show that almost all collections of observations have a unique swaps preference; that is, the measure of

[6] Notice that, since P is a linear order, if there exist $a, b \in A$ with $a \neq b$ such that $f(A, a) > 0$ and $f(A, b) > 0$, then f is not rationalizable.

[7] The purpose here is to create a bijection between $\mathcal{P}$ and a subset of the rationalizable collections. The set $\mathcal{R}$ is one way of creating this bijection, which comes without loss of generality.

all collections with a nonunique swaps preference is zero. We then typically talk about the swaps preference without further considerations unless the distinction is relevant.[8] We now illustrate the swaps index and the swaps preference by way of two examples.

EXAMPLE 1. Consider the set of alternatives $X = \{1, \ldots, k\}$. Suppose that the collection f contains observations involving all the subsets of X and is completely consistent with the preference relation P, ranking the alternatives as $1P2P \cdots Pk$. Now consider the collection of observations g involving the consistent evidence f with a high frequency $(1 - \alpha)$ and the extra observation (X, x), $x > 1$, with a low frequency α. That is, $g = (1 - \alpha)f + \alpha\mathbf{1}_{(X,x)}$, where $\mathbf{1}_{(X,x)}$ denotes the collection with all the mass centered on the observation (X, x). Clearly, the collection g is not rationalizable. In order to determine the swaps index and the swaps preference for g, notice that for any $P' \neq P$ there is at least one pair of alternatives, z and y, with $y < z$ and $zP'y$. Hence, the weighted sum of swaps for P' is at least $(1 - \alpha)f(\{y, z\}, y)$. Meanwhile, preference P requires $x - 1$ swaps in the observation (X, x), and hence the weighted sum of swaps for P is exactly $\alpha(x - 1)$. For small values of α, it is clearly the case that $\alpha(x - 1) < (1 - \alpha)f(\{y, z\}, y)$, and therefore, $I_S(g) = \alpha(x - 1)$ and $P_S(g) = P$. Hence, in such cases the swaps preference coincides with the rational preference P, and the inconsistency attributed to g by the swaps index is the mass of the inconsistent observation α weighted by the number of swaps required to rationalize the inconsistent observation.

EXAMPLE 2. Let

$$f(\{x, y\}, x) = f(\{y, z\}, y) = \frac{1 - 2\alpha}{2}$$

and

$$f(\{x, y, z\}, y) = f(\{x, z\}, z) = \alpha,$$

where α is small. That is, there is large evidence that x is better than y and that y is better than z and some small evidence that y is better than x and z and that z is better than x. Notice that any preference in which y is ranked above x or z is ranked above y has a weighted sum of swaps of at least $(1 - 2\alpha)/2$. There is only one more preference to be analyzed, namely, $xPyPz$. This preference requires exactly one swap in menu $\{x, y, z\}$, where y is chosen, and also one swap in menu $\{x, z\}$ where z is chosen. The weighted sum of swaps of P is therefore 2α, which, for small values of α, is smaller than $(1 - 2\alpha)/2$, and hence $I_S(f) = 2\alpha$ and

[8] In addition, in App. C, we deal with the computational complexity of obtaining P_S.

$P_S(f) = P$. That is, the swaps preference rationalizes the large evidence of data,

$$f(\{x, y\}, x) = f(\{y, z\}, y) = \frac{1 - 2\alpha}{2},$$

and incurs some relatively small errors in $f(\{x, y, z\}, y) = f(\{x, z\}, z) = \alpha$.

III. Comparison to Alternative Measures

A. The Measurement of Rationality

In a consumer setting, Afriat (1973) suggests measuring the degree of relative wealth adjustment that, when applied to all budget constraints, avoids all violations of the maximization principle. The idea is that, when a portion of wealth is considered, all budget sets shrink, thus eliminating some revealed information, and thereby possibly removing some inconsistencies from the data. Thus, Afriat's proposal associates the degree of inconsistency in a collection of observations with the minimal wealth adjustment needed to make all the data consistent with the maximization principle.

We now formally define Afriat's index for our setting. Let $w_x^A \in (0, 1]$ be the minimum proportion of income in budget set A that must be removed in order to make x unaffordable. Then, given a menu A, if a proportion w of income is removed, all alternatives $x \in A$ with $w_x^A \leq w$ become unaffordable. We say that a collection f is w-rationalizable if there exists a preference relation P such that $f(A, a) > 0$ implies that aPx for every $x \in A \setminus \{a\}$ with $w_x^A > w$. Notice that when $w = 0$, this is but the standard definition of rationalizability. Afriat's inconsistency measure is defined as the minimum value w^* such that f is w^*-rationalizable. Note that we can alternatively represent this index in terms of preference relations, making its representation closer in spirit to that of the swaps index. To see this, suppose that P^* is a preference that w^*-rationalizes f. Then, for all observations (A, a) with $f(A, a) > 0$, all alternatives xP^*a must be unaffordable at w^*. Hence

$$w^* = \max_{(A,a):f(A,a)>0} \max_{x \in A, xP^* a} w_x^A.$$

Since no other preference can w-rationalize f for $w < w^*$, it is clearly the case that we can define Afriat's index as[9]

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

$$I_A(f) = \min_{P} \max_{(A,a):f(A,a)>0} \max_{x \in A, xPa} w_x^A.$$

Varian (1990) considers vectors of wealth adjustments **w**, with potentially different adjustments in the various observations. Then, Varian's index identifies the closest vector **w** to 0 that, under a certain norm, **w**-rationalizes the data. Here, given the structure of the swaps index, we consider the 1-norm and define Varian's index as follows:

$$I_V(f) = \min_{P} \sum_{(A,a)} f(A,a) \max_{x \in A, xPa} w_x^A.$$

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 the minimal subset to be discarded suggests itself as a measure of inconsistency. It follows immediately that, in our setting, the Houtman-Maks index, which we denote by I_{HM}, is but a special case of Varian's index when $w_x^A = 1$ for every A and every $x \in A$.

Finally, rationality has also been measured by counting the number of times in the data a consistency property is violated (see, e.g., Swofford and Whitney 1987; Famulari 1995). Consider for instance the case of the weak axiom of revealed preference (WARP). In our context, WARP is violated whenever there are two menus A and B and two distinct elements a and b in $A \cap B$ such that $f(A,a) > 0$ and $f(B,b) > 0$. Hence, we can measure the mass of violations of WARP by means of

$$I_W = \sum_{(A,a),(B,b):\{a,b\} \subseteq A \cap B, a \neq b} f(A,a) f(B,b).$$

Recently, Echenique et al. (2011) made 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 the generalized axiom of revealed preference (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. To illustrate the structure of this index in our framework, let us contemplate only violations of WARP (i.e., cycles of length 2). Consider a violation of WARP involving observations (A, a) and (B, b). The money pump reasoning evaluates the wealth lost in this cycle by adding up the minimal wealth $\tilde{w}_b^A$ that must be removed to make b unaffordable in A and the minimal wealth $\tilde{w}_a^B$ that must be removed to make a unaffordable in B.[10] Then,

[10] Notice that $\tilde{w}_x^A$, assumed to be strictly positive, is measured in dollars while Afriat's and Varian's weights w_x^A are proportions of wealth.

$\tilde{w}_b^A + \tilde{w}_a^B$ represents the money that could be pumped by an arbitrager from the WARP violation. Now, given the vector of weights $\tilde{\mathbf{w}}$, the WARP money pump index can be defined as[11]

$$I_{W-MP} = \sum_{(A,a),(B,b):\{a,b\}\subseteq A\cap B, a\neq b} f(A,a)f(B,b)(\tilde{w}_b^A + \tilde{w}_a^B).$$

In order to illustrate the differences between all these indices and the swaps index, let us reconsider example 1. Consider then two different scenarios in which $x = k$ and $x = 2$, that is, $g_k = (1-\alpha)f + \alpha\mathbf{1}_{(X,k)}$ and $g_2 = (1-\alpha)f + \alpha\mathbf{1}_{(X,2)}$. Intuitively, collection g_k involves a more severe inconsistency, since the observation in question is one in which the individual chooses the worst possible alternative, alternative k, while ignoring all the rest. Collection g_2 also shows some inconsistency with the maximization principle, but this inconsistency is orders of magnitude lower, since it involves choosing the second-best available option, that is, option 2. It follows immediately from the discussion in example 1 that the swaps index ranks these two collections in accordance with the above intuition, that is, $I_S(g_k) = \alpha(k-1) > \alpha = I_S(g_2)$. Afriat's and Varian's judgment of these collections depends crucially on the monetary values of the alternatives, which need not necessarily coincide with the welfare ranking and hence may lead to counterintuitive conclusions. For example, if 1 is the least expensive alternative in menu X, that is, $w_1^X \geq w_t^X$ for all $t \leq k$, Varian's approach involves removing income until alternative 1 becomes unaffordable, regardless of the scenario. Hence, both collections would be equally inconsistent. Note that, for Afriat, the mass of violations is irrelevant, and hence if removing option 1 from X is costly and removing alternative k from all menus is cheaper, it may be the case that g_k is w-rationalizable for some value $w < w_1^X$, while g_2 is not. Therefore, Afriat's index may judge g_k as being less inconsistent than g_2. With respect to Houtman-Maks's index, since the inconsistencies in both scenarios are of identical size, α, I_{HM} does not discriminate between them. Finally, the assessment provided by WARP violation index I_W depends on the specific nature of f. To illustrate, consider, for example, that $k = 3$ and that

$$f(X,1) = f(\{1,3\},1) = f(\{2,3\},2) = \beta$$

and

$$f(\{1,2\},1) = 1 - 3\beta.$$

It follows immediately that

[11] It is immediate to extend this index to consider cycles of any length, something that we avoid here for notational convenience.

$$I_W(g_k) = 3\alpha\beta(1-\alpha) < (1-\alpha)\alpha(1-2\beta) = I_W(g_2)$$

whenever $\beta < 1/5$, and hence scenario 2 is regarded as the more inconsistent of the two. Although index I_{W-MP} weights both sides of the above inequality by $\tilde{\mathbf{w}}$, the inequality still holds for certain nonnegligible values of β.

B. *The Measurement of Welfare*

Let us illustrate our approach to welfare analysis by contrasting it first with two proposals: Bernheim and Rangel (2009) and Green and Hojman (2009). Although these two papers tackle the problem from different angles, they independently suggest the same notion of welfare. Let us denote by $\overline{P}$ the Bernheim-Rangel-Green-Hojman welfare relation, 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 swaps preference P_S. It turns out to be the case that the two welfare relations are fundamentally different. It follows immediately that P_S is not contained in $\overline{P}$ because 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 in which 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 ways.

Nishimura (2014) has recently proposed a different approach, the transitive core. Given a complete nonnecessarily transitive relation $\succcurlyeq$, the transitive core declares an alternative x preferred to alternative y whenever, for every z, (i) $y \succcurlyeq z$ implies $x \succcurlyeq z$ and (ii) $z \succcurlyeq x$ implies $z \succcurlyeq y$. Like $\overline{P}$, the transitive core may be incomplete, and since relative frequencies are not considered, the transitive core may go in the opposite direction to P_S.

We illustrate the differences between the swaps preference and the proposals here presented, using example 2 above. We argued there that xP_SyP_Sz. Note now that $z\overline{P}x$ since x is never chosen in the presence of z, and hence $\overline{P}$ and P_S follow different directions. Moreover, if $\succcurlyeq$ is understood to be the revealed preference, y is ranked above x by the transitive core, and hence this is different from P_S too.

Finally, notice that the swaps preference P_S comes, by construction, with the associated inconsistency I_S, which provides a measure of the credibility of P_S. A low value of I_S naturally gives credit to P_S, while high values of I_S may call for more cautious conclusions regarding the true welfare of the individual, either by focusing on subsets of alternatives over which violations are less dramatic (in the spirit of the aforementioned approaches) or by adopting a particular boundedly rational model of choice.

IV. Recoverability of Preferences and the Swaps Index

Consider a decision maker who evaluates alternatives according to the preference relation P but when it comes to selecting the preferred option sometimes chooses a suboptimal alternative. Mistakes can occur for various reasons, such as lack of attention, errors of calculation, misunderstanding of the choice situation, trembling hand when about to select the desired alternative, inability to implement the desired choice, and so forth. Whatever the specific model, mistakes generate a potentially inconsistent collection of observations f. This raises the issue of whether the swaps index has the capacity to recover the preference relation P from the observed choices f.

We show below that the swaps index identifies the true underlying preference for models that generate collections of observations in which, for any pair of alternatives, the better one is revealed preferred to the worse one more often than the reverse. Formally, we say that the collection f generated by a model satisfies *P-monotonicity* if xPy implies that $\sum_{A\supseteq\{x,y\}} f(A,x) \geq \sum_{A\supseteq\{x,y\}} f(A,y)$, where the inequality is strict whenever $\sum_{a\in\{x,y\}} f(\{x,y\},a) > 0$. In order to assess the generality of this result, we first show that a diverse number of highly influential classes of stochastic choice models satisfy this property.

Random utility models.—Suppose that the individual evaluates the alternatives by way of a utility function $u: X \to \mathbb{R}_{++}$.[12] At the moment of choice, this valuation is subject to an additive random error component. That is, when choosing from A, the true valuation of alternative x, $u(x)$, is subject to a random independent and identically distributed (i.i.d.) term, $\epsilon_A(x)$, which follows a continuous distribution, resulting in the final valuation $U(x) = u(x) + \epsilon_A(x)$. Then, the probability by which alternative a is chosen from A is the probability of a being maximal in A according to U, that is, $\Pr[a = \arg\max_{x\in A} U(x)]$.[13] Let ρ denote the probability distribution over the menus of options available to the individual, where

[12] In consonance with our analysis, assume that $u(x) \neq u(y)$ for every $x, y \in X$, $x \neq y$. Also, notice that the preference relation P of the individual is simply the one for which $u(x) > u(y) \Leftrightarrow xPy$. This also applies for the utility function used in the choice control models below.

[13] Notice that, since $\epsilon_A(x)$ is continuously distributed, the probability of ties is zero and hence $\Pr[a = \arg\max_{x\in A} U(x)]$ is well defined. Classic references for this class of models are Luce (1959) and McFadden (1974). See also Gul, Natenzon, and Pesendorfer (2014).

$\rho(A)$ denotes the probability of confronting $A \subseteq X$. We can now define the collection of observations generated by a random utility model as

$$f_{\mathrm{RUM}}(A, a) = \rho(A)\Pr\left[a = \arg\max_{x\in A} U(x)\right]$$

for every $(A, a) \in \mathcal{O}$. While the most widely used random utility models (logit, probit) have menu-independent errors, our formulation allows for menu-dependent utility errors, as in the contextual utility model of Wilcox (2011).

Tremble models.—The mistake structure in random utility models depends on the cardinal utility values of the options. Another way to model mistakes is as constant probability shocks that perturb the selection of the optimal alternative. That is, an individual facing menu A chooses her optimal option with high probability $1 - \mu_A > 1/2$ and, with probability μ_A, trembles and overlooks the optimal option. In the spirit of the trembling hand perfect equilibrium concept in game theory, in the event of a tremble, any other option is selected with equal probability. Formally, $f_{\text{TM-per}}(A, a) = \rho(A)(1 - \mu_A)$ when $a = m(P, A)$, and $f_{\text{TM-per}}(A, a) = \rho(A)[\mu_A/(|A| - 1)]$ otherwise, where ρ is defined as above. Alternatively, in line with the notion of proper equilibrium in game theory, one may entertain that the perturbation process recurs among the surviving alternatives. That is, conditional on a shock involving the best option, with probability $1 - \mu_A$ the individual chooses the second-best option from A and with probability μ_A overlooks the second-best option, and so forth. In this case, the resulting collection of observations is $f_{\text{TM-pro}}(A, a) = \rho(A)(1 - \mu_A)\mu_A^{|\{x\in A: xPa\}|}$ for any alternative a other than the worst one in menu A, and $f_{\text{TM-pro}}(A, a) = \rho(A)\mu_A^{|A|-1}$ otherwise. We write f_{TM} to refer to both models, $f_{\text{TM-per}}$ and $f_{\text{TM-pro}}$.[14] As in the previous case, the class of tremble models that we are contemplating allows the error to depend on the particular menus.

Choice control models.—Consider the case in which being able to control the implementation of choice involves a cost. In such a situation, the agent evaluates the trade-off between the cost of control and the cost of deviating from her preferences and maximizes accordingly.[15] Following Fudenberg et al. (2014), consider a utility function $u : X \to \mathbb{R}_{++}$ and a continuous control function $c_A : [0, 1] \to \mathbb{R}$ that describes the cost of choosing any alternative from menu A with a given probability. The utility associated with the individual choosing a probability distribution p_A over A is therefore

[14] See Selten (1975) and Myerson (1978). See Harless and Camerer (1994) for a first treatment of the tremble notion in the stochastic choice literature.

[15] Alternative motivations for the models in this category include a desire for randomization, the cost of deviating from a social exogenous choice distribution, etc. See Mattsson and Weibull (2002) and Fudenberg, Iijima, and Strzalecki (2014) for a discussion.

$$\sum_{x\in A}[p_A(x)u(x) - c_A(p_A(x))].$$

The individual then selects a probability distribution p_A^* that maximizes this utility, that is,

$$p_A^* \in \arg\max_{p_A} \sum_{x\in A}[p_A(x)u(x) - c_A(p_A(x))].$$

Thus, by using ρ as above, we can define the collection generated by the choice control model as $f_{\text{CCM}}(A, a) = \rho(A)p_A^*(a)$.

Proposition 1 establishes that all the above models satisfy P-monotonicity.

PROPOSITION 1. f_{RUM}, f_{TM}, and f_{CCM} satisfy P-monotonicity.

Proof. We first analyze random utility models. Consider a menu A and alternatives $x, y \in A$ with xPy. Take a realization of the error terms such that U is maximized at y over the menu A. That is, $u(y) + \epsilon_A(y) > u(x) + \epsilon_A(x)$ and $u(y) + \epsilon_A(y) > u(z) + \epsilon_A(z)$ for any other option $z \in A\setminus\{x, y\}$. Then, consider the alternative realization of the errors, where y receives the shock $\epsilon_A(x)$, x receives the shock $\epsilon_A(y)$, and z receives the same shock $\epsilon_A(z)$. Since $u(x) > u(y)$, $u(x) + \epsilon_A(y) > u(y) + \epsilon_A(y) > u(z) + \epsilon_A(z)$ for all $z \in A\setminus\{x, y\}$, and also $u(x) + \epsilon_A(y) > u(y) + \epsilon_A(x)$. Then, the continuous i.i.d. nature of the errors within menu A guarantees that

$$\Pr\left[x = \arg\max_{w\in A} U(w)\right] > \Pr\left[y = \arg\max_{w\in A} U(w)\right].$$

This implies that $f_{\text{RUM}}(A, x) \geq f_{\text{RUM}}(A, y)$ with strict inequality if the menu A is such that $\rho(A) > 0$. Consequently,

$$\sum_{A\supseteq\{x,y\}} f_{\text{RUM}}(A, x) \geq \sum_{A\supseteq\{x,y\}} f_{\text{RUM}}(A, y),$$

with strict inequality whenever $\rho(A) > 0$ for at least one set A containing x and y and, clearly, P-monotonicity holds.

We now study tremble models. Consider a menu A and alternatives $x, y \in A$ with xPy. In the case of $f_{\text{TM-per}}$, notice that $x = m(P, A)$ implies that

$$\begin{aligned} f_{\text{TM-per}}(A, x) &= \rho(A)(1-\mu_A) \geq \rho(A)\mu_A \geq \rho(A)\frac{\mu_A}{|A|-1} \\ &= f_{\text{TM-per}}(A, y) \end{aligned}$$

while $x \neq m(P, A)$ implies that

$$f_{\text{TM-per}}(A, x) = \rho(A)\frac{\mu_A}{|A|-1} = f_{\text{TM-per}}(A, y).$$

In the case of $f_{\text{TM-pro}}$, if y is not the worst alternative in A,

$$f_{\text{TM-pro}}(A,x) = \rho(A)(1-\mu_A)\mu_A^{|\{z\in A:zPx\}|} \geq \rho(A)(1-\mu_A)\mu_A^{|\{z\in A:zPy\}|} = f_{\text{TM-pro}}(A,y).$$

If y is the worst alternative in A,

$$f_{\text{TM-pro}}(A,x) = \rho(A)(1-\mu_A)\mu_A^{|\{z\in A:zPx\}|} \geq \rho(A)\mu_A^{|A|-1} = f_{\text{TM-pro}}(A,y).$$

Then

$$\sum_{A\supseteq\{x,y\}} f_{\text{TM}}(A,x) \geq \sum_{A\supseteq\{x,y\}} f_{\text{TM}}(A,y),$$

with strict inequality whenever $\rho(A) > 0$ for at least one set A such that (i) in the case of $f_{\text{TM-per}}$, x is the best alternative in $A\supseteq\{x,y\}$, and (ii) in the case of $f_{\text{TM-pro}}$, $A\supseteq\{x,y\}$. This is clearly the case for $\{x,y\}$, and hence P-monotonicity holds.

Finally, we analyze choice control models. Consider a menu A and alternatives $x, y\in A$ with xPy. We first prove that $f_{\text{CCM}}(A,x)\geq f_{\text{CCM}}(A,y)$. Suppose, by contradiction, that $f_{\text{CCM}}(A,x) < f_{\text{CCM}}(A,y)$ or, equivalently, $p_A^*(x) < p_A^*(y)$. Consider p_A' with $p_A'(x) = p_A^*(y)$, $p_A'(y) = p_A^*(x)$, and $p_A'(z) = p_A^*(z)$ for all $z\in A\setminus\{x,y\}$. Since, by assumption, $u(x) > u(y)$, it is the case that

$$\sum_{w\in A}[p_A'(w)u(w) - c_A(p_A'(w))] > \sum_{w\in A}[p_A^*(w)u(w) - c_A(p_A^*(w))],$$

thus contradicting the optimality of p^*. Since this is true for every menu,

$$\sum_{A\supseteq\{x,y\}} f_{\text{CCM}}(A,x) \geq \sum_{A\supseteq\{x,y\}} f_{\text{CCM}}(A,y)$$

holds. For the strict part, notice that continuity of $c_{\{x,y\}}$ prevents the optimal solution p^* from being constant in $\{x,y\}$, and hence P-monotonicity follows. QED

We now show that the swaps index always identifies the true underlying preference in models that satisfy P-monotonicity and, particularly, that the presence of all the menus in the data guarantees that the swaps index uniquely identifies the preference.

THEOREM 1. If f satisfies P-monotonicity, then P is a swaps preference of f. If, moreover, $\sum_{a\in A} f(A,a) > 0$ holds for every menu A, then P is the unique swaps preference of f.

Proof. Let f be P-monotone. Consider any preference P' different from P. Then, there exist at least two alternatives a_1 and a_2 that are consecutive in P', with $a_2P'a_1$ but a_1Pa_2. Define a new preference P'' by $xP''y \Leftrightarrow xP'y$ whenever $\{x, y\} \neq \{a_1, a_2\}$ and $a_1P''a_2$. That is, P'' is simply defined by changing the position of the consecutive alternatives a_1 and a_2 in P', reconciling their comparison with that of preference P and leaving all else the same. We now show that P'' rationalizes data with fewer swaps than P'. To see this, simply notice that the swaps computation will be affected only by menus A such that $A \supseteq \{a_1, a_2\}$. Also, for any of such sets, since both alternatives are consecutive in both P' and P'', the swaps computation will be affected only by observations of the form (A, a_1) and (A, a_2) and clearly,

$$\sum_{(A,a)} f(A, a) \mid \{x \in A : xP''a\} \mid \; = \sum_{(A,a)} f(A, a) \mid \{x \in A : xP'a\} \mid + \sum_{A \supseteq \{a_1, a_2\}} f(A, a_2) - \sum_{A \supseteq \{a_1, a_2\}} f(A, a_1).$$

Since f is P-monotone, the latter is smaller than or equal to $\sum_{(A,a)} f(A, a) \mid \{x \in A : xP'a\} \mid$, as desired. Given the finiteness of X, repeated application of this algorithm leads to preference P and proves that

$$\sum_{(A,a)} f(A, a) \mid \{x \in A : xPa\} \mid \; \leq \sum_{(A,a)} f(A, a) \mid \{x \in A : xP'a\} \mid .$$

Hence, P is an argument that minimizes the swaps index. Whenever $\sum_{a \in A} f(A, a) > 0$ holds for every menu A, it is in particular satisfied for the menus $\{a_1, a_2\}$ involved in each step of the previous algorithm. By P-monotonicity, the corresponding inequalities are strict, and therefore P is the unique swaps preference. QED

Theorem 1 provides a simple test to guarantee that the swaps index identifies the true preference of a particular choice model. Two questions naturally arise at this point. The first is whether other indices may also systematically recover it when P-monotonicity holds. It is easy to see that the Afriat and Varian indices do not possess this recovery property in general, since they depend on the monetary structure of the alternatives, which is not necessarily aligned with preferences. To see this, consider the simplest case in which $X = \{x, y\}$ and suppose xPy. Notice that if $f(X, y) \neq 0$, I_A recovers P if and only if $w_x^X \leq w_y^X$. Similarly, I_V recovers P if and only if $w_x^X / w_y^X \leq f(X, x)/f(X, y)$. Without these extra conditions, I_A and I_V are unable to recover P. Moreover, indices based on the number of violations of a rationality property, such as I_W or the money pump index, are also unable to recover the preference, since these indices are not built to identify any particular preference, nor can they be written in this form.[16] Finally, since I_{HM} does not take into consideration the severity of the inconsistencies, it is also unable to recover P from P-monotone models. To see this, let $X = \{x, y, z\}$ with $xPyPz$, and consider a model generating a P-monotone collection f such that $f(\{y, z\}, y) < f(\{y, z\}, z)$. It is immediate that the mass of inconsistent observations in f with respect to $xP'zP'y$ is strictly lower than that of P, and hence the optimal preference for I_{HM} cannot be P.[17]

The next question concerns choice models not satisfying P-monotonicity for which I_S does not recover the true preference. A leading case is consideration set models.[18] In this setting, the individual considers each alternative with a given probability and then chooses the maximal alternative from those that have been considered, and hence good alternatives may be chosen with low probability. We can address this case by using a slight generalization of I_S, the nonneutral swaps index I_{NNS} proposed in Section VI.B.

V. Axiomatic Foundations for the Swaps Index

Here, we propose seven properties that shape the way in which an inconsistency index I treats different types of collections of observations. We then show that the swaps index is characterized by this set of properties.

Continuity (*CONT*).—The index 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, which is justified in the standard fashion. That is, it is desirable that a small variation in the data does not cause an abrupt change in the inconsistency value.

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 can be explained by maximizing the same preference relation.

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

[16] In Sec. V we discuss the axiom piecewise linearity, which allows for the recoverability of preferences.

[17] Again, in Sec. V we discuss the axiom disjoint composition, which allows us to account for the severity of the inconsistencies.

[18] See Masatlioglu et al. (2012) for a deterministic modeling and Manzini and Mariotti (2014) for a recent stochastic model.

$$I(\alpha f + (1 - \alpha)g) \geq \alpha I(f) + (1 - \alpha)I(g).$$

.

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

Piecewise linearity (*PWL*).—The index I is a piecewise linear function over $|\mathcal{P}|$ pieces. That is, there are $|\mathcal{P}|$ subsets of $\mathcal{F}$, the union of which is $\mathcal{F}$ such that for every pair f, g belonging to the same subset and every $\alpha \in [0, 1]$,

$$I(\alpha f + (1 - \alpha)g) = \alpha I(f) + (1 - \alpha)I(g).$$

Piecewise linearity brings two features: the piecewise nature of the index and the linear structure of the index over each piece. Let us now elaborate on the desirability of these two features.

Notice that the piecewise assumption in piecewise linearity is attractive from the recoverability of preferences perspective and hence is critical for predicting behavior and enabling individual welfare analysis. An index satisfying the piecewise assumption divides the set of collections of observations $\mathcal{F}$ into $|\mathcal{P}|$ classes. Thus, as any preference is linked to one and only one of such classes, every single collection of observations, even the nonrationalizable ones, can be linked to a specific preference relation.

Within each of the pieces, piecewise linearity makes the index react monotonically with respect to inconsistencies, whether they are (i) of the same type, thus making the index react to the mass of an inconsistency, or (ii) of different types, thus making the index react to the accumulation of several different inconsistencies. To enable formal study of these implications, we introduce a useful class of collections of observations, which we describe as perturbed. Consider a rationalizable collection of observations $r \in \mathcal{R}$ and an observation $(A, a) \in \mathcal{O}$. An ϵ-perturbation of r in the direction of (A, a) involves replacing an ϵ-mass of optimal choices $(A, m(P^r, A))$ with the possibly suboptimal choices (A, a).[19] We denote

[19] Obviously, the value of ϵ must be lower than $r(A, m(P^r, A))$.

such a perturbed collection by $r^{\epsilon(A,a)} = r + \epsilon\mathbf{1}_{(A,a)} - \epsilon\mathbf{1}_{(A,m(P^r,A))}$ and the collection in which two different ϵ-perturbations take place by

$$r^{\epsilon(A,a)}_{\epsilon(B,b)} = r + \epsilon\mathbf{1}_{(A,a)} - \epsilon\mathbf{1}_{(A,m(P^r,A))} + \epsilon\mathbf{1}_{(B,b)} - \epsilon\mathbf{1}_{(B,m(P^r,B))}.$$

The following lemma, proved in Appendix A, establishes the above implications.

Lemma 1. Let I be an inconsistency index satisfying PWL, CONT, and RAT. Consider any collection $r \in \mathcal{R}$ and any two different observations (A, a), (B, b) such that $a \neq m(P^r, A)$ and $b \neq m(P^r, B)$. For any two sufficiently small real values $\epsilon_1 > \epsilon_2 \geq 0$,

1. reactivity to the mass of an inconsistency: $I(r^{\epsilon_1(A,a)}) > I(r^{\epsilon_2(A,a)})$;
2. reactivity to several inconsistencies: $I(r^{\epsilon_1(A,a)}_{\epsilon_1(B,b)}) > \max\{I(r^{\epsilon_1(A,a)}), I(r^{\epsilon_1(B,b)})\}$.

The proof of lemma 1 explicitly shows how PWL implies reactivity of the index to both the mass and the types of inconsistencies in a linear fashion, namely, $I(r^{\epsilon_1(A,a)}) = (\epsilon_1/\epsilon_2)I(r^{\epsilon_2(A,a)})$ and $I(r^{\epsilon_1(A,a)}_{\epsilon_1(B,b)}) = I(r^{\epsilon_1(A,a)}) + I(r^{\varepsilon_1(B,b)})$.

Ordinal consistency (*OC*).—For every $(A, a) \in \mathcal{O}$ and every $r, \tilde{r} \in \mathcal{R}$ such that $r(\{x, y\}, x) = \tilde{r}(\{x, y\}, x)$ whenever $x, y \in A$, it is $I(r^{\epsilon(A,a)}) = I(\tilde{r}^{\epsilon(A,a)})$ for any sufficiently small $\epsilon > 0$.

Ordinal consistency is in the spirit of the classical properties of independence of irrelevant alternatives. A small perturbation of the type (A, a) generates the same inconsistency in two rationalizable collections r and $\tilde{r}$ that coincide in the ranking of the alternatives within A but may diverge in the ranking of alternatives outside A. In other words, the order of alternatives not involved in the inconsistency is inconsequential. In line with the standard justification for such a property, one may simply contend that when evaluating a perturbed collection, any alternative not involved in the perturbation at hand should not matter.

Disjoint composition (*DC*).—For every (A_1, a), $(A_2, a) \in \mathcal{O}$ such that $A_1 \cap A_2 = \{a\}$ and every $r \in \mathcal{R}$, it is $I(r^{\epsilon(A_1\cup A_2,a)}) = I(r^{\epsilon(A_1,a)}_{\epsilon(A_2,a)})$ for any sufficiently small $\epsilon > 0$.

In words, disjoint composition states that, given a rationalizable collection r, a small perturbation of the type $(A_1 \cup A_2, a)$ can be broken down into two small perturbations of the form (A_1, a) and (A_2, a), provided that A_1 and A_2 share no alternative other than a. By iteration, an index having this property is able to reduce the inconsistency of the observation into inconsistencies involving binary comparisons. This property is desirable for several reasons. First, from a purely normative point of view, notice that the standard welfare approach is constructed precisely on the basis of binary comparisons. Hence, an index that aims to capture the severity of an inconsistency in terms of the welfare loss involved must likewise be based on binary comparisons. Second, from a practical point of view, this decomposition facilitates the tractability of the data by compacting it into a unique matrix of binary choices. To illustrate, notice that both $r^{\epsilon(A_1 \cup A_2, a)}$ and $r_{\epsilon(A_2,a)}^{\epsilon(A_1,a)}$ correspond to the following summary of binary revealed choices. Whenever $xP^r a$ and $x \in A_1 \cup A_2$, ϵ percent of the data is inconsistent with x being preferred to a. No inconsistencies arise in any other comparison of two alternatives. Disjoint composition implies that this summary is the only relevant information and hence declares the two collections equally inconsistent.

In order to introduce our last property, let us consider the following notation. Given a permutation σ over the set of alternatives X, for any collection f we denote by $\sigma(f)$ the permuted collection 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 2 states the characterization result.

THEOREM 2. 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.

Proof. It is immediate to see that any positive scalar transformation of the swaps index satisfies the axioms. By way of seven steps, we show that an index satisfying the axioms is a transformation of the swaps index.

Step 1: Following the proof of lemma 1, consider the convex hulls of the closure of the $|\mathcal{P}|$ subsets of collections. Reasoning analogously, for every $r \in \mathcal{R}$, there exists $\alpha^r \in (0, 1)$ such that, for every observation (A, a) and every $\alpha \in [0, \alpha^r]$, the collection $\alpha\mathbf{1}_{(A,a)} + (1 - \alpha)r$ belongs to the convex hull of r. We then define, for every r and (A, a), the weight

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

Now notice that, whenever $aP^r x$ for all $x \in A \setminus \{a\}$, the collection $\alpha^r\mathbf{1}_{(A,a)} + (1 - \alpha^r)r$ is rationalizable by P^r and RAT implies $w(P^r, A, a) = 0$. Otherwise, it follows that $r(A, x) > 0$ with $x \neq a$, which implies that observations (A, a) and (A, x) have positive mass in the collection $\alpha^r\mathbf{1}_{(A,a)} + (1 - \alpha^r)r$, and RAT guarantees that $w(P^r, A, a) > 0$.

Step 2: We now prove that, whenever $f \in \mathcal{F}$ and $r \in \mathcal{R}$ belong to the same convex hull, it is the case that $I(f) = \sum_{(A,a)} f(A, a)w(P^r, A, a)$. By RAT and PWL,

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

Notice that

$$\begin{aligned}\alpha^r f + (1-\alpha^r)r &= \alpha^r\left(\sum_{(A,a)} f(A,a)\mathbf{1}_{(A,a)}\right) + (1-\alpha^r)r \\ &= \sum_{(A,a)} f(A,a)[\alpha^r \mathbf{1}_{(A,a)} + (1-\alpha^r)r].\end{aligned}$$

By definition of α^r, all collections $\alpha^r\mathbf{1}_{(A,a)} + (1-\alpha^r)r$ belong to the convex hull of r and all convex combinations of such collections must also lie in it. We can thus apply linearity repeatedly to obtain

$$\begin{aligned}I(f) &= \frac{I(\alpha^r f + (1-\alpha^r)r)}{\alpha^r} = \frac{\sum_{(A,a)} f(A,a) I(\alpha^r \mathbf{1}_{(A,a)} + (1-\alpha^r)r)}{\alpha^r} \\ &= \sum_{(A,a)} f(A,a) w(P^r, A, a).\end{aligned}$$

Step 3: Here, we prove that, for every $f \in \mathcal{F}$, $I(f) = \min_P \sum_{(A,a)} f(A,a)w(P,A,a)$. We first prove that, for every $r \in \mathcal{R}$, $I(f) \leq \sum_{(A,a)} f(A,a)w(P^r,A,a)$. By RAT and CONC,

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

By definition of α^r, all collections $\alpha^r\mathbf{1}_{(A,a)} + (1-\alpha^r)r$ belong to the convex hull of r, and hence $\alpha^r f + (1-\alpha^r)r$ also belongs to the hull. By steps 1 and 2, we know that

$$I(\alpha^r f + (1-\alpha^r)r) = \alpha^r \sum_{(A,a)} f(A,a)w(P^r,A,a)$$

and hence,

$$I(f) \leq \sum_{(A,a)} f(A,a)w(P^r,A,a).$$

By the proof of lemma 1 we know that each convex hull contains one and only one collection in $\mathcal{R}$, and then for every $f \in \mathcal{F}$, there exists $\hat{r} \in \mathcal{R}$ such that f and $\hat{r}$ lie in the same convex hull. Hence, step 2 and the above reasoning guarantee that

$$I(f) = \sum_{(A,a)} f(A,a)w(P^{\hat{r}},A,a) = \min_P \sum_{(A,a)} f(A,a)w(P,A,a).$$

Step 4: We now prove that, for every (A, a), and every pair P^r and $P^{\tilde{r}}$ such that $xP^r y \Leftrightarrow xP^{\tilde{r}} y$ whenever $x, y \in A$, it is the case that

$w(P^r, A, a) = w(P^{\tilde{r}}, A, a)$. To see this, notice that there exists a sufficiently small α such that $r^{\alpha(A,a)}$ and $\tilde{r}^{\alpha(A,a)}$ belong to the convex hulls of r and $\tilde{r}$, respectively. By steps 1 and 2 and OC, it is the case that

$$\alpha w(P^r, A, a) = I(r^{\alpha(A,a)}) = I(\tilde{r}^{\alpha(A,a)}) = \alpha w(P^{\tilde{r}}, A, a)$$

or, equivalently, $w(P^r, A, a) = w(P^{\tilde{r}}, A, a)$.

Step 5: Here we prove that, for every (A, a) and P^r, $w(P^r, A, a) = \sum_{x\in A} w(P^r, \{x, a\}, a)$. To do this, we prove that for any two menus A_1, A_2 such that $A_1 \cap A_2 = \{a\}$ and $A_1 \cup A_2 = A$, it is the case that $w(P^r, A, a) = w(P^r, A_1, a) + w(P^r, A_2, a)$. The recursive application of this idea, given the finiteness of X, concludes the step. Again, there exists a sufficiently small α such that $r^{\alpha(A,a)}$, $r^{\alpha(A_1,a)}$, and $r^{\alpha(A_2,a)}$ all belong to the convex hull of r. By steps 1 and 2 and DC, it is the case that

$$\begin{aligned}\alpha w(P^r, A, a) &= I(r^{\alpha(A,a)}) = I(r^{\alpha(A_1,a)}_{\alpha(A_2,a)})\\ &= \alpha w(P^r, A_1, a) + \alpha w(P^r, A_2, a),\end{aligned}$$

which implies $w(P^r, A, a) = w(P^r, A_1, a) + w(P^r, A_2, a)$.

Step 6: Here we prove that $w(P^r, \{x, y\}, y) = w(P^{\tilde{r}}, \{z, t\}, t)$ holds for every $x, y, z, t \in X$ and every pair P^r and $P^{\tilde{r}}$ such that the ranking of x (respectively, of y) in P^r is the same as the ranking of z (respectively, of t) in $P^{\tilde{r}}$. Consider the bijection $\sigma : X \rightarrow X$, which assigns, to the alternative ranked at s in P^r, the alternative ranked at s in $P^{\tilde{r}}$. Then, it is $\sigma(x) = z$ and $\sigma(y) = t$ and also, $\sigma(r) = \tilde{r}$. There exists a sufficiently small α such that $r^{\alpha(\{x,y\},y)}$ belongs to the convex hull of r and $\tilde{r}^{\alpha(\{z,t\},t)}$ belongs to the convex hull of $\tilde{r}$. By steps 1 and 2 and NEU, we have that

$$\begin{aligned}\alpha w(P^r, \{x, y\}, y) &= I(r^{\alpha(\{x,y\},y)}) = I(\sigma(r^{\alpha(\{x,y\},y)})) = I(\tilde{r}^{\alpha(\{z,t\},t)})\\ &= \alpha w(P^{\tilde{r}}, \{z, t\}, t),\end{aligned}$$

that is, $w(P^r, \{x, y\}, y) = w(P^{\tilde{r}}, \{z, t\}, t)$.

Step 7: We finally prove that I is a positive scalar transformation of the swaps index. Let P and P' be any two preferences and consider $x, y, z, t \in X$ with xPy and $zP't$. Thanks to step 4, consider without loss of generality that x and y (respectively, z and t) are the first two elements of P (respectively, P'). Steps 1 and 6 guarantee that $w(P, \{x, y\}, y) = w(P', \{z, t\}, t) > 0$ and steps 3 and 5 lead to

$$\begin{aligned}I(f) &= \min_P \sum_{(A,a)} f(A, a) w(P, A, a)\\ &= \min_P \sum_{(A,a)} f(A, a) \sum_{x\in A: xPa} w(P, \{x, a\}, a)\\ &= K \min_P \sum_{(A,a)} f(A, a) \mid \{x \in A : xPa\} \mid,\end{aligned}$$

TABLE 1
SUMMARY OF THE RELATIONSHIP BETWEEN AXIOMS AND INCONSISTENCY INDICES

	CONT	RAT	CONC	PWL	OC	NEU	DC
I_S	✓	✓	✓	✓	✓	✓	✓
I_{HM}	✓	✓	✓	✓	✓	✓	X
I_V	✓	✓	✓	✓	✓	X	X
I_W	✓	✓	✓[a]	X	X	✓	X
I_{W-MP}	✓	✓	✓[a]	X	X	X	X
I_A	X	✓	✓	X	X	X	X

[a] I_W and I_{W-MP} do not satisfy CONC, but a simple transformation would do. See Sec. VII for how to build the transformation in an application.

with $K > 0$, which shows that I is a positive scalar transformation of the swaps index. QED

Table 1 illustrates the structural relationship of the swaps index with the other rationality indices discussed in Section III.A. We do this by stating which axioms, among those characterizing the swaps index, they satisfy.

VI. A General Class of Indices

A. *General Weighted Index*

The swaps index relies exclusively on the endogenous information contained in the revealed choices. On occasions, however, the analyst may have more information and may wish to use it to assess the consistency of choice and identify the optimal welfare ranking. We now offer a generalization of the swaps index that is able to incorporate other information. The *general weighted index* considers 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 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.

It turns out that the general weighted index is characterized by the first four axioms used in the characterization of the swaps index.[20]

PROPOSITION 2. An inconsistency index I satisfies CONT, RAT, CONC, and PWL if and only if it is a general weighted index.

[20] The proof of this result, and all the ones that follow, can be found in App. A.

B. Nonneutral Swaps Index and Positional Swaps Index

We now present two indices from the class of general weighted indices that may be especially relevant. We start by considering settings in which the analyst has information on the nature of the alternatives, such as their monetary values, attributes, and so forth. Under these circumstances, the property of NEU may lose its appeal, since one now may wish to treat different pairs of alternatives differently, using the exogenous information that is available on them. It turns out that the remaining six properties in theorem 2 characterize a class of indices that we call the *nonneutral swaps index*. Let $w_{x,a} \in \mathbb{R}_{++}$ denote the weight of the ordered pair of alternatives x and a; that is, $w_{x,a}$ represents the cost of swapping the preferred alternative x with the chosen alternative a. Then

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

Proposition 3. An inconsistency index I satisfies CONT, RAT, CONC, PWL, OC, and DC if and only if it is a nonneutral swaps index.

Now suppose that the analyst has information on the cardinal utility values of the different alternatives, based on their position in the ranking, and wants to use it. Then, OC, which completely disregards this type of information, immediately obliterates its appeal. We show that the elimination of OC from the system of properties characterizes the following index, which we call the *positional swaps index*:

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

where $w_{i,j} \in \mathbb{R}_{++}$ denotes the weight associated with positions i and j and $\hat{x}(P)$ is the ranking of alternative x in P. Again, $w_{i,j}$ is interpreted as the cost of swapping the preferred alternative, the one that occupies position i in the ranking, with the chosen alternative, that occupies position j in the ranking.

Proposition 4. An inconsistency index I satisfies CONT, RAT, CONC, PWL, DC, and NEU if and only if it is a positional swaps index.

C. Varian and Houtman-Maks

As introduced in Section III.A, two popular measures of the consistency of behavior are due to Varian (1990) and Houtman and Maks (1985). We have already shown that these indices satisfy the properties that, by theorem 3, characterize the general weighted indices. We now provide their complete characterizations.

Let us start with the case of Varian. Its characterization requires a structure related to the search for the maximum weight in a given upper contour set. Let us then consider the following notation. For any $r \in \mathcal{R}$ and any $(A, a) \in \mathcal{O}$, denote by $\mathcal{R}^r_{(A,a)}$ all rationalizable collections $\tilde{r}$ such that (i) the top two alternatives in $P^{\tilde{r}}$ belong to A, and (ii) the top alternative in $P^{\tilde{r}}$ belongs to the strict upper contour set of a with respect to P^r.

Varian's consistency (*VC*).—For every $(A, a) \in \mathcal{O}$ and every $r \in \mathcal{R}$, it is $I(r^{\epsilon(A,a)}) = \max_{\tilde{r} \in \mathcal{R}^r_{(A,a)}} I(\tilde{r}^{\epsilon(A,z_{\tilde{r}})})$ for any sufficiently small $\epsilon > 0$, where $z_{\tilde{r}}$ is the second-best alternative according to $P^{\tilde{r}}$.[21]

Varian's consistency imposes that the inconsistency generated by a small perturbation of r in the direction of (A, a) can be related to that of perturbed collections of observations in which the inconsistency involves only the top alternative, which is ranked higher than a according to P^r. Varian's consistency is stronger than ordinal consistency because, whenever r and r' treat all the alternatives in A equally, the classes $\mathcal{R}^r_{(A,a)}$ and $\mathcal{R}^{r'}_{(A,a)}$ are the same. The following result establishes the characterization of Varian's index I_V.

PROPOSITION 5. 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-Maks's index, recalling that, in our setting, it is but a special case of Varian's index when $w^A_x = 1$ for every A and every $x \in A$. Consequently, the characterization of I_{HM} builds on that of I_V and imposes some additional structure. First, notice that I_{HM} does not discriminate between the alternatives, and hence any relabeling of the alternatives should have no effect on the level of inconsistency, thus reinstating the appeal of neutrality. However, I_{HM} requires further structure.

Houtman-Maks's composition (*HMC*).—For every $(A_1, a), (A_2, a) \in \mathcal{O}$ with $A_1 \cap A_2 = \{a\}$ and every $r \in \mathcal{R}$, $I(r^{\epsilon(A_1 \cup A_2, a)}) = \max\{I(r^{\epsilon(A_1,a)}), I(r^{\epsilon(A_2,a)})\}$ for any sufficiently small $\epsilon > 0$.

Houtman-Maks's composition establishes that, under the same conditions of disjoint composition, a small perturbation of type $(A_1 \cup A_2, a)$ is equal to the maximum of the two small perturbations that appear when breaking down the former observation into (A_1, a) and (A_2, a). We can now establish the characterization result of I_{HM}.

PROPOSITION 6. 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.

VII. An Application

In this section we use the experimental study of Harbaugh et al. (2001) to see the applicability of the swaps index.[22] The paper develops a test of

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

[22] We are very grateful to the authors for sharing all their material with us.

consistency with rationality for three different age groups: 31 7-year-old participants, 42 11-year-old participants, and 55 21-year-old participants. The experimental choice task presents the participants with 28 different bundles of two goods confronted in 11 different menus.[23] By counting the number of GARP violations, the main result is that, although violations of rationality are significantly more frequent in the youngest age group, they are present in all three age groups: 74 percent, 38 percent, and 35 percent, in the 7-, 11-, and 21-year-old groups, respectively.

We now report on I_S, together with I_{HM} and I_W. Given that the alternatives are defined by the quantities of two different goods, we compute I_S and I_{HM} by considering the set of all linear orders that satisfy quantity monotonicity. Note that $f(A, a)$ is either 1/11 or zero, given that the individuals make choices from 11 different menus. With respect to I_W, we say that there is a violation of WARP between observations (A, a) and (B, b) if there are alternatives x, y with $a \leq x \in B$ and $b \leq y \in A$. We normalize the number of WARP violations dividing them by the total number of observations.[24] The results for all 128 subjects are reported in table 2 in the online appendix. The main conclusions reached in Harbaugh et al. (2001) are reproduced here.

We now contrast I_S with the other indices. First, among the 128 subjects, 70 are rational, and clearly, I_S coincides with I_{HM}, I_W, and I_A over them, since all these indices satisfy RAT.[25] Over the remaining subjects, the Spearman's rank correlation coefficient between I_S and I_{HM} is .97, between I_S and I_W is .83, and between I_S and I_A is .51. We now illustrate the differences in the rationality judgment of the indices, by using some particular participants.

Swaps versus Houtman-Maks.—Consider individual 119.[26] It turns out that all the inconsistencies generated by this individual can be eliminated by dropping only two observations, $(A_6, (4, 1))$ and $(A_9, (2, 0))$, which leads us to $I_{HM}(f_{119}) = 2/11$. However, by focusing on the number of inconsistencies, I_{HM} disregards their severity, which can be seen to be relevant since $I_S(f_{119}) = 5/11$, which is one of the highest inconsistency levels (see app. table 2). In fact, it is easy to find other individuals with a higher I_{HM} index but still arguably less inconsistent than individual 119. One example

[23] $A_1 = \{(6, 0), (3, 1), (0, 2)\}$, $A_2 = \{(9, 0), (6, 1), (3, 2), (0, 3)\}$, $A_3 = \{(6, 0), (4, 1), (2, 2), (0, 3)\}$, $A_4 = \{(8, 0), (6, 1), (4, 2), (2, 3), (0, 4)\}$, $A_5 = \{(4, 0), (3, 1), (2, 2), (1, 3), (0, 4)\}$, $A_6 = \{(5, 0), (4, 1), (3, 2), (2, 3), (1, 4), (0, 5)\}$, $A_7 = \{(6, 0), (5, 1), (4, 2), (3, 3), (2, 4), (1, 5), (0, 6)\}$, $A_8 = \{(3, 0), (2, 2), (1, 4), (0, 6)\}$, $A_9 = \{(2, 0), (1, 3), (0, 6)\}$, $A_{10} = \{(4, 0), (3, 2), (2, 4), (1, 6), (0, 8)\}$, and $A_{11} = \{(3, 0), (2, 3), (1, 6), (0, 9)\}$.

[24] Notice that our definition in the text would divide it by 11×11 instead of 11. This normalization is vacuous when comparing the inconsistency of individuals.

[25] Notice that the computation of I_A (or I_V) would require the explicit assumption of certain weights. Harbaugh et al. (2001) provide a computation of I_A under certain assumptions regarding the budget sets. We use their computations here.

[26] The choices of the individual from menus A_1 to A_{11} are given in the following ordered vector: ((3, 1), (3, 2), (0, 3), (2, 3), (1, 3), (4, 1), (3, 3), (1, 4), (2, 0), (2, 4), (2, 3)).

is subject 60, who presents three mild inconsistencies, and $I_{HM}(f_{60}) = 3/11 = I_S(f_{60})$.[27]

Swaps versus WARP.—Individual 28, with $I_W(f_{28}) = 6/11$, represents one of the cases with the largest number of cycles.[28] However, by merely counting the number of cycles, I_W is unable to determine the number and severity of the mistakes that need to be cancelled in order to break the cycles. Closer inspection shows that this can be done by eliminating only two mild inconsistencies. This is what the swaps index does, $I_S(f_{28}) = 2/11$, with inconsistent observations $(A_4, (2, 3))$ and $(A_{11}, (3, 0))$, where, according to P_S, only (4, 2) is preferred to (2, 3) in the first and (2, 3) ranks above (3, 0) in the second. In this respect, there are a number of individuals that are classed by I_W as less inconsistent than individual 28 but whose choices are nevertheless harder to reconcile with preference maximization, and whose inconsistency values in terms of I_S are therefore higher.

Swaps versus Afriat.—Consider subject 12, who according to Afriat has a relatively low inconsistency index, $I_A(f_{12}) = .125$.[29] By considering only the largest violation and, within it, focusing on nonwelfare information, Afriat ignores (i) that individual 12 commits a relatively large number of mistakes (three, to be precise, since $I_{HM}(f_{12}) = 3/11$) and (ii) that the subject is committing relatively serious mistakes by choosing alternatives that are dominated by many others in the menu (leading to $I_S(f_{12}) = 6/11$). Once again, it is easy to find cases that are incorrectly ordered by Afriat. Consider individual 28 or 119, for example, who requires larger income adjustments but, according to I_S, fewer preference adjustments.

Appendix A

Remaining Proofs

Proof of Lemma 1

PWL guarantees that there are $|\mathcal{P}|$ pieces of $\mathcal{F}$, over every one of which the index is linear. The repeated application of linearity and CONT guarantee that the index is also linear over the convex hull of the closure of each piece. We now prove that each of these convex hulls contains one and only one collection in $\mathcal{R}$. Suppose, by contradiction, that this is not the case. Since $|\mathcal{P}| = |\mathcal{R}|$, there must exist two distinct r, r' belonging to the same convex hull. Then, PWL and RAT

[27] The choices of individual 60 are ((3, 1), (3, 2), (2, 2), (4, 2), (3, 1), (4, 1), (4, 2), (2, 2), (2, 0), (3, 2), (3, 0)).

[28] The choices of individual 28 are ((3, 1), (9, 0), (2, 2), (2, 3), (2, 2), (3, 2), (3, 3), (2, 2), (2, 0), (3, 2), (3, 0)).

[29] The choices of individual 12 are ((3, 1), (3, 2), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3), (0, 6), (1, 3), (3, 2), (3, 0)). Afriat's inconsistency is driven by the critical observation $(A_4, (2, 3))$, where (3, 2) is feasible and costs .875 times as much as the chosen element (2, 3), and in its counterpart $(A_{10}, (3, 2))$, where (2, 3) is feasible and costs .875 times as much as the chosen element.

guarantee that, for every $\alpha \in (0, 1)$, it is the case that $I(\alpha r + (1-\alpha)r') = \alpha I(r) + (1-\alpha)I(r') = 0$. However, since $r \neq r'$, there must exist at least one menu A and two distinct alternatives, a and b, such that $r(A, a) > 0$ and $r'(A, b) > 0$. Consequently, for $\alpha \in (0, 1)$, $[\alpha r + (1-\alpha)r'](A, a) > 0$ and $[\alpha r + (1-\alpha)r'](A, b) > 0$, and hence, $\alpha r + (1-\alpha)r'$ is not rationalizable. RAT implies that $I(\alpha r + (1-\alpha)r') \neq 0$, which is a contradiction. Now, given r and (A, a), the collections $r^{\epsilon(A,a)}$ converge to r as ϵ goes to zero, and since there is a finite number of hulls, CONT guarantees that these collections belong to the same convex hull as r for sufficiently small values of ϵ. In the same vein, given r and two different observations (A, a) and (B, b), the collections r and $r^{\epsilon(A,a)}_{\epsilon(B,b)}$ belong to the same convex hull for sufficiently small values of ϵ. Then, for sufficiently small perturbations $\epsilon_1 > \epsilon_2 \geq 0$, PWL guarantees that

$$I(r^{\epsilon_2(A,a)}) = I\left(\frac{\epsilon_2}{\epsilon_1} r^{\epsilon_1(A,a)} + \left(1 - \frac{\epsilon_2}{\epsilon_1}\right) r\right) = \frac{\epsilon_2}{\epsilon_1} I(r^{\epsilon_1(A,a)}) + \left(1 - \frac{\epsilon_2}{\epsilon_1}\right) I(r).$$

Under the assumption of RAT, whenever $a \neq m(P^r, A)$, this is but

$$I(r^{\epsilon_2(A,a)}) = \frac{\epsilon_2}{\epsilon_1} I(r^{\epsilon_1(A,a)}) < I(r^{\epsilon_1(A,a)}),$$

as desired. Now consider r, two different observations, (A, a) and (B, b), and a sufficiently small perturbation $\epsilon_1 > 0$. From PWL and the previous reasoning,

$$\begin{aligned} I(r^{\epsilon_1(A,a)}_{\epsilon_1(B,b)}) &= I\left(\frac{1}{2} r^{2\epsilon_1(A,a)} + \frac{1}{2} r^{2\epsilon_1(B,b)}\right) = \frac{1}{2} I(r^{2\epsilon_1(A,a)}) + \frac{1}{2} I(r^{2\epsilon_1(B,b)}) \\ &= I(r^{\epsilon_1(A,a)}) + I(r^{\epsilon_1(B,b)}). \end{aligned}$$

Whenever $a \neq m(P^r, A)$ and $b \neq m(P^r, B)$, the latter is strictly larger than $\max\{I(r^{\epsilon_1(A,a)}), I(r^{\epsilon_1(B,b)})\}$, as desired. QED

Proof of Proposition 2

Immediate from the proof of theorem 2. QED

Proof of Proposition 3

It is easy to see that nonneutral swaps indices satisfy the axioms. To prove the converse statement, we use steps 1–5 in the proof of theorem 2. By steps 1 and 5,

$$\sum_{(A,a)} f(A, a) w(P, A, a) = \sum_{(A,a)} f(A, a) \sum_{x \in A: xPa} w(P, \{x, a\}, a).$$

By steps 1 and 4, $w(P, \{x, a\}, a) > 0$ is independent of P, provided that xPa, and then we can write $\sum_{(A,a)} f(A, a) \sum_{x \in A: xPa} w_{x,a}$. Step 3 proves that the index is a nonneutral swaps index. QED

Proof of Proposition 4

It is easy to see that positional swaps indices satisfy the axioms. To prove the converse, we use the proof of theorem 2 except steps 4 and 7. By steps 1 and 5,

$$\sum_{(A,a)} f(A,a) w(P,A,a) = \sum_{(A,a)} f(A,a) \sum_{x \in A : xPa} w(P,\{x,a\},a).$$

By steps 1 and 6, $w(P, \{x, a\}, a) > 0$ depends only on the rank of alternatives x and a in P. This, together with step 3, shows that the index is a positional swaps index. QED

Proof of Proposition 5

It is easy to see that Varian's index satisfies the axioms. For the converse, we use the first three steps in the proof of theorem 2. Consider a set A, alternatives x, y, and z in A, and a pair, P^r and $P^{\bar{r}}$, such that (i) x and y are, respectively, the first- and second-best alternatives in P^r, and (ii) x and z are, respectively, the first- and second-best alternatives in $P^{\bar{r}}$. We claim that $w(P^r, A, y) = w(P^{\bar{r}}, A, z)$. There exists a sufficiently small α such that $r^{\alpha(A,y)}$ and $\bar{r}^{\alpha(A,z)}$ belong, respectively, to the convex hulls of r and $\bar{r}$. Since the upper contour sets of y in P^r and z in $P^{\bar{r}}$ are both equal to $\{x\}$, it is the case that $\mathcal{R}^r_{(A,y)} = \mathcal{R}^{\bar{r}}_{(A,z)}$, and hence, steps 1 and 2 in the proof of theorem 2 and VC imply

$$\begin{aligned} \alpha w(P^r, A, y) = I(r^{\alpha(A,y)}) = \max_{\tilde{r} \in \mathcal{R}^r_{(A,y)}} = \max_{\tilde{r} \in \mathcal{R}^{\bar{r}}_{(A,z)}} = I(\bar{r}^{\alpha(A,z)}) \\ = \alpha w(P^{\bar{r}(A,z)}). \end{aligned}$$

We then denote this value by w^A_x. Now, given (A, a) and r, there exists α sufficiently small such that $r^{\alpha(A,a)}$ belongs to the convex hull of r and for every $\tilde{r} \in \mathcal{R}^r_{(A,a)}$, $\tilde{r}^{\alpha(A,a_{\tilde{r}})}$ belongs to the convex hull of $\tilde{r}$, where $a_{\tilde{r}}$ is the second-best alternative in $P^{\tilde{r}}$. We can apply steps 1 and 2 in the proof of theorem 2 and VC to see that

$$\begin{aligned} \alpha w(P^r, A, a) = I(r^{\alpha(A,a)}) = \max_{\tilde{r} \in \mathcal{R}^r_{(A,a)}} I(\tilde{r}^{\alpha(A,a_{\tilde{r}})}) = \max_{\tilde{r} \in \mathcal{R}^r_{(A,a)}} \alpha w(P^{\tilde{r}}, A, a_{\tilde{r}}) \\ = \alpha \max_{x \in A, xP^r a} w^A_x. \end{aligned}$$

This proves that the index is Varian's index. QED

Proof of Proposition 6

Clearly, Houtman-Maks's index satisfies the axioms. To see the converse, from the proof of proposition 5, we now show that for every menu A and any pair of alternatives x and y belonging to A, it is the case that $w^A_x = w^{\{x,y\}}_x$. To see this, consider any r such that x is the top alternative in P^r and $y \in A$ is the second top alternative in P^r. For a sufficiently small α, it is the case that $r^{\alpha(A,y)}$, $r^{\alpha(\{x,y\},y)}$, and

$r^{\alpha(A\setminus\{y\},y)}$ all belong to the convex hull of r. By steps 1 and 2 in the proof of theorem 2, HMC, and RAT, it is the case that

$$\alpha w_x^A = I(r^{\alpha(A,y)}) = \max\{I(r^{\alpha(\{x,y\},y)}, I(r^{\alpha(A\setminus\{y\},y)})\} = I(r^{\alpha(\{x,y\},y)}) = \alpha w_x^{\{x,y\}}.$$

NEU guarantees that $w_x^{\{x,y\}} = w_z^{\{z,t\}}$ for every $x, y, z, t \in X$, and given the strict positivity of these weights, the index is a scalar transformation of the Houtman-Maks index. QED

Appendix B

Uniqueness

Here we establish that almost all collections of observations have a unique swaps preference.

Proposition 7. The Lebesgue measure of the set of all collections of observations for which P_S is not unique is zero.

Proof. The set of all collections $\mathcal{F}$ is the simplex over all possible observations (A, a). Consider two different preference relations, P_i and P_j, over X. We describe the set of collections $\mathcal{F}_{ij}$, for which the number of swaps associated with preference P_i is equal to the number of swaps associated with preference P_j, that is,

$$\sum_{(A,a)} f(A,a) \mid \{x \in A : xP_i a\} \mid = \sum_{(A,a)} f(A,a) \mid \{x \in A : xP_j a\} \mid$$

or, equivalently,

$$\sum_{(A,a)} f(A,a)(\mid \{x \in A : xP_i a\} \mid - \mid \{x \in A : xP_j a\} \mid) = 0.$$

Consider the interior of the simplex $\mathcal{F}$ and notice that, since P_i and P_j are different, there exists at least one observation such that $(\mid \{x \in A : xP_i a\} \mid - \mid \{x \in A : xP_j a\} \mid) \neq 0$. Hence, the interior of $\mathcal{F}_{ij}$ is defined as the intersection of a hyperplane with the interior of the simplex $\mathcal{F}$, and consequently, $\mathcal{F}_{ij}$ has volume zero. Since there is a finite number of preferences, the set $\cup_i \cup_j \mathcal{F}_{ij}$ also measures zero. Finally, notice that the set of all collections for which P_S is not unique is contained in $\cup_i \cup_j \mathcal{F}_{ij}$ and, hence, also measures zero. QED

Proposition 7 considers all possible collections of observations, and one may wonder whether the result rests on the domain assumptions. To illustrate, consider the simple case in which we have a finite number of data points, one for each menu of alternatives. Then, obviously, the measure of all collections of observations for which P_S is not unique is no longer zero. However, as the number of alternatives grows, this measure can also be proved to go to zero.

Appendix C

Computational Considerations

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 (1985) is equivalent to a well-known problem in the computer science literature, namely, the minimum set covering problem. Smeulders et al. (2012) relate Varian's and Houtman-Maks's indices to the independent set problem. Thus, one can draw from a wide range of algorithms developed by the operations research literature to solve these potential problems in the computation of the desired index.

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 some of an economic nature, a particular example being the triangularization of input-output matrices for examining the hierarchical structures of the productive sectors in an economy (see Korte and Oberhofer 1970; Fukui 1986). 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}, involves 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 swaps preference are equivalent.

PROPOSITION 8.

1. For every $f \in \mathcal{F}$, one can define a LOP with vertices in X, the solution of which provides the swaps preference.
2. For every LOP with vertices in X, one can define an $f \in \mathcal{F}$ in which the swaps preference provides the solution to the LOP.

Proof. For the first part, consider the collection 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

$$\begin{aligned}\sum_{\pi(x)<\pi(y)} c_{xy} &= \sum_{\pi(x)<\pi(y)} \sum_{(A,y):x\in A} f(A, y)\\ &= \sum_{(A,y)} f(A, y)\mid \{x \in A : \pi(x) < \pi(y)\}\mid,\end{aligned}$$

and hence, by solving the LOP, we obtain the swaps preference. 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 f such that $f(\{x, y\}, y) = c_{xy}/c$ and 0 otherwise. Since f is defined only over binary problems,

$$\begin{aligned}\sum_{(A,a)} f(A, a)\mid \{x \in A : \pi(x) < \pi(a)\}\mid &= \sum_{(\{x,y\},y):\pi(x)<\pi(y)} f(\{x, y\}, y)\\ &= \sum_{\pi(x)<\pi(y)} c_{xy},\end{aligned}$$

as desired. QED

Proposition 8 enables the techniques offered by the literature for the solution of the LOP to be used directly in the computation of the swaps preference. These techniques involve an ample array of algorithms for finding the globally optimal solution.[30] Alternatively, the literature also offers methods, which, while not com-

[30] 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.

puting the globally optimal solution, are much lighter in computational intensity and provide good approximations.[31]

References

Afriat, Sydney N. 1973. "On a System of Inequalities in Demand Analysis: An Extension of the Classical Method." *Internat. Econ. Rev.* 14:460–72.

Baldiga, Katherine A., and Jerry R. Green. 2013. "Assent-Maximizing Social Choice." *Soc. Choice and Welfare* 40:439–60.

Bernheim, B. Douglas, and Antonio Rangel. 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *Q.J.E.* 124: 51–104.

Bossert, Walter, and Yves Sprumont. 2003. "Efficient and Non-deteriorating Choice." *Math. Soc. Sci.* 45:131–42.

Brusco, Michael J., Hans-Friedrich Kohn, and Stephanie Stahl. 2008. "Heuristic Implementation of Dynamic Programming for Matrix Permutation Problems in Combinatorial Data Analysis." *Psychometrika* 73:503–22.

Chambers, Christopher P., and Takashi Hayashi. 2012. "Choice and Individual Welfare." *J. Econ. Theory* 147 (5): 1818–49.

Chaovalitwongse, Wanpracha A., Carlos A. S. Oliveira, Bruno Chiarini, Panos M. Pardalos, and Mauricio G. C. Resende. 2011. "Revised GRASP with Path-Relinking for the Linear Ordering Problem." *J. Combinatorial Optimization* 22: 572–93.

Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman. 2014. "Who Is (More) Rational?" *A.E.R.* 104:1518–50.

Dean, Mark, and Daniel Martin. 2012. "Testing for Rationality with Consumption Data." Manuscript, Brown Univ.

Echenique, Federico, Sangmok Lee, and Matthew Shum. 2011. "The Money Pump as a Measure of Revealed Preference Violations." *J.P.E.* 119 (6): 1201–23.

Famulari, Melissa. 1995. "A Household-Based, Nonparametric Test of Demand Theory." *Rev. Econ. and Statis.* 77:372–83.

Fudenberg, Drew, Ryota Iijima, and Tomasz Strzalecki. 2014. "Stochastic Choice and Revealed Perturbed Utility." Manuscript, Harvard Univ.

Fukui, Yukio. 1986. "A More Powerful Method for Triangularizing Input-Output Matrices and the Similarity of Production Structures." *Econometrica* 54:1425–33.

Green, Jerry R., and Daniel A. Hojman. 2009. "Choice, Rationality and Welfare Measurement." Manuscript, Harvard Univ.

Grötschel, Martin, Michael Jünger, and Gerhard Reinelt. 1984. "A Cutting Plane Algorithm for the Linear Ordering Problem." *Operations Res.* 2:1195–1220.

Gul, Faruk, Paulo Natenzon, and Wolfgang Pesendorfer. 2014. "Random Choice as Behavioral Optimization." *Econometrica* 82:1873–1912.

Halevy, Yoram, Dotan Persitz, and Lanny Zrill. 2012. "Parametric Recoverability of Preferences." Manuscript, Univ. British Columbia.

Harbaugh, William T., Kate Krause, and Timothy R. Berry. 2001. "GARP for Kids: On the Development of Rational Choice Behavior." *A.E.R.* 91:1539–45.

Harless, David, and Colin F. Camerer. 1994. "The Predictive Utility of Generalized Expected Utility Theories." *Econometrica* 62:1251–89.

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

Houtman, Martijn, and J. A. H. Maks. 1985. "Determining All Maximal Data Subsets Consistent with Revealed Preference." *Kwantitatieve Methoden* 19:89–104.

Iyengar, Sheena S., and Mark R. Lepper. 2000. "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?" *J. Personality and Soc. Psychology* 79:995–1006.

Kalai, Gil, Ariel Rubinstein, and Ran Spiegler. 2002. "Rationalizing Choice Functions by Multiple Rationales." *Econometrica* 70:2481–88.

Korte, Bernhard, and Walter Oberhofer. 1970. "Triangularizing Input-Output Matrices and the Structure of Production." *European Econ. Rev.* 1:482–511.

Luce, Robert D. 1959. *Individual Choice Behavior: A Theoretical Analysis.* New York: Wiley.

Manzini, Paola, and Marco Mariotti. 2007. "Sequentially Rationalizable Choice." *A.E.R.* 97:477–81.

———. 2012. "Categorize Then Choose: Boundedly Rational Choice and Welfare." *J. European Econ. Assoc.* 10 (5): 1141–65.

———. 2014. "Stochastic Choice and Consideration Sets." *Econometrica* 82:1153–76.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay. 2012. "Revealed Attention." *A.E.R.* 102 (5): 2183–2205.

Masatlioglu, Yusufcan, and Efe A. Ok. 2005. "Rational Choice with Status Quo Bias." *J. Econ. Theory* 121:1–29.

———. 2014. "A Canonical Model of Choice with Initial Endowments." *Rev. Econ. Studies* 81 (2): 851–83.

Mattsson, Lars-Göran, and Jörgen W. Weibull. 2002. "Probabilistic Choice and Procedurally Bounded Rationality." *Games and Econ. Behavior* 41:61–78.

May, Kenneth O. 1954. "Intransitivity, Utility, and the Aggregation of Preference Patterns." *Econometrica* 22:1–13.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers of Econometrics*, edited by Paul Zarembka. New York: Academic Press.

Myerson, Roger B. 1978. "Refinements of the Nash Equilibrium Concept." *Internat. J. Game Theory* 7:73–80.

Nishimura, Hiroki. 2014. "The Transitive Core: Inference of Welfare from Nontransitive Preference Relations." Manuscript, New York Univ.

Ok, Efe A., Pietro Ortoleva, and Gil Riella. 2012. "Revealed (P)reference Theory." Manuscript, New York Univ.

Palacios-Huerta, Ignacio, and Oscar Volij. 2004. "The Measurement of Intellectual Influence." *Econometrica* 72 (3): 963–77.

Rubinstein, Ariel. 1980. "Ranking the Participants in a Tournament." *SIAM J. Appl. Math.* 38:108–11.

Rubinstein, Ariel, and Yuval Salant. 2012. "Eliciting Welfare Preferences from Behavioral Datasets." *Rev. Econ. Studies* 79:375–87.

Salant, Yuval, and Ariel Rubinstein. 2008. "(A, f): Choice with Frames." *Rev. Econ. Studies* 75:1287–96.

Selten, Reinhard. 1975. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games." *Internat. J. Game Theory* 4:25–55.

Smeulders, Bart, Laurens Cherchye, Bram De Rock, and Frits C. R. Spieksma. 2012. "Goodness of Fit Measures for Revealed Preference Tests: Complexity Results and Algorithms." Manuscript, Univ. Leuven.

Swofford, James L., and Gerald A. Whitney. 1987. "Nonparametric Test of Utility Maximization and Weak Separability for Consumption, Leisure and Money." *Rev. Econ. and Statis.* 69:458–64.

Thaler, Richard. 1980. "Toward a Positive Theory of Consumer Behavior." *J. Econ. Behavior and Org.* 1:39–60.
Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211:453–58.
Varian, Hal R. 1990. "Goodness-of-Fit in Optimizing Models." *J. Econometrics* 46: 125–40.
Wilcox, Nathaniel T. 2011. "Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice under Risk." *J. Econometrics* 162:89–104.
Xu, Yongsheng, and Lin Zhou. 2007. "Rationalizability of Choice Functions by Game Trees." *J. Econ. Theory* 134:548–56.