fedea

Fundación de
Estudios de
Economía Aplicada

Do we agree? Measuring the cohesiveness of preferences

by

Jorge Alcalde-Unzu^{*}

Marc Vorsatz^{}**

Documento de Trabajo 2010-23

September 2010

* University of the Basque Country.

** EDEA.

Los Documentos de Trabajo se distribuyen gratuitamente a las Universidades e Instituciones de Investigación que lo solicitan. No obstante están disponibles en texto completo a través de Internet: <http://www.fedea.es>.

These Working Paper are distributed free of charge to University Department and other Research Centres. They are also available through Internet: <http://www.fedea.es>.

ISSN:1696-750X

Do we agree? Measuring the cohesiveness of preferences*

Jorge Alcalde-Unzu[†] and Marc Vorsatz[‡]

February 10, 2010

Abstract

In this paper, we suggest new ways of how to measure the similarity of preferences in a group of individuals. For simplicity, we refer to this as the *cohesiveness* (of preferences). We propose axioms a cohesiveness measure should satisfy and show that these properties fully characterize a family of measures. According to it, the similarities between each pair of objects in a preference profile —calculated as the proportion margin by which one objects wins against the other in a pairwise comparison— are aggregated by a weighted mean. The weight of each pair of objects depends on their importance at the social level.

Keywords: Cohesiveness, characterization, distance, measure, similarity, preferences.

JEL-Numbers: D39, D70.

*We are very grateful to the audiences in Barcelona, Bilbao, Montreal, Soria, and Vigo. We specially thank Katherine Baldiga, Miguel A. Ballester, Casilda Lasso de la Vega, Conchita D'Ambrosio, José Luis García-Lapresta, Jerry Green, Jordi Massó, Juan Moreno-Ternero, Jorge Nieto, David Pérez-Roman, and Ana Urrutia for many fruitful discussions.

[†]Corresponding author. Department of Applied Economics IV, University of the Basque Country, Avenida Lehendakari Agirre 83. 48015 Bilbao, Spain. Email: jorge.alcalde@ehu.es. Financial support from the Spanish Ministry of Education and Science, through the Juan de la Cierva program and the projects ECO2009-11213 and ECO2009-12836, is gratefully acknowledged.

[‡]Fundación de Estudios de Economía Aplicada (FEDEA), Calle Jorge Juan 46, 28001 Madrid, Spain. Email: mvorsatz@fedea.es. Financial support from the Spanish Ministry of Education and Science, through the Ramón y Cajal program and the project ECO2009-07530, is gratefully acknowledged.

1 Introduction

Motivation The cohesion in a group depends, among many other factors, on how similar its members are, where the similarity can be evaluated using external characteristics such as age, ethnicity, income, or the access to opportunities and internal characteristics such as values, attitudes, or preferences. Since it has been found in several studies that some aspects of group cohesiveness like income disparity and ethnic polarization affect important socio-economic variables such as growth and the probability of conflicts (see Alesina and Rodrik [2], Persson and Tabellini [22], and Montalvo and Reynol-Querol [21]), it is no surprise that a significant body of economic literature aims at constructing the measures used in empirical research to measure the similarity of a group of individuals in each dimension from a normative point of view. For example, Gini [14] and Atkinson [3] propose measures of income inequality (also see Cowell [5] for an overview of this literature); Esteban and Ray [12] and Duclos et al. [9] suggest ways to assess income, ethnic, or religious polarization in a society; Hutchens [16] and Echenique and Fryer [10] propose measures for quantifying the segregation of individual occupations or geographical districts; and, finally, Kranich [20] and Roemer [23] evaluate the distribution of opportunities among individuals.

The objective of this paper is to address the question of how to measure the similarity of individual preferences (or opinions) on a set of objects (candidates, alternatives, *etc.*). For simplicity, we will call this the *cohesiveness*. In order to measure the similarity of or, in dual terms, the distance between a group of individuals in any dimension, one intuitive approach is to calculate the similarity between each pair of individuals and then aggregate these numbers by taking, for example, the mean. When considering this approach, the following two issues have to be considered: (i) Which is the appropriate similarity measure for pairs of individuals? And (ii), is all relevant information of the social cohesiveness in this dimension

taken into account and aggregated correctly?

For the case when preferences can be represented by means of an ordinal ranking, many distance measures for pairs of individuals have been studied. The most important one is probably the Kemeny [18] distance that calculates the proportion of pairwise comparisons the two rankings do not coincide upon. This idea has been followed up by Kendall [17] who suggests to measure the cohesiveness of any two rankings by one minus the Kemeny distance; that is, the number of pairwise comparisons the two individuals agree upon. This measure is usually referred to in the literature as Kendall's τ . Alternative distance measures for the case of two individuals have been proposed, for example, by Cook and Seiford [6, 7] and, more recently, by Klamler [19]. The former authors suggest to sum the differences in the Borda score of each object in the two rankings and divide the resulting number by the maximal possible value of this sum. Klamler [19], on the other hand, considers all possible subsets (with a size of at least two) generated by the set of objects and calculates the proportion of different maximal objects between the two rankings. Hence, in this study equal weight is put on all possible subsets (or, to say it differently, all subsets are equally likely *ex ante*), an assumption that has recently been relaxed by Baldiga and Green [4].¹

With respect to point (ii) mentioned above, it is indeed true that the usual approach (and, probably, the most intuitive one at a first sight) consists in selecting a similarity measure between two rankings of the ones cited above together with an aggregator (usually the mean). For example, the average tau ($\bar{\tau}$) introduced by Hays [15] calculates first Kendall's tau for all possible pairs of individuals. The cohesiveness is then determined by taking the average of all those numbers.²

¹The corresponding similarity measures can be constructed in the same straightforward way as Kendall's τ is constructed from the Kemeny distance.

²See Alcalde-Unzu and Vorsatz [1] for a related axiomatic characterization. See also García-Lapresta and Pérez-Roman [13] for an overview of cohesiveness measures constructed in this way.

However, we will argue in this paper that selecting a measure defined over two individuals together with the mean aggregator seems too restrictive as an approach as it leaves important information that is only contained in the whole preference profile unconsidered. The following example will illustrate this point.

Example 1: Consider the society $N = \{1, 2, 3\}$ and the set of objects $K = \{x, y, z\}$. The following table presents two strict preference profiles P_N (to the left) and $\bar{P}_N$ (to the right) and we would like to know how they should be ordered in terms of their cohesiveness.

P_1	P_2	P_3	$\bar{P}_1$	$\bar{P}_2$	$\bar{P}_3$
x	x	y	x	x	y
y	y	z	y	z	x
z	z	x	z	y	z

The two preference profiles are constructed in such a way that $\bar{P}_N$ can be obtained from P_N by inverting the binary relation yP_2z for individual 2 and the binary relation zP_3x for individual 3. The relevant question is therefore whether the cohesiveness is higher if y is better than z for individual 2 and x is worse than z for individual 3 or if y is worse than z for individual 2 and x is better than z for individual 3. We argue that this double inversion from P_N to $\bar{P}_N$ should increase the cohesiveness. One way to see this starts from the observation that the preference profiles restricted to the individuals 2 and 3 are identical under joint permutations of objects and individuals (with the roles of x , y and z in $P_{\{2,3\}}$ played by y , x and z , respectively, in $\bar{P}_{\{2,3\}}$). Hence, the restricted profiles should naturally have the same cohesiveness. Adding individual 1 with the preference xP_1y to the restricted profiles should then imply that the cohesiveness is higher in $\bar{P}_N$ than in P_N .

However, if the mean of the distances between all pairs of individuals is applied to this example, we obtain that the cohesiveness is the same for both preference profiles independently of which of the formerly mentioned distance measures is

applied. This is because when objects x and y are compared with object z , it is not taken into account whether x or y is more accepted by individual 1.³ $\square$

The conclusion that can be drawn from this example is that a cohesiveness measure has to be constructed focusing on the preference profile as a whole (and not on pairs of individuals alone) and that a weakly higher weight has to be given to objects that are more important at the social level. The axiomatic model we present next is motivated by these two points.

Model We approach the problem of measuring the cohesiveness of preferences in a group from a general point of view. In particular, we assume that each group member has a linear ranking that reflects her/his preferences on a universal set of objects. A group of individuals is represented by a preference profile, a list of individual preferences. With these primitives at hand, we aim at constructing measures that assign to each preference profile of every possible society and every possible subset of objects a value from the unit interval, with the implicit assumption that higher values imply more cohesiveness.

Following an axiomatic approach, we propose a set of properties a cohesiveness measure should meet. The first property, *Replication Invariance*, states that if a group of individuals together with their preferences is copied a number of times, then the cohesiveness should not change. Second, *Neutrality* says that the measure is not biased towards certain objects; that is, relabeling objects does not affect the

³For both preference profiles, the cohesiveness is $\frac{5}{9}$ under $\bar{\tau}$, $\frac{1}{3}$ if the distance of Cook and Seiford [6, 7] is selected, $\frac{1}{2}$ if the choice-based distance of Klamler [19] is considered, and $\frac{3+2(1-\mu_3)}{9}$ with the generalization of Baldiga and Green [4] (where μ_3 and $\frac{1-\mu_3}{3}$ are the probabilities that the set of three objects and, respectively, any set of two objects occurs). Obviously, one can think of variations of these measures that would assign a higher cohesiveness to $\bar{P}_N$. For example, one could modify the approach of Cook and Seiford [6, 7] by allowing for scoring rules different from the Borda count or one could allow for different probabilities for subsets of the same size, but these measures naturally fail in many other examples.

cohesiveness. Third, *Monotonicity* states that if more individuals prefer x to y than the other way around and one of the individuals who prefers y to x changes her/his opinion only on this pair, then the cohesiveness increases. Fourth, *Ranking* is a key property in our analysis as it requires that a cohesiveness measure should be more sensitive to the inversion of binary relations whenever the objects that are involved in the change are socially more important.⁴ Since socially more important objects are on average ranked higher, this property makes it ultimately possible to differentiate between preference profiles as those presented in Example 1. Finally, *Consistency* studies the case in which one individual changes her/his opinion only on the pair $\{x, y\}$ and this change induces that a tie between these objects at the social level is broken. For these cases, the property says that the impact on the cohesiveness may be different if the tie is broken in favor of x or in favor of y . In particular, the impact depends on how much cohesiveness there is in the society between each of these objects and the remaining ones.

Our main result shows that these five properties fully characterize a family of measures. If there are only two objects, the cohesiveness measure is exactly the proportion margin by which one object is preferred to the other in the society. If there are more than two objects, the values of these proportion margins for each pair of objects are aggregated by a weighted mean. Thereby, the weights indicate how important objects are at the social level. In particular, the weights are a function that depends on the number of objects each of the two objects under consideration wins and ties against (with a tie counting half a win) according to pairwise comparisons. The measures of the family differ in the degree of this positive dependence. To be more exact, each measure is associated with a vector of non-negative real numbers (one number for each possible size of the set of objects) and

⁴Formally, the idea is that object x is socially more important than object y if x wins against more and loses against less objects than y in pairwise comparisons.

higher values imply that more weight is given to the similarity between objects that are socially more important.

Remainder We proceed as follows: In the next section, we introduce basic notation and definitions. Section 3 presents the set of axioms and the characterization result. Finally, we conclude. The proofs of the results are relegated to the Appendix.

2 Notation and Definitions

Consider an infinite universal set of objects X and an infinite set of individuals equal to the set of natural numbers $\mathbb{N}$.⁵ Elements of X are usually denoted by x , y , and z ; generic individuals are indexed by i and j . The main objective of this work is to evaluate the cohesiveness for all finite *societies* $N \subset \mathbb{N}$ of size $n \geq 2$ and for all finite subsets $K \subset X$ of size $k \geq 2$.⁶ That is, we aim at determining the cohesiveness in all finite non-trivial *situations* (N, K) .

Let P_i be the *strict preference relation* of individual i on X . Formally, P_i is a complete, transitive, and antisymmetric binary relation on X . The set of all strict preference relations on X is denoted by $\mathcal{P}$. Given a strict preference relation $P_i \in \mathcal{P}$ and a subset $K \subset X$, the preference relation P_i^K is obtained by restricting P_i to K . Obviously, P_i^K is a complete, transitive, and antisymmetric binary relation on K . The set of all K -restricted preference relations is denoted by $\mathcal{P}^K$.

A *profile* $P = (P_i)_{i \in \mathbb{N}} \in \mathcal{P}^{\mathbb{N}}$ is a list of individual preference relations. Given a situation (N, K) and a profile $P \in \mathcal{P}^{\mathbb{N}}$, a *preference profile* $P_N^K = (P_i^K)_{i \in N} \in (\mathcal{P}^K)^N$ is an n -tuple of K -restricted preference relations, one for every individual belonging

⁵We have represented the universal set of individuals by a countable sets. However, the results also hold if we assume that it is an uncountable set. With respect to X , it does not matter whether it is countable or uncountable.

⁶We will also make frequent use of the capital letters A , B , and C to denote societies. The capital letters S and T , on the other hand, are reserved to indicate subsets of X .

to N .⁷ We will say that two preference profiles $P_A^K \in (\mathcal{P}^K)^A$ and $\bar{P}_B^K \in (\mathcal{P}^K)^B$, corresponding to the societies A and B of equal size, are *isomorphic* whenever there exists a one-to-one mapping $\pi : A \rightarrow B$ such that for all $i \in A$, $\bar{P}_{\pi(i)}^K = P_i^K$.

Given a subset $K \subset X$, let $\bar{K} = \{\{x, y\} \in K^2 : x \neq y\}$ be the set of all pairs of distinct objects in K . Given the restricted profile P_N and a pair of distinct objects $\{x, y\}$, $\#(xP_N y) = \#\{i \in N : xP_i y\}$ denotes the number of individuals who prefer x to y at P_N . Then, $n_{x,y}(P_N) = \#(xP_N y) - \#(yP_N x)$ is the difference between the number of individuals who prefer x to y and the ones who prefer y to x and $\sigma_{x,y}(P_N) = \frac{|n_{x,y}(P_N)|}{n}$ is the proportion margin by which one object wins against the other in a pairwise comparison. We will say that x *wins-or-ties* against y in society N (denoted by $x \succsim_{P_N} y$) whenever $n_{x,y}(P_N) \geq 0$. The asymmetric part (the *win* relation) and the symmetric part (the *tie* relation) of $\succsim_{P_N}$ —referred to as $\succ_{P_N}$ and $\sim_{P_N}$, respectively—are defined in the natural way.

Given a preference profile P_N^K and an object $x \in K$, $r_x(P_N^K)$ is the number of objects x wins-or-ties against in N ; that is, $r_x(P_N^K) = \#\{y \in K \setminus \{x\} : x \succsim_{P_N} y\}$. Similarly, let $p_x(P_N^K) = \#\{y \in K \setminus \{x\} : x \succ_{P_N} y\}$ be the number of objects x wins against, $i_x(P_N^K) = \#\{y \in K \setminus \{x\} : x \sim_{P_N} y\}$ the number of objects x ties against, and $r_{x,y}(P_N^K) = p_x(P_N^K) + p_y(P_N^K) + \frac{1}{2}(i_x(P_N^K) + i_y(P_N^K))$ the number of objects x and y win-or-tie against when ties count only half a win. Finally, given two preference profiles $P_N^K, \bar{P}_N^K \in (\mathcal{P}^K)^N$, object $x \in K$ is *at least as important* in $\bar{P}_N^K$ as in P_N^K if $r_x(\bar{P}_N^K) \geq r_x(P_N^K)$ and $p_x(\bar{P}_N^K) \geq p_x(P_N^K)$. That is, x is at least as important in $\bar{P}_N^K$ as in P_N^K whenever the number of objects that win against x does not increase from $\bar{P}_N^K$ to P_N^K , while, at the same time, the number of objects that x wins against does not decrease from $\bar{P}_N^K$ to P_N^K .

Given the preference relation $P_i^K \in \mathcal{P}^K$ and the pair of objects $\{x, y\} \in \bar{K}$, the ordered pair (x, y) is a *contiguous pair* at P_i^K if $xP_i y$ and there is no other object

⁷If we only restrict the set of individuals to N , the preference profile will be denoted by P_N .

$z \in K$ such that xP_izP_iy . For any two preference relations $P_i^K, \bar{P}_i^K \in \mathcal{P}^K$ and any pair of objects $\{x, y\} \in \bar{K}$, $\bar{P}_i^K$ is said to be (x, y) -different from P_i^K if (y, x) is a contiguous pair at P_i^K , (x, y) is a contiguous pair at $\bar{P}_i^K$, and $zP_iw \Leftrightarrow z\bar{P}_iw$ for all $\{z, w\} \in \bar{K} \setminus \{x, y\}$. Finally, given the preference profiles $P_N^K, \bar{P}_N^K \in (\mathcal{P}^K)^N$, an individual $i \in N$, and a pair of objects $\{x, y\} \in \bar{K}$, $\bar{P}_N^K$ is said to be (x, y) -different from P_N^K for individual i if $\bar{P}_i^K$ is (x, y) -different from P_i^K and $\bar{P}_j^K = P_j^K$ for all $j \stackrel{K}{\sim} N \setminus \{i\}$. Intuitively, $\bar{P}_N^K$ is (x, y) -different from P_N^K for individual i if $\bar{P}_N^K$ can be derived from P_N^K by only reversing the binary relation yP_ix .

A *cohesiveness measure for the situation* (N, K) is a function $M^{N,K} : \mathcal{P}^N \rightarrow [0, 1]$ that assigns to every profile $P \in \mathcal{P}^N$ a real number $M^{N,K}(P)$ from the unit interval with the property that for all profiles $P, \bar{P} \in \mathcal{P}^N$ such that $P_N^K = \bar{P}_N^K, M^{N,K}(\bar{P}) = M^{N,K}(P)$. For notational purposes, we will write $M(P_N^K)$ instead of $M^{N,K}(P)$. A *cohesiveness measure* M is a family of functions $\{M^{N,K} : \mathcal{P}^N \rightarrow [0, 1]\}_{N \subset \mathbb{N}, K \subset X}$. Since we have fixed the range of the measure to the unit interval (which can be interpreted as a kind of normalization), we assume that for all subsets $K \subset X$, there are two preference profiles P_A^K and $\bar{P}_B^K$ corresponding to the possibly different societies $A, B \subset \mathbb{N}$ such that $M(P_A^K) = 0$ and $M(\bar{P}_B^K) = 1$. That is, for all subsets of objects there are preference profiles such that the extreme values of the measure are indeed reached.

3 Axioms and Characterization

We present five properties and show that they fully characterize a general class of measures. The first axiom, Replication Invariance, requires that if a society and their preferences are copied a number of times, the cohesiveness in the new society, which consists of the union of all copied groups, should be the same as in the original society. This property have the same spirit of a classical property

in the literature of income inequality measurement (see Dalton [8]). Observe that Replication Invariance implies the well-known property of anonymity, which in this context would ask the measure to be independent to permutations of individuals.

REPLICATION INVARIANCE (REP): The cohesiveness measure M satisfies *replication invariance* if for all societies $A, B \subset \mathbb{N}$, all subsets $K \subset X$, and all preference profiles $P_A^K \in (\mathcal{P}^K)^A$ and $\bar{P}_B^K \in (\mathcal{P}^K)^B$ such that $\bar{P}_B^K$ consists of the union of $m \geq 1$ disjoint isomorphic copies of P_A^K ,

$$M(\bar{P}_B^K) = M(P_A^K).$$

According to the second axiom, Neutrality, the measure should not be biased towards certain objects. Formally, the axiom requires that if it is possible to derive one preference profile from a distinct one by only relabelling objects, then the cohesiveness should be the same in both preference profiles.

NEUTRALITY (NEU): The cohesiveness measure M is *neutral* if for all societies $N \subset \mathbb{N}$, all subsets $S, T \subset X$, and all preference profiles $P_N^S \in (\mathcal{P}^S)^N$, $\bar{P}_N^T \in (\mathcal{P}^T)^N$ such that there exists a permutation $\mu : X \rightarrow X$ with the properties that $x \in S \Leftrightarrow \mu(x) \in T$ and $x P_i y \Leftrightarrow \mu(x) \bar{P}_i \mu(y)$ for all individuals $i \in N$,

$$M(\bar{P}_N^T) = M(P_N^S).$$

The third property, Monotonicity, regards the situation when there are more individuals who prefer x over y than the other way around but there is still some individual i who prefers y over x , and (y, x) is a contiguous pair of the preference ranking of this individual. If this individual i now changes her/his opinion only on the pair $\{x, y\}$, then the preferences in the group become more alike on that particular pair of objects everything else being equal. Monotonicity asks then that the cohesiveness should strictly increase.

MONOTONICITY (MON): The cohesiveness measure M is *monotone* if for all situations (N, K) , there exists an individual $i \in N$ and a pair of distinct objects $\{x, y\} \in \bar{K}$ such that for all preference profiles $P_N^K, \bar{P}_N^K \in (\mathcal{P}^K)^N$ satisfying that $\bar{P}_N^K$ is (x, y) -different from P_N^K for individual i ,

$$x \succ_{P_N} y \Rightarrow M(\bar{P}_N^K) > M(P_N^K).$$

The fourth property, Ranking, states that the effect on the cohesiveness if some individual i only changes her/his opinion on the pair $\{x, y\}$ is higher if the two objects are more important at the social level. To express this more formally, let P_N^K and $\bar{P}_N^K$ be such that x wins against y in the pairwise comparison in both preference profiles. Suppose also that (y, x) is a contiguous pair of objects for the same individual i at both preference profiles. Consider now the two preference profiles $\tilde{P}_N^K$ and $\hat{P}_N^K$ that are obtained from P_N^K and $\bar{P}_N^K$, respectively, by only changing the ordering between x and y for individual i . Now, if both x and y are at least as important in $\bar{P}_N^K$ as in P_N^K , then the change in the cohesiveness from $\bar{P}_N^K$ to $\hat{P}_N^K$ has to be at least as large as the change from P_N^K to $\tilde{P}_N^K$.

RANKING (RAN): The cohesiveness measure M satisfies *ranking* if for all situations (N, K) , there exists an individual $i \in N$ and a pair of distinct objects $\{x, y\} \in \bar{K}$ such that for all preference profiles $P_N^K, \bar{P}_N^K, \tilde{P}_N^K, \hat{P}_N^K \in (\mathcal{P}^K)^N$ satisfying that $\tilde{P}_N^K$ is (x, y) -different from P_N^K for individual i , $\hat{P}_N^K$ is (x, y) -different from $\bar{P}_N^K$ for individual i , $x \succ_{P_N} y$, $x \succ_{\bar{P}_N} y$, and both x and y are at least as important in $\bar{P}_N^K$ as in P_N^K ,

$$M(\hat{P}_N^K) - M(\bar{P}_N^K) \geq M(\tilde{P}_N^K) - M(P_N^K).$$

The last property, Consistency, expresses the idea that, in presence of more objects, it may not be equivalent to break a tie between two objects, x and y , at the social level in favor of one or the other. It should be natural that if the society

agrees more on how to order x and the remaining objects than y and the remaining objects, the cohesiveness increases more if the tie is broken in favor of x than if it is broken in favor of y . The property exactly says that if one individual changes her/his opinion only on the pair $\{x, y\}$ and this provokes that a tie between these objects at the social level is broken in favor of x , the cohesiveness may change differently to what it would have changed if there were no tie (*i.e.*, x won previously) *ceteris paribus*. This difference depends proportionally on how the society orders each of these objects and remaining ones; *i.e.*, the sum over all $M^{N, \{x, v\}} - M^{N, \{y, v\}}$, where $v \in K \setminus \{x, y\}$.

CONSISTENCY (CON): The cohesiveness measure M is *consistent* if for all situations (N, K) , there exists $t_N^K \in \mathbb{R}$, an individual $i \in N$, and a pair of distinct objects $\{x, y\} \in \bar{K}$ such that for all preference profiles $P_N^K, \bar{P}_N^K, \hat{P}_N^K, \check{P}_N^K \in (\mathcal{P}^K)^N$ satisfying that $\hat{P}_N^K$ is (x, y) -different from P_N^K for individual i , $\check{P}_N^K$ is (x, y) -different from $\bar{P}_N^K$ for individual i , $x \sim_{P_N} y, x \succ_{\bar{P}_N} y$, and $w \succsim_{P_N} z \Leftrightarrow w \succsim_{\bar{P}_N} z$ for all $\{w, z\} \in \bar{K} \setminus \{x, y\}$,

$$[M(\hat{P}_N^K) - M(P_N^K)] - [M(\check{P}_N^K) - M(\bar{P}_N^K)] = t_N^K \sum_{v \in K \setminus \{x, y\}} \left(M(P_N^{\{x, v\}}) - M(P_N^{\{y, v\}}) \right).$$

These are the five axioms used in our analysis. Next, we formally introduce the class of measures Φ that depends on the vector of non-negative real numbers $\alpha \equiv \{\alpha_k\}_{k \in \mathbb{N}, k \geq 2}$.

Definition 1 *The cohesiveness measure M is said to belong to the class Φ if there exists a vector $\{\alpha_k\}_{k \in \mathbb{N}, k \geq 2}$, with $\alpha_k \geq 0$ for all k , such that for all situations (N, K) and all profiles $P \in \mathcal{P}^{\mathbb{N}}$,*

$$M(P_N^K) = \frac{\sum_{\{x, y\} \in \bar{K}} \omega_k(r_{x, y}(P_N^K)) \cdot \sigma_{x, y}(P_N^K)}{\sum_{\{x, y\} \in \bar{K}} \omega_k(r_{x, y}(P_N^K))},$$

where for all $v \in \{1, \frac{3}{2}, 2, \frac{5}{2}, \dots, 2k-3\}$, $\omega_k(v) = 1 + 2\alpha_k(v-1)$.

The generic measure of the class Φ will be denoted by Φ_α . Any such measure can be described as follows: (i) First, one calculates for all pairs of objects the proportion margin by which one objects wins against the other in the pairwise comparison. For example, if $n = 6$ and 4 individuals prefer x to y , the proportion margin is $\frac{4-2}{6} = \frac{1}{3}$. (ii) Second, weights are assigned to every pair of objects as a function of the size of the set of objects K in such a way that if the sum of the number of objects that lose against each object of the pair is equal to a and the sum of the number of objects that ties with each object of the pair is equal to b , a weight of $(2a + b)\alpha_k + 1$ is given to that pair. So, for example, if both objects in a pair lose against all the remaining objects get an absolute weight of 1. On the other hand, if both x and y win against all the remaining objects, the absolute weight of the pair $\{x, y\}$ will be $2(2k - 4)\alpha_k + 1$. (iii) Finally, the cohesiveness is determined as the weighted average of the proportion margins determined in (i) with the absolute weights defined in (ii).

As it can be easily seen, the class Φ is parameterized by the vector α . If $\alpha_k = 0$ for all k , then all pairs of objects receive always an absolute weight of 1, independently of the set of objects K . Hence, for all situations (N, K) and all profiles $P \in \mathcal{P}^N$, $\Phi_0(P_N^K) = \frac{2}{k(k-1)} \sum_{\{x,y\} \in \bar{K}} \sigma_{x,y}(P_N^K)$.⁸ We call this measure *average sigma* and denote it by $\bar{\sigma}$.⁹ The average sigma can be considered the baseline comparison as it is the only measure belonging to Φ that does not take into account how important objects are at the social level. If $\alpha_k = \frac{1}{2}$ for all k , $\omega_k(r_{x,y}(P_N^K)) = r_{x,y}(P_N^K)$. Thus, the weight given to any pair of objects is equal to the sum of the number of objects each of the members of the pair win and tie against (where a tie counts only half a win). Finally, it should also be remarked that the higher the parameter α_k , the more weight is put on pairs of objects that are socially more important.

⁸The notation Φ_α will be used to indicate the measure Φ_α , where the vector α is such that $\alpha_k = x$ for all possible k .

⁹See Alcalde-Unzu and Vorsatz [1] for an alternative axiomatic characterization of $\bar{\sigma}$.

Theorem 1 is a full characterization of Φ in terms of the axioms introduced before.¹⁰

Theorem 1 *The cohesiveness measure M satisfies REP, NEU, MON, RAN, and CON if, and only if, it belongs to the class Φ .*

Proof: See the Appendix. □

We also show that all properties used in Theorem 1 are necessary.

Proposition 1 *The properties used in Theorem 1 are independent.*

Proof: See the Appendix. □

Finally, we conclude our main section by applying the family Φ to the example presented in the Introduction.

Example 1 (cont.): To remind the reader, we present here again the two preference profiles P_N and $\bar{P}_N$.

$\mathbf{P_1}$	$\mathbf{P_2}$	$\mathbf{P_3}$	$\mathbf{\bar{P}_1}$	$\mathbf{\bar{P}_2}$	$\mathbf{\bar{P}_3}$
x	x	y	x	x	y
y	y	z	y	z	x
z	z	x	z	y	z

Observe first that $\sigma_{x,y}(P_N) = \sigma_{x,y}(\bar{P}_N) = \frac{1}{3}$, $\sigma_{x,z}(P_N) = \sigma_{y,z}(\bar{P}_N) = \frac{1}{3}$, and $\sigma_{y,z}(P_N) = \sigma_{x,z}(\bar{P}_N) = 1$. Given any $\alpha_3 \in \mathbb{R}_+$, we obtain that $\Phi_\alpha(P_N) = \frac{(1+4\alpha_3) \cdot \frac{1}{3} + (1+2\alpha_3) \cdot \frac{1}{3} + 1 \cdot 1}{3+6\alpha_3} = \frac{5+6\alpha_3}{9+18\alpha_3}$ and $\Phi_\alpha(\bar{P}_N) = \frac{(1+4\alpha_3) \cdot \frac{1}{3} + (1+2\alpha_3) \cdot 1 + 1 \cdot \frac{1}{3}}{3+6\alpha_3} = \frac{5}{9}$.

¹⁰One comment regarding the characterized family is at stake. As we have explained above, every measure belonging to Φ is parametrized by a vector α and, a priori, there is no connection between two parameters that correspond to subsets of objects of different sizes. One intuitive approach would therefore be to consider additional properties that put restrictions on the vector α (*i.e.*, in such a way that α_k is independent of k). However, it turned out that it is difficult to come up with axioms that naturally connect the parameters for subsets of objects of different sizes. Consequently, we decided to present the family in its most general form.

So, $\Phi_\alpha(P_N) = \Phi_\alpha(\bar{P}_N)$ only if $\alpha_3 = 0$. If $\alpha_3 > 0$, we get (as desired) that $\Phi_\alpha(P_N) < \Phi_\alpha(\bar{P}_N)$. $\square$

4 Conclusion

Group cohesiveness can be defined as the degree to which the members of the group stick together, both emotional and task-related. If we think, for example, of an organizational unit at the workplace, group cohesiveness plays a central role not only because it leads to a better performance (*i.e.*, there is a higher likelihood that a goal is met and/or the quality of the output is higher), but also because it decreases the probability of internal conflicts.

Eisenberg [11] reviewed various aspects affecting group cohesiveness; the most important ones are the similarity of the members of the group (a higher degree of similarity leads to more trust), the group size (social norms are easier to maintain in smaller groups), the entry difficulty (forming part of an exclusive group has a high emotional value), the group success, and the existence of external competition and threats (more competition makes the individuals focus more on the joint goals). Among these factors, the similarity of the group members—one distinguishes here between *external* variables such as age, gender, or religion and *internal* variables such as values and attitudes—is often regarded a key variable.

The objective of this study was to measure the similarity of individuals in a group on one particular dimension: Preferences. To do so, we assumed the existence of an unstructured set of objects and supposed all individuals to have strict preferences on this set that can be represented by an ordinal ranking. This abstract setting allows our study to be applicable in a variety of instances like decision problems in organizations where objects could be candidates (the problem of electing representatives) or alternatives (the problem of choosing among different investment

strategies) or more leisure related activities where objects could represent tastes about theater, sports, or arts.

Within this general framework, we proposed five axioms and showed that these properties fully characterize a family of (cohesiveness) measures. In particular, the axioms imply that if there are only two objects, then the cohesiveness has to be calculated as the proportion margin by which one object is preferred to the other in the society. If there are more than two objects, these proportion margins are calculated for each pair of objects and are then weighted by a function that depends positively on the number of remaining objects each of the two objects wins and ties against according to a pairwise comparison (with a tie counting half a win). Consequently, this positive dependence takes explicitly into account how important objects are at the social level, which we believe to be a considerable innovation with respect to the existing literature.

Finally, we would like to indicate one direction for future research that seems promising. In the current study, groups are *exogenously* given, which is a natural assumption in many instances. However, it is also true that workers in a firm look for a new employer if their job satisfaction is low. Also, one can generally choose freely whether or not to go to the movies with a group of people. Consequently, we would like to extend our model by allowing for *endogenous group formation* and propose cohesiveness measures for the situation when various groups co-exists. In that approach one will have to carefully balance intra-group cohesiveness with inter-group distances at the advantage of being able to make predictions about which kind of groups are more likely to form.

Appendix

Proof of Theorem 1

It is easy to see that the measures belonging to Φ satisfy REP, NEU, MON, RAN, and CON. Hence, we concentrate on showing that these five properties indeed imply the measure to be in Φ . The proof will be developed in three main parts:

1. It will be shown that if K only consists of two objects, then the cohesiveness has to be measured by $\sigma_{x,y}(P_N)$, where $\{x, y\} = K$.
2. It will be shown that if K consists of any arbitrary number of objects, the difference in cohesiveness between two preference profiles of the same situation (N, K) , P_N^K and $\bar{P}_N^K$, has to take the form of the difference between the formulae of any cohesiveness measure of the family Φ .
3. It will be shown that the minimal cohesiveness is always attained when $\sigma_{x,y} = 0$ for all $\{x, y\} \in \bar{K}$. Plugging this result back into the former part, yields the theorem.

Before starting, we need to introduce the following remark:

Remark 1 *One can see that REP, NEU and MON imply the following stronger version of Monotonicity: The cohesiveness measure M is monotone if for all situations (N, K) , all individuals $i \in N$, all pairs of distinct objects $\{x, y\} \in \bar{K}$, and all preference profiles $P_N^K, \bar{P}_N^K \in (\mathcal{P}^K)^N$ such that $\bar{P}_N^K$ is (x, y) -different from P_N^K for individual i ,*

$$x \succ_{P_N} y \Rightarrow M(\bar{P}_N^K) > M(P_N^K).$$

In the rest of the proof, when we refer to MON we will be applying this stronger version. Similar stronger versions of Ranking and Consistency can be deduced from the application of REP and NEU jointly with the original properties RAN and CON, respectively. We will always refer to these stronger properties in the rest of the proof.

Part 1: We show first in Lemma 1 that the cohesiveness for a subset of objects $K = \{x, y\}$ is zero if and only if $x \sim_{P_N} y$. Afterwards, in Lemma 2, we establish that the cohesiveness is maximal if and only if all individuals in the society agree on how to order x and y . Finally, Proposition 2 establishes that for all preference profiles $P_N^{\{x,y\}}$, $M(P_N^{\{x,y\}}) = \sigma_{x,y}(P_N)$.

Lemma 1 *Suppose that the cohesiveness measure M satisfies REP, NEU and MON. Then, for all situations $(N, \{x, y\})$ and all profiles $P \in \mathcal{P}^N$, $M(P_N^{\{x,y\}}) = 0$ if, and only if, $x \sim_{P_N} y$.*

Proof: The proof of the lemma is divided into two parts.

$\Rightarrow$ Suppose otherwise; that is, there is some situation $(N, \{x, y\})$ and some profile

$P \in \mathcal{P}^N$ such that $\neg(x \sim_{P_N} y)$ and $M(P_N^{\{x,y\}}) = 0$. Suppose first that $x \succ_{P_N} y$. Consider a society A (of size $a = 3n$) and a preference profile $\bar{P}_A^{\{x,y\}}$ such that $\bar{P}_A^{\{x,y\}}$ consists of the union of three isomorphic and disjoint copies of $P_N^{\{x,y\}}$. Since $x \succ_{P_N} y$ is equivalent to $n_{x,y}(P_N) > 0$ by definition, it must be the case that $n_{x,y}(\bar{P}_A) \geq 3$. Now, let the preference profile $\tilde{P}_A^{\{x,y\}}$ be such that $\bar{P}_A^{\{x,y\}}$ is (x, y) -different from $\tilde{P}_A^{\{x,y\}}$ for some individual $i \in A$. We have that $n_{\{x,y\}}(\tilde{P}_A) > 0$ and, therefore, $M(\bar{P}_A^{\{x,y\}}) > M(\tilde{P}_A^{\{x,y\}})$ by MON. Since $M(\bar{P}_A^{\{x,y\}}) = M(P_N^{\{x,y\}})$ by REP and $M(P_N^{\{x,y\}}) = 0$ by assumption, it follows that $M(\tilde{P}_A^{\{x,y\}}) < 0$. This contradicts the definition of M . If $y \succ_{P_N} x$, a similar reasoning applies.

$\Leftarrow$ Take any situation $(N, \{x, y\})$ and any profile $P \in \mathcal{P}^N$ with the property that

$x \sim_{P_N} y$. By the definition of M , there is some society B (of size b) and some preference profile $\bar{P}_B^{\{x,y\}}$ for which $M(\bar{P}_B^{\{x,y\}}) = 0$. By the first part of this proof, $x \sim_{\bar{P}_B} y$. Consider now the society C (of size $c = b \cdot n$) and the preference profile $\tilde{P}_C^{\{x,y\}}$ which consists of the union of n isomorphic and disjoint copies of $\bar{P}_B^{\{x,y\}}$. Then, $M(\tilde{P}_C^{\{x,y\}}) = M(\bar{P}_B^{\{x,y\}}) = 0$ by REP. Since

$n_{x,y}(\tilde{P}_C) = n_{x,y}(P_N) = 0$ by construction, $\tilde{P}_C^{\{x,y\}}$ consists of the union of b isomorphic and disjoint copies of $P_N^{\{x,y\}}$. Hence, $M(\tilde{P}_C^{\{x,y\}}) = M(P_N^{\{x,y\}})$ by REP. Thus, $M(P_N^{\{x,y\}}) = 0$.

This concludes the proof of the lemma. $\square$

Lemma 2 *Suppose that the cohesiveness measure M satisfies REP, NEU, and MON. Then, for all situations $(N, \{x, y\})$ and all profiles $P \in \mathcal{P}^N$, $M(P_N^{\{x,y\}}) = 1$ if, and only if, $|n_{x,y}(P_N)| = n$.*

Proof: The proof of the lemma is divided into two parts.

$\Rightarrow$ Suppose otherwise; that is, there is some situation $(N, \{x, y\})$ and some profile $P \in \mathcal{P}^N$ such that $|n_{x,y}(P_N)| < n$ and $M(P_N^{\{x,y\}}) = 1$. By NEU, assume without loss of generality that $0 \leq n_{x,y}(P_N) < n$. Hence, there is some individual $i \in N$ who prefers y to x while, at the same time, $x \succsim_{P_N} y$. If $x \succ_{P_N} y$, consider the preference profile $\bar{P}_N^{\{x,y\}}$ such that $\bar{P}_N^{\{x,y\}}$ is (x, y) -different from $P_N^{\{x,y\}}$ for individual i . It follows from MON that $M(\bar{P}_N^{\{x,y\}}) > M(P_N^{\{x,y\}}) = 1$. This contradicts the definition of M . On the other hand, if $x \sim_{P_N} y$, it follows from Lemma 1 that $M(P_N^{\{x,y\}}) = 0$. This contradicts that $M(P_N^{\{x,y\}}) = 1$.

$\Leftarrow$ Take any situation $(N, \{x, y\})$ and any profile $P \in \mathcal{P}^N$ with the property that $|n_{x,y}(P_N)| = n$. By NEU, assume without loss of generality that $n_{x,y}(P_N) = n$. By the definition of M , there is some society B (of size b) and some preference profile $\bar{P}_B^{\{x,y\}}$ for which $M(\bar{P}_B^{\{x,y\}}) = 1$. By the first part of this proof, $\bar{P}_B^{\{x,y\}}$ must be unanimous; that is, $|n_{x,y}(\bar{P}_B)| = b$. By NEU, assume without loss of generality that $n_{x,y}(\bar{P}_B) = b$. Consider now the society C (of size $c = b \cdot n$) and the preference profile $\tilde{P}_C^{\{x,y\}}$ which consists of the union of n isomorphic and disjoint copies of $\bar{P}_B^{\{x,y\}}$. Then, $M(\tilde{P}_C^{\{x,y\}}) = M(\bar{P}_B^{\{x,y\}}) = 1$ by REP. Finally, since $\tilde{P}_C^{\{x,y\}}$ consists of the union of b isomorphic and

disjoint copies of $P_N^{\{x,y\}}$, we have that $M(\tilde{P}_C^{\{x,y\}}) = M(P_N^{\{x,y\}})$ by REP.

Hence, $M(P_N^{\{x,y\}}) = 1$.

This concludes the proof of the lemma. $\square$

Now, we are ready to provide a full characterization of the measure in case the subset only contains two objects.

Proposition 2 *If the cohesiveness measure M satisfies REP, NEU, MON, RAN, and CON then, for all situations $(N, \{x, y\})$ and all profiles $P \in \mathcal{P}^{\mathbb{N}}$,*

$$M(P_N^{\{x,y\}}) = \sigma_{x,y}(P_N).$$

Proof: Take any situation $(N, \{x, y\})$ and any profile $P \in \mathcal{P}^{\mathbb{N}}$. Since REP implies that $M^{N, \{x,y\}}$ is independent of the identity of the individuals (the classical anonymity property for a fixed society), the measure $M^{N, \{x,y\}}$ can be written as a function of the number of individuals who prefer x over y [*i.e.*, $n_{x,y}(P_N^K)$] and the number of individuals who prefer y over x [*i.e.*, $n_{y,x}(P_N^K) = n - n_{x,y}(P_N^K)$]. Given that n is fixed for $M^{N, \{x,y\}}$ and that the measure is also symmetric in the objects (by NEU), we have that $M^{N, \{x,y\}}$ only depends on $|n_{x,y}(P_N)|$. The remainder of the proof is now divided into two parts, depending on whether n is even or odd.

a. Suppose that n is even. If $n = 2$, the values of $M^{N, \{x,y\}}$ are entirely determined by the first two lemmas. Suppose that $n = 4$. Let the preference profiles $P_N^{\{x,y\}}$, $\bar{P}_N^{\{x,y\}}$, and $\hat{P}_N^{\{x,y\}}$ be such that $|n_{x,y}(P_N)| = 0$, $|n_{x,y}(\bar{P}_N)| = 2$ and $|n_{x,y}(\hat{P}_N)| = 4$. Given that that $M(P_N^{\{x,y\}}) = 0$ by Lemma 1, it follows from CON that $M(\hat{P}_N^{\{x,y\}}) = 2 \cdot M(\bar{P}_N^{\{x,y\}})$. We also know from Lemma 2 that $M(\hat{P}_N^{\{x,y\}}) = 1$. Then, we obtain that $M(\bar{P}_N^{\{x,y\}}) = \frac{1}{2}$. This concludes the proof for the case when $n = 4$.

Finally, suppose from now on that $n \geq 6$. Let the preference profiles $P_N^{\{x,y\}}$, $\bar{P}_N^{\{x,y\}}$, $\hat{P}_N^{\{x,y\}}$, and $\tilde{P}_N^{\{x,y\}}$ be such that $|n_{x,y}(P_N)| = 0$, $|n_{x,y}(\bar{P}_N)| = 2$,

$|n_{x,y}(\hat{P}_N)| = 4$, and $6 \leq |n_{x,y}(\tilde{P}_N)| = b \leq n$. We know from Lemma 1 that $M(P_N^{\{x,y\}}) = 0$. Then, it follows from CON that $M(\hat{P}_N^{\{x,y\}}) = 2 \cdot M(\bar{P}_N^{\{x,y\}})$. The iterative application of RAN implies in addition that $M(\tilde{P}_N^{\{x,y\}}) = \frac{b}{2} \cdot M(\bar{P}_N^{\{x,y\}})$. By Lemma 2, we have that if $b = n$, $M(\tilde{P}_N^{\{x,y\}}) = 1$. Therefore, we have that $M(\bar{P}_N^{\{x,y\}}) = \frac{2}{n} = \sigma_{x,y}(\bar{P}_N)$. Then, we obtain that $M(\hat{P}_N^{\{x,y\}}) = \frac{4}{n} = \sigma_{x,y}(\hat{P}_N)$ and that for any $b \leq n$, $M(\tilde{P}_N^{\{x,y\}}) = \frac{b}{n} = \sigma_{x,y}(\tilde{P}_N)$.

- b. Suppose that n is odd. Consider any society A (of size $a = 2n$) and the preference profile $\bar{P}_A^{\{x,y\}}$ that consists of the union of two isomorphic and disjoint copies of $P_N^{\{x,y\}}$. By REP, $M(\bar{P}_A^{\{x,y\}}) = M(P_N^{\{x,y\}})$. Since $|n_{x,y}(\bar{P}_A)| = 2 \cdot |n_{x,y}(P_N)|$ by construction and $M(\bar{P}_A^{\{x,y\}}) = \frac{|n_{x,y}(\bar{P}_A)|}{a}$ by the first part of the proof, it can be concluded that $M(P_N^{\{x,y\}}) = \frac{|n_{x,y}(\bar{P}_A)|}{a} = \frac{|n_{x,y}(P_N)|}{n} = \sigma_{x,y}(P_N)$.

This concludes the proof of the proposition. $\square$

Part 2: We focus now on the general case where K is allowed to be any subset of X with at least two objects. Our particular objective will be to calculate the difference in the cohesiveness between any two preference profiles of the same situation (N, K) . In the main step of the proof, we are going to show that this difference can be expressed as a weighted average over the values of σ for all possible pairs of objects, where the weight ω_k assigned to the pair $\{x, y\}$ depends on $r_{x,y}(P_N^K) = p_x(P_N^K) + p_y(P_N^K) + \frac{1}{2}(i_x(P_N^K) + i_y(P_N^K))$. Finally, we will see that $\omega_k(r_{x,y}(P_N^K)) = 1 + 2\alpha_k(r_{x,y}(P_N^K) - 1)$ for some $\alpha_k \geq 0$. To be more concrete, the objective is to prove the following proposition.

Proposition 3 *If the cohesiveness measure M satisfies REP, NEU, MON, RAN, and CON, then there exists a vector $\alpha \equiv \{\alpha_k\}_{k \in \mathbb{N}, n \geq 2}$ such that for all situations (N, K) and all profiles $\bar{P}, P \in \mathcal{P}^{\mathbb{N}}$,*

$$M(\bar{P}_N^K) - M(P_N^K) = \frac{\sum_{\{x,y\} \in \bar{K}} \omega_k(r_{x,y}(\bar{P}_N^K)) \cdot \sigma_{x,y}(\bar{P}_N)}{\sum_{\{x,y\} \in \bar{K}} \omega_k(r_{x,y}(\bar{P}_N^K))} - \frac{\sum_{\{x,y\} \in K} \omega_k(r_{x,y}(P_N^K)) \cdot \sigma_{x,y}(P_N)}{\sum_{\{x,y\} \in K} \omega_k(r_{x,y}(P_N^K))},$$

where for all $v \in \{1, \frac{3}{2}, 2, \frac{5}{2}, \dots, 2k-3\}$, $\omega_k(v) = 1 + 2\alpha_k(v-1)$.

The proof of this proposition is rather lengthy as we will have to establish several claims as we proceed. In order to maintain the overall readability, the argument will therefore be developed in three different subparts.

Part 2.1: We derive the magnitude of the difference $M(\bar{P}_N^K) - M(P_N^K)$.

Observe that it is possible to arrive at $\bar{P}_N^K$ starting from P_N^K by means of a (not unique) succession of $m \geq 1$ changes of contiguous pairs of objects. During the proof, we are going to decompose the difference $M(\bar{P}_N^K) - M(P_N^K)$ into $\frac{k(k-1)}{2}$ quantities, $\{A_{x,y}\}_{\{x,y\} \in \bar{K}}$, where we accumulate the differences that accrue during the process. Suppose that in the first step of the process, the preference profile P_N^K is changed to $\tilde{P}_N^K$, where $\tilde{P}_N^K$ is (x,y) -different from P_N^K for some individual $i \in N$. There are four possible cases:

- a. If $x \succ_{P_N} y$, then $M(\tilde{P}_N^K) > M(P_N^K)$ by MON. By RAN, the change in the measure only depends on the numbers $p_x(P_N^K)$, $p_y(P_N^K)$, $i_x(P_N^K)$, and $i_y(P_N^K)$; that is, $M(\tilde{P}_N^K) - M(P_N^K) = \delta_N^K(p_x(P_N^K), p_y(P_N^K), i_x(P_N^K), i_y(P_N^K))$. We also define the function ρ_N^K as $\rho_N^K(\cdot) = \frac{n}{2} \delta_N^K(\cdot)$. We classify the differences in terms of pairs of objects as follows: $A_{x,y}^1 = \frac{2}{n} \rho_N^K(p_x(P_N^K), p_y(P_N^K), i_x(P_N^K), i_y(P_N^K))$ and $A_{w,z}^1 = 0$ for all $\{w,z\} \in \bar{K} \setminus \{\{x,y\}\}$.
- b. If $x \sim_{P_N} y$ (and, then, $x \succ_{\bar{P}_N} y$), the properties of RAN and CON imply together that the change in measure, $M(\tilde{P}_N^K) - M(P_N^K)$, is equal to $\delta_N^K(p_x(\tilde{P}_N^K), p_y(\tilde{P}_N^K), i_x(\tilde{P}_N^K), i_y(\tilde{P}_N^K)) + t_N^K \sum_{z \in K \setminus \{x,y\}} (M(P_N^{\{x,z\}}) - M(P_N^{\{y,z\}}))$. It can be seen that the sum consists of two parts: The function $\delta_N^K(\cdot)$ refers to the change in the measure in case x would win against y in the pairwise comparison ceteris paribus, while the second part adapts $\delta_N^K(\cdot)$ to the fact

that the two objects tie at P_N^K . Also, by Proposition 2, for all $z \in K \setminus \{x, y\}$, $M(P_N^{\{x,z\}}) = \sigma_{x,z}(P_N)$ and $M(P_N^{\{y,z\}}) = \sigma_{y,z}(P_N)$. If we define the function $\rho_N^K(\cdot)$ in the same way as before, the difference can be decomposed as follows: $A_{x,y}^1 = \frac{2}{n} \rho_N^K(p_x(\tilde{P}_N^K), p_y(\tilde{P}_N^K), i_x(\tilde{P}_N^K), i_y(\tilde{P}_N^K))$, $A_{x,z}^1 = t_N^K \sigma_{x,z}(P_N)$ and $A_{y,z}^1 = -t_N^K \sigma_{y,z}(P_N)$ for all $z \in K \setminus \{x, y\}$, and $A_{v,w}^1 = 0$ for all $\{v, w\} \in \bar{T}$, where $T = K \setminus \{x, y\}$.

- c. If $y \succ_{\tilde{P}_N} x$ (and, then, $y \succ_{P_N} x$), $M(P_N^K) > M(\tilde{P}_N^K)$ by MON. It follows then from RAN that $M(\tilde{P}_N^K) - M(P_N^K) = -\delta_N^K(p_y(\tilde{P}_N^K), p_x(\tilde{P}_N^K), i_y(\tilde{P}_N^K), i_x(\tilde{P}_N^K))$. Since $p_z(\tilde{P}_N^K) = p_z(P_N^K)$ and $i_z(\tilde{P}_N^K) = i_z(P_N^K)$ for all $z \in \{x, y\}$, we finally obtain that $A_{x,y}^1 = -\frac{2}{n} \rho_N^K(p_y(P_N^K), p_x(P_N^K), i_y(P_N^K), i_x(P_N^K))$ and $A_{w,z}^1 = 0$ for all $\{w, z\} \in \bar{K} \setminus \{\{x, y\}\}$.
- d. If $x \sim_{\tilde{P}_N} y$ (and, then, $y \succ_{P_N} x$), the properties of RAN and CON imply together that $M(\tilde{P}_N^K) - M(P_N^K) = -\delta_N^K(p_y(P_N^K), p_x(P_N^K), i_y(P_N^K), i_x(P_N^K)) - t_N^K \sum_{z \in K \setminus \{x, y\}} (M(\tilde{P}_N^{\{y,z\}}) - M(\tilde{P}_N^{\{x,z\}}))$. Also, by Proposition 2, $M(\tilde{P}_N^{\{w,z\}}) = M(P_N^{\{w,z\}}) = \sigma_{w,z}(P_N)$ for all $w \in \{x, y\}$ and $z \in K \setminus \{x, y\}$. Consequently, the difference in the cohesiveness between the two preference profiles can be decomposed as follows: $A_{x,y}^1 = -\frac{2}{n} \rho_N^K(p_y(P_N^K), p_x(P_N^K), i_y(P_N^K), i_x(P_N^K))$, $A_{x,z}^1 = -t_N^K \sigma_{x,z}(P_N)$ and $A_{y,z}^1 = t_N^K \sigma_{y,z}(P_N)$ for all $z \in K \setminus \{x, y\}$, and $A_{v,w}^1 = 0$ for all $\{v, w\} \in \bar{T}$, where $T = K \setminus \{x, y\}$.

The remaining $(m-1)$ steps of the process of arriving at $\bar{P}_N^K$ are very similar to the first step. Consequently, we get that $M(\bar{P}_N^K) - M(P_N^K) = \sum_{\{x,y\} \in \bar{K}} \sum_{i=1}^m A_{x,y}^i$. This concludes Part 2.1.

Part 2.2: Three symmetry properties of the function ρ are established.

Since we have calculated the difference $M(\bar{P}_N^K) - M(P_N^K)$ in a very direct way, the formula obtained does not provide many insights on how to calculate the cohesiveness. In three claims, we are going to establish symmetry properties of the

function ρ_N^K that will help us to derive the exact weights and thereby to simplify the expression obtained at the end of Part 2.1.

Claim 1 *For all situations (N, K) and all possible vectors (a, b, c, d) , $\rho_N^K(a, b, c, d) = \rho_N^K(a, b, c + 1, d - 1)$.*

Proof: Consider any situation (N, K) and any preference profile P_N^K that satisfy the following conditions for some $x, y, z \in K$: (a) $x \succ_{P_N} y$ and (y, x) is a contiguous pair of objects for individual 3; (b) $n_{z,x}(P_N) - n_{x,z}(P_N) = 2$ and (z, x) is a contiguous pair of objects for individual 1; (c) $y \sim_{P_N} z$ and (y, z) is a contiguous pair of objects for individual 2; and (d) $p_x(P_N^K) = a, p_y(P_N^K) = b, i_x(P_N^K) = c$ and $i_y(P_N^K) = d$.^{11,12}

We also introduce the preference profile $\bar{P}_N^K$ which is equal to P_N^K apart from that the three individuals mentioned before change one pair of objects: Individual 3 switches the pair (y, x) , individual 1 switches the pair (z, x) , and individual 2 switches the pair (y, z) . Consequently, this preference profile satisfies $(p_x(\bar{P}_N^K), p_y(\bar{P}_N^K), i_x(\bar{P}_N^K), i_y(\bar{P}_N^K)) = (a, b, c + 1, d - 1)$.

The proof now consists in exploiting the differences in the cohesiveness that accrue along two different paths from P_N^K to $\bar{P}_N^K$. This is done in such a way that all changes along the paths apart from $\rho_N^K(a, b, c, d)$ and $\rho_N^K(a, b, c + 1, d - 1)$ cancel out. To provide the necessary intuition, individual 3 changes her/his preferences on the pair (y, x) first in the first path and last in the second path. The other two individuals change preferences in the same order: Individual 1 changes preferences before individual 2 in both paths. In the following, we sometimes use the notation

¹¹Obviously, we made the implicit assumption that $\{1, 2, 3\} \subseteq N$. However, this is not crucial as it is always possible to increase the size of the society or change the identity of its individuals applying REP.

¹²Since we can work with arbitrary large societies due to REP, it is always possible to construct a preference profile for any possible vector (a, b, c, d) such that there is a preference profile $\bar{P}_N^K \in (\mathcal{P}^K)^N$ and a contiguous pair (y, x) , with $\{x, y\} \in \bar{K}$ and $x \succ_{\bar{P}_N} y$, in a preference ranking of an individual $i \in N$ with $(p_x(\bar{P}_N^K), p_y(\bar{P}_N^K), i_x(\bar{P}_N^K), i_y(\bar{P}_N^K)) = (a, b, c, d)$. Details can be provided upon request.

$P_{K,N}$ alternatively for P_N^K .

The first path

1. The preference profile $P_{K,N}^{1,1}$ is (x, y) -different from $P_{K,N}^{1,0} \equiv P_N^K$ for individual 3. By step (a) of Part 2.1, $M(P_{K,N}^{1,1}) - M(P_N^K) = \frac{2}{n} \rho_N^K(a, b, c, d)$.
2. The preference profile $P_{K,N}^{1,2}$ is (x, z) -different from $P_{K,N}^{1,1}$ for individual 1. By step (d) of Part 2.1, $M(P_{K,N}^{1,2}) - M(P_{K,N}^{1,1}) = -\frac{2}{n} \rho_N^K(p_z(P_{K,N}^{1,1}), a, i_z(P_{K,N}^{1,1}), c) - t_N^K \sum_{w \in K \setminus \{x, z\}} (\sigma_{z,w}(P_N^{1,1}) - \sigma_{x,w}(P_N^{1,1}))$.
3. The preference profile $\bar{P}_N^K \equiv P_{K,N}^{1,3}$ is (z, y) -different from $P_{K,N}^{1,2}$ for individual 2. It can be concluded from step (b) of Part 2.1, $M(P_{K,N}^{1,3}) - M(P_{K,N}^{1,2}) = \frac{2}{n} \rho_N^K(p_z(P_{K,N}^{1,2}), b, i_z(P_{K,N}^{1,2}), d) + t_N^K \sum_{w \in K \setminus \{z, y\}} (\sigma_{z,w}(P_N^{1,2}) - \sigma_{y,w}(P_N^{1,2}))$. This concludes the description of the first path.

The second path

1. The preference profile $P_{K,N}^{2,1}$ is (x, z) -different from $P_N^K \equiv P_{K,N}^{2,0}$ for individual 1. It can then be concluded from step (d) of Part 2.1 that $M(P_{K,N}^{2,1}) - M(P_N^K) = -\frac{2}{n} \rho_N^K(p_z(P_N^K), a, i_z(P_N^K), c) - t_N^K \sum_{w \in K \setminus \{x, z\}} (\sigma_{z,w}(P_N) - \sigma_{x,w}(P_N))$.
2. The preference profile $P_{K,N}^{2,2}$ is (z, y) -different from $P_{K,N}^{2,1}$ for individual 2. By step (b) of Part 2.1, $M(P_{K,N}^{2,2}) - M(P_{K,N}^{2,1}) = \frac{2}{n} \rho_N^K(p_z(P_{K,N}^{2,1}), b, i_z(P_{K,N}^{2,1}), d) + t_N^K \sum_{w \in K \setminus \{z, y\}} (\sigma_{z,w}(P_N^{2,1}) - \sigma_{y,w}(P_N^{2,1}))$.
3. The preference profile $P_{N,K}^{2,3} \equiv \bar{P}_N^K$ is (x, y) -different from $P_{K,N}^{2,2}$ for individual 3. By step (a) of Part 2.1, $M(\bar{P}_N^K) - M(P_{K,N}^{2,2}) = \frac{2}{n} \rho_N^K(a, b, c + 1, d - 1)$. This concludes the description of the second path.

Next, we compare the differences that accrue along the two paths. We have that $\rho_N^K(p_z(P_{K,N}^{1,1}), a, i_z(P_{K,N}^{1,1}), c)$ is equal to $\rho_N^K(p_z(P_N^K), a, i_z(P_N^K), c)$ because z has not been involved in any change in preferences until that point. Similarly,

$\rho_N^K(p_z(P_{K,N}^{1,3}), b, i_z(P_{K,N}^{1,3}), d)$ is equal to $\rho_N^K(p_z(P_{K,N}^{2,2}), b, i_z(P_{K,N}^{2,2}), d)$ because z has been switched along both paths in exactly the same way: It has only been made indifferent with x . Taking into account these relationships, one can then see that

$$\begin{aligned} M(P_{K,N}^{1,2}) - M(P_{K,N}^{1,1}) &= M(P_{K,N}^{2,1}) - M(P_{K,N}^{2,0}) + t_N^K \cdot \frac{2}{n} \\ M(P_{K,N}^{1,3}) - M(P_{K,N}^{1,2}) &= M(P_{K,N}^{2,2}) - M(P_{K,N}^{2,1}) - t_N^K \cdot \frac{2}{n} \end{aligned}$$

Consequently, all changes apart from $M(P_{K,N}^{1,1}) - M(P_{K,N})$ and $M(\bar{P}_{K,N}) - M(P_{K,N}^{2,2})$ cancel out along the two paths. Finally, since the starting and ending profiles in both paths are the same, it has to be the case that $M(P_{K,N}^{1,1}) - M(P_{K,N}) = M(\bar{P}_{K,N}) - M(P_{K,N}^{2,2})$. This equation is equivalent to $\rho_N^K(a, b, c, d) = \rho_N^K(a, b, c + 1, d - 1)$, which concludes the proof of the claim. $\square$

Claim 2 *For all situations (N, K) and all possible vectors (a, b, c, d) , $\rho_N^K(a, b, c, d) = \rho_N^K(a - 1, b + 1, c, d)$.*

Proof: Consider any situation (N, K) and any preference profile P_N^K that satisfy the following conditions for some $x, y, z \in K$: (a) $p_x(P_N^K) = a$, $p_y(P_N^K) = b$, $i_x(P_N^K) = c$, and $i_y(P_N^K) = d$; (b) $x \succ_{P_N} y$ and (y, x) is a contiguous pair of objects for individual 1; (c) $n_{z,x}(P_N) - n_{x,z}(P_N) = n_{y,z}(P_N) - n_{z,y}(P_N) = 2$; (d) the ordered pair (z, x) is contiguous for individuals 2 and 4; and (e) the ordered pair (y, z) is contiguous for individuals 3 and 5.¹³

Consider the preference profile $\bar{P}_N^K$ that differs from P_N^K only because the individuals 1 to 5 change the order of the above mentioned pairs of objects in their preferences. We consider again two different paths from P_N^K to $\bar{P}_N^K$. Individuals invert preferences along the first path following the order of the natural numbers. The second path differs from the first only because individual 1 now inverts her/his preference on the pair $\{x, y\}$ last instead of first. In the next step, we collect the differences in the cohesiveness that accrue along the two different paths.

¹³The argument in the footnotes of Claim 1 also apply here.

If we denote by $P_{K,N}^{i,j}$ the profile obtained after the j -th change in the i -th path, for $i \in \{1, 2\}$ and $j \in \{1, \dots, 5\}$,¹⁴ we have that:

$$M(P_{K,N}^{1,2}) - M(P_{K,N}^{1,1}) = M(P_{K,N}^{2,1}) - M(P_{K,N}^{2,0}) + t_N^K \cdot \frac{2}{n}$$

$$M(P_{K,N}^{1,3}) - M(P_{K,N}^{1,2}) = M(P_{K,N}^{2,2}) - M(P_{K,N}^{2,1}) - t_N^K \cdot \frac{2}{n}$$

$$M(P_{K,N}^{1,4}) - M(P_{K,N}^{1,3}) = M(P_{K,N}^{2,3}) - M(P_{K,N}^{2,2}) + t_N^K \cdot \frac{2}{n}$$

$$M(P_{K,N}^{1,5}) - M(P_{K,N}^{1,4}) = M(P_{K,N}^{2,4}) - M(P_{K,N}^{2,3}) - t_N^K \cdot \frac{2}{n}$$

Then, since the starting and ending profiles in both paths are the same, we have that

$M(P_{K,N}^{1,1}) - M(P_{K,N}^{1,0}) = M(P_{K,N}^{2,5}) - M(P_{K,N}^{2,4})$. Given that $M(P_{K,N}^{1,1}) - M(P_{K,N}^{1,0}) = \frac{2}{n} \rho_N^K(a, b, c, d)$ and that $M(P_{K,N}^{2,5}) - M(P_{K,N}^{2,4}) = \frac{2}{n} \rho_N^K(a - 1, b + 1, c, d)$, we have that $\rho_N^K(a, b, c, d) = \rho_N^K(a - 1, b + 1, c, d)$, which concludes the proof of the claim. $\square$

Claims 1 and 2 imply that we can now introduce the function β_N^K , which satisfies the condition that $\beta_N^K(p_x(P_N^K) + p_y(P_N^K), i_x(P_N^K) + i_y(P_N^K))$ equals the former $\rho_N^K(p_x(P_N^K), p_y(P_N^K), i_x(P_N^K), i_y(P_N^K))$. In the final claim of this part, we show that a win at the social level counts double an indifference; that is, $\beta_N^K(p_x(P_N^K) + p_y(P_N^K), i_x(P_N^K) + i_y(P_N^K)) = \beta_N^K(p_x(P_N^K) + p_y(P_N^K) + 1, i_x(P_N^K) + i_y(P_N^K) - 2)$.

Claim 3 *For all situations (N, K) and all possible vectors (a, b) , $\beta_N^K(a, b) = \beta_N^K(a + 1, b - 2)$.*

Proof: Consider any situation (N, K) and any preference profile P_N^K that satisfy the following conditions for some $x, y, z \in K$: (a) $x \succ_{P_N} y$ and (y, x) is a contiguous pair of objects for individual 1; (b) $x \sim_{P_N} z$, and (z, x) is a contiguous pair of objects for individual 2; (c) $y \sim_{P_N} z$ and (y, z) is a contiguous pair of objects for individual 3; and (d) $p_x(P_N^K) = c, p_y(P_N^K) = d, i_x(P_N^K) = e$ and $i_y(P_N^K) = f$ such that $c + d = a$ and $e + f = b$.¹⁵ Consider also the preference profile $\bar{P}_N^K$ which

¹⁴Note that $P_{K,N}^{i,0} \equiv P_N^K$ and $P_{K,N}^{i,5} \equiv \bar{P}_N^K$ for all $i \in \{1, 2\}$.

¹⁵The argument in the footnotes of Claim 1 also apply here.

is equal to P_N^K apart from that individuals 1, 2, and 3 switch the preferences on the pairs of objects mentioned before. The proof consists again in exploiting the differences in the cohesiveness that accrue along two different paths from P_N^K to $\bar{P}_N^K$.

The first path

1. The preference profile $P_{K,N}^{1,1}$ is (x, y) -different from P_N^K for individual 1. By step (a) of Part 2.1 and Claims 1 and 2, $M(P_{K,N}^{1,1}) - M(P_N^K) = \frac{2}{n} \beta_N^K(a, b)$.
2. The preference profile $P_{K,N}^{1,2}$ is (x, z) -different from $P_{K,N}^{1,1}$ for individual 2. By step (b) of Part 2.1 and Claims 1 and 2, $M(P_{K,N}^{1,2}) - M(P_{K,N}^{1,1}) = \frac{2}{n} \beta_N^K(c + p_z(P_{K,N}^{1,1}) + 1, e + i_z(P_{K,N}^{1,1}) - 2) + t_N^K \sum_{w \in K \setminus \{x, z\}} (\sigma_{x,w}(P_N^{1,1}) - \sigma_{z,w}(P_N^{1,1}))$.
3. The preference profile $\bar{P}_N^K$ is (z, y) -different from $P_{K,N}^{1,2}$ for individual 3. By step (b) of Part 2.1 and Claims 1 and 2, $M(\bar{P}_N^K) - M(P_{K,N}^{1,2}) = \frac{2}{n} \beta_N^K(p_z(\bar{P}_N^K) + d + 1, i_z(\bar{P}_N^K) + f - 2) + t_N^K \sum_{w \in K \setminus \{z, y\}} (\sigma_{z,w}(P_N^{1,2}) - \sigma_{y,w}(P_N^{1,2}))$.

The second path

1. The preference profile $P_{K,N}^{2,1}$ is (x, z) -different from P_N^K for individual 2. By step (b) of Part 2.1 and Claims 1 and 2, $M(P_{K,N}^{2,1}) - M(P_N^K) = \frac{2}{n} \beta_N^K(c + p_z(P_N^K) + 1, e + i_z(P_N^K) - 2) + t_N^K \sum_{w \in K \setminus \{x, z\}} (\sigma_{x,w}(P_N) - \sigma_{z,w}(P_N))$.
2. The preference profile $P_{K,N}^{2,2}$ is (z, y) -different from $P_{K,N}^{2,1}$ for individual 3. From step (b) of Part 2.1 and Claims 1 and 2 we get $M(P_{K,N}^{2,2}) - M(P_{K,N}^{2,1}) = \frac{2}{n} \beta_N^K(p_z(P_{K,N}^{2,1}) + d + 1, i_z(P_{K,N}^{2,1}) + f - 2) + t_N^K \sum_{w \in K \setminus \{z, y\}} (\sigma_{z,w}(P_N^{2,1}) - \sigma_{y,w}(P_N^{2,1}))$.
3. The preference profile $\bar{P}_N^K$ is (x, y) -different from $P_{K,N}^{2,2}$ for individual 1. By step (a) of Part 2.1 and Claims 1 and 2, $M(\bar{P}_N^K) - M(P_{K,N}^{2,2}) = \frac{2}{n} \beta_N^K(a + 1, b - 2)$.

As in the previous claims, it is easy to see that

$$M(P_{K,N}^{1,2}) - M(P_{K,N}^{1,1}) = M(P_{K,N}^{2,1}) - M(P_N^K) + t_N^K \cdot \frac{2}{n}$$

$$M(\bar{P}_N^K) - M(P_{K,N}^{1,2}) = M(P_{K,N}^{2,2}) - M(P_{K,N}^{2,1}) - t_N^K \cdot \frac{2}{n}.$$

Then, $\beta_N^K(a+1, b-2) = \beta_N^K(a, b)$, which concludes the proof. $\square$

Claim 3 establishes that the weight of a pair $\{x, y\} \in \bar{K}$ depends on $r_{x,y}(P_N^K)$; that is, we can introduce the function ω_N^K which is such that for all (a, b) , $\omega_N^K(v) = \beta_N^K(a, b)$, where $v = a + \frac{b}{2}$. This concludes Part 2.2.

Part 2.3: We derive the exact expression of $M(\bar{P}_N^K) - M(P_N^K)$.

Remember that the objective of Part 2.1 has been to classify the differences accumulated in the process of arriving at $\bar{P}_N^K$ from P_N^K in pairs of objects. In effect, we have seen that $M(\bar{P}_N^K) - M(P_N^K) = \sum_{\{x,y\} \in \bar{K}} \sum_{i=1}^m A_{x,y}^i$. Using the results from Part 2.2 we then get that the accumulated difference in the pair $\{x, y\} \in \bar{K}$ along all changes of contiguous pairs equals $\sum_{i=1}^m A_{x,y}^i = \sigma_{x,y}(P_N) (\omega_N^K(r_{x,y}(\bar{P}_N^K)) - \omega_N^K(r_{x,y}(P_N^K))) + (\sigma_{x,y}(\bar{P}_N) - \sigma_{x,y}(P_N)) \omega_N^K(r_{x,y}(\bar{P}_N^K))$. Summing this difference over all possible pairs of objects, we obtain that

$$M(\bar{P}_N^K) - M(P_N^K) = \frac{\sum_{\{x,y\} \in \bar{K}} \omega_N^K(r_{x,y}(\bar{P}_N^K)) \cdot \sigma_{x,y}(\bar{P}_N)}{\sum_{\{x,y\} \in \bar{K}} \omega_N^K(r_{x,y}(\bar{P}_N^K))} - \frac{\sum_{\{x,y\} \in \bar{K}} \omega_N^K(r_{x,y}(P_N^K)) \cdot \sigma_{x,y}(P_N)}{\sum_{\{x,y\} \in \bar{K}} \omega_N^K(r_{x,y}(P_N^K))}.$$

To fully complete the proof of the proposition, it remains to be shown that (a) for all situations (A, S) and (B, T) such that $|S| = |T|$, $\omega_A^K(v) = \omega_B^K(v)$ for all $v \in \{1, \frac{3}{2}, 2, \dots, 2k-3\}$ and (b) for each situation (N, K) , there exists $\alpha_k \geq 0$ such that $\omega_N^K(v) = 1 + 2\alpha_k(v-1)$ for all possible v .

First, since the weights have to be positive by MON, it is possible to normalize and fix $\omega_N^K(1) = 1$ for all (N, K) . We now proof the following claim.

Claim 4 *For each situation (N, K) , there is $\alpha_k \geq 0$ so that for all $v \in \{\frac{3}{2}, 2, \dots, 2k-3\}$, $\omega_N^K(v) = \omega_N^K(v - \frac{1}{2}) + \alpha_k$.*

Proof: We will only show that $\omega_N^K(\frac{3}{2}) = \omega_N^K(1) + \alpha_k$, with $\alpha_k \geq 0$, because all other equations can be derived using similar arguments. Consider a preference profile P_N^K which satisfies the following conditions for some $x, y, z \in K$: (a) $x \succ_{P_N} y$ and (y, x)

is contiguous pair of objects for individual 1; (b) $n_{x,z}(P_N) + 2 = n_{z,x}(P_N)$ and (z, x) is a contiguous pair of objects for individual 2; and (c) for all $w \in K \setminus \{x, y\}$, $w \succ_{P_N} x$ and $w \succ_{P_N} y$.¹⁶ Consider also the preference profile $\bar{P}_N^K$ which is equal to P_N^K apart from that individuals 1 and 2 have changed their preferences over the pairs (y, x) and (z, x) , respectively. The proof consists in exploiting the differences in the cohesiveness that accrue along two different paths from P_N^K to $\bar{P}_N^K$.

The first path

1. The preference profile $\hat{P}_N^K$ is (x, y) -different from P_N^K for individual 1. Hence, by step (a) in Part 2.1 and Claim 3, $M(\hat{P}_N^K) - M(P_N^K) = \frac{2}{n} \omega_N^K(1)$.
2. The preference profile $\bar{P}_N^K$ is (x, z) -different from $\hat{P}_N^K$ for individual 2. Hence, it follows from step (d) in Part 2.1 and Claim 3 that $M(\bar{P}_N^K) - M(\hat{P}_N^K) = -\frac{2}{n} \omega_N^K(r_{x,z}(\hat{P}_N^K)) - t_N^K \sum_{w \in K \setminus \{x,z\}} (\sigma_{z,w}(\hat{P}_N^K) - \sigma_{x,w}(\hat{P}_N^K))$.

The second path

1. The preference profile $\tilde{P}_N^K$ is (x, z) -different from P_N^K for individual 2. Hence, it follows from step (d) in Part 2.1 and Claim 3 that $M(\tilde{P}_N^K) - M(P_N^K) = -\frac{2}{n} \omega_N^K(r_{x,z}(P_N^K)) - t_N^K \sum_{w \in K \setminus \{x,z\}} (\sigma_{z,w}(P_N^K) - \sigma_{x,w}(P_N^K))$.
2. The preference profile $\bar{P}_N^K$ is (x, y) -different from $\tilde{P}_N^K$ for individual 1. Hence, by step (a) in Part 2.1 and Claim 3, $M(\bar{P}_N^K) - M(\tilde{P}_N^K) = \frac{2}{n} \omega_N^K(\frac{3}{2})$.

As in the previous claims, it is easy to see that $M(\bar{P}_N^K) - M(\hat{P}_N^K) = M(\tilde{P}_N^K) - M(P_N^K) + t_N^K \cdot \frac{2}{n}$. Since the difference in the cohesiveness between $\bar{P}_N^K$ and P_N^K has to be the same independently of the path used, we get that $\frac{2}{n} \omega_N^K(1) + t_N^K \frac{2}{n} = \frac{2}{n} \omega_N^K(\frac{3}{2})$. Hence, $\omega_N^K(\frac{3}{2}) = \omega_N^K(1) + t_N^K$. Finally, it is easy to see that t_N^K has to be independent of N (by REP), that it depends only on the size of the set of objects and not the set itself (by NEU), and that it is non-negative (by RAN). Consequently, we can set t_N^K to be equal to α_k . This concludes the proof of the claim. $\square$

¹⁶The argument in the footnotes of Claim 1 also apply here.

To finally conclude the proof of the proposition, observe that by Claim 4 and the normalization by which $\omega_N^K(1) = 1$ for all (N, K) , ω_N^K is only function of the size of the set of objects; that is, for all $v \in \{1, 2, \dots, 2k-3\}$, we can set $\omega_N^K(v) = \omega_k(v)$.

Part 3: Finally, we show that the cohesiveness is zero whenever $\sigma_{x,y} = 0$ for all $\{x, y\} \in \bar{K}$. Plugging this result back into Proposition 3 concludes the proof of the theorem.

Lemma 3 *If the cohesiveness measure M satisfies REP, NEU, MON, RAN, and CON, then for all situations (N, K) and all profiles $P \in \mathcal{P}^{\mathbb{N}}$ such that $\sigma_{x,y}(P_N) = 0$ for all $\{x, y\} \in \bar{K}$,*

$$M(P_N^K) = 0.$$

Proof: Take any situation (N, K) and any preference profile P_N^K such that for all $\{x, y\} \in \bar{K}$, $\sigma_{x,y}(P_N) = 0$. By definition of M , there exists a society A (of size a) and a profile $\bar{P}$ such that $M(\bar{P}_A^K) = 0$. Consider the society B (of size $b = a \cdot n$) and two preference profiles $\tilde{P}_B^K$ and $\hat{P}_B^K$, where $\tilde{P}_B^K$ is obtained from a isomorphic and disjoint copies of P_N^K and $\hat{P}_B^K$ is obtained from n isomorphic and disjoint copies of $\bar{P}_A^K$. By REP, $M(\hat{P}_B^K) = M(\bar{P}_A^K) = 0$. Hence, by Proposition 3,

$$M(\tilde{P}_B^K) - 0 = \frac{\sum_{\{x,y\} \in \bar{K}} \omega_k(r_{x,y}(\tilde{P}_B^K)) \cdot \sigma_{x,y}(\tilde{P}_B)}{\sum_{\{x,y\} \in \bar{K}} \omega_k(r_{x,y}(\tilde{P}_B^K))} - \frac{\sum_{\{x,y\} \in \bar{K}} \omega_k(r_{x,y}(\hat{P}_B^K)) \cdot \sigma_{x,y}(\hat{P}_B)}{\sum_{\{x,y\} \in \bar{K}} \omega_k(r_{x,y}(\hat{P}_B^K))}.$$

Since $\sigma_{x,y}(P_N) = 0$ for all $\{x, y\} \in \bar{K}$ by assumption, it follows from the construction of $\tilde{P}_B^K$ that $\sigma_{x,y}(\tilde{P}_B) = 0$ for all $\{x, y\} \in \bar{K}$ as well. Consequently, the former equation reduces to

$$M(\tilde{P}_B^K) = - \frac{\sum_{\{x,y\} \in \bar{K}} \omega_k(r_{x,y}(\hat{P}_B^K)) \cdot \sigma_{x,y}(\hat{P}_B)}{\sum_{\{x,y\} \in \bar{K}} \omega_k(r_{x,y}(\hat{P}_B^K))}.$$

Observe that all weights $\omega_k(\cdot)$ are strictly positive by Proposition 3 and that the values of σ cannot be negative. Given that the values of the cohesiveness should be

non-negative, the only solution to the equation is $\sigma_{x,y}(\hat{P}_B) = 0$ for all $\{x, y\} \in \bar{K}$ and, therefore, $M(\tilde{P}_B^K) = 0$. Finally, $M(P_N^K) = M(\tilde{P}_B^K) = 0$ by REP. $\square$

Independence of the Axioms (Proposition 1)

We will provide cohesiveness measures that satisfy all the properties but one.

Replication Invariance: Let the cohesiveness measure M_1 be such that for all situations (N, K) and all preference profiles P_N^K ,

$$M_1(P_N^K) = \Phi_0(P_N^K) \text{ if } 1 \in N \text{ and } M_1(P_N^K) = \Phi_{0.5}(P_N^K) \text{ if } 1 \notin N.$$

This cohesiveness measure satisfies NEU, MON, RAN and CON. The following example shows that it does not satisfy replication invariance.

Let $N = \{1, 2\}$, $S = \{2, 3\}$, and $K = \{x, y, z\}$. Suppose that the profile P is such that xP_1yP_1z , xP_2zP_2y , and $P_3 = P_1$. Then, $M_1(P_N^K) = \frac{2}{3}$ and $M_1(P_S^K) = \frac{5}{6}$. REP would imply that $M_1(P_N^K) = M_1(P_S^K)$.

Neutrality: Let $q : \bar{X} \rightarrow \mathbb{R}_{++}$ be a function that assigns to each pair of alternatives $\{x, y\} \in \bar{X}$ a strictly positive weight $q_{x,y} > 0$ in such a way that $q_{x,y} \neq q_{w,z}$ for some $\{w, z\} \neq \{x, y\}$. Now, let cohesiveness measure M_2 be such that for all situations (N, K) and all preference profiles P_N^K ,

$$M_2(P_N^K) = \sum_{\{x,y\} \in \bar{K}} \frac{q_{x,y}}{\sum_{\{w,z\} \in \bar{K}} q_{w,z}} \sigma_{x,y}(P_N).$$

This cohesiveness measure satisfies REP, MON, RAN, and CON. The following example shows that it is not neutral.

Let $N = \{1, 2\}$ and $K = \{x, y, z\}$. Suppose that the preference profiles P_N^K and $\bar{P}_N^K$ are such that xP_1yP_1z , yP_2xP_2z , $z\bar{P}_1y\bar{P}_1x$, and $y\bar{P}_2z\bar{P}_2x$. Moreover, let $q_{x,z} = q_{y,z} = 1$ and $q_{x,y} = 2$. Then, $M_2(P_N^K) = \frac{2}{4}$ and $M_2(\bar{P}_N^K) = \frac{3}{4}$. NEU would imply that $M_2(P_N^K) = M_2(\bar{P}_N^K)$.

Monotonicity: Let the cohesiveness measure M_3 be such that for all situations (N, K) and all preference profiles P_N^K ,

$$M_3(P_N^K) = 1 - \frac{2}{k(k-1)} \sum_{\{x,y\} \in \bar{K}} \sigma_{x,y}(P_N).$$

This cohesiveness measure satisfies REP, NEU, RAN, and CON. The following example shows that it is not monotone.

Let $N = \{1, 2\}$ and $K = \{x, y\}$. Suppose that the preference profiles P_N^K and $\bar{P}_N^K$ are such that xP_1y , yP_2x , $x\bar{P}_1y$, and $x\bar{P}_2y$. Then, $M_3(P_N^K) = 1$ and $M_3(\bar{P}_N^K) = 0$. MON would imply that $M_3(\bar{P}_N^K) > M_3(P_N^K)$.

Ranking: Let the cohesiveness measure M_4 be such that for all situations (N, K) and all preference profiles P_N^K ,

$$M_4(P_N^K) = \frac{2}{k(k-1)} \sum_{\{x,y\} \in \bar{K}} (2k-2-r_{x,y}(P_N^K)) \cdot \sigma_{x,y}(P_N^K).$$

This cohesiveness measure satisfies REP, NEU, MON, and CON. The following example shows that it does not satisfy ranking.

Let $N = \{1, 2, 3\}$ and $K = \{x, y, z\}$. Suppose that the preference profile P_N^K is such that xP_1yP_1z , $P_2 = P_1$ and zP_3yP_3x . Also, let the preference profiles $\bar{P}_N^K$ and $\hat{P}_N^K$ be such that $\bar{P}_N^K$ is (x, y) -different from P_N^K for individual 3 and $\hat{P}_N^K$ is (y, z) -different from $\bar{P}_N^K$ for individual 3. Then, $M_4(P_N^K) = \frac{1}{3}$, $M_4(\bar{P}_N^K) = \frac{4}{9}$, and $M_4(\hat{P}_N^K) = \frac{2}{3}$. So, $M_4(\bar{P}_N^K) - M_4(P_N^K) = \frac{1}{9}$ and $M_4(\hat{P}_N^K) - M_4(P_N^K) = \frac{1}{3}$. RAN would imply that $M_4(\bar{P}_N^K) - M_4(P_N^K) \geq M_4(\hat{P}_N^K) - M_4(P_N^K)$.

Consistency: Let $\{a_n\}_{n \in \{4, 6, 8, \dots\}}$ be the sequence of real numbers in the unit interval determined by the equations $a_4 = 0.6$ and $a_n = a_{n \cdot s} + (s-1) \cdot \frac{2(1-a_{n \cdot s})}{s \cdot n - 2}$ for all $n \in \{4, 6, 8, \dots\}$ and all $s \in \mathbb{N}$. Consider now, for each pair of objects $\{x, y\}$, the function $\Omega_{x,y}$ such that for each preference profile P_N , $\Omega_{x,y}(P_N) = 0$ whenever $\sigma(P_N) = 0$; $\Omega_{x,y}(P_N) = a_n + \frac{2(1-a_n)}{n-2} \cdot (\frac{|n_{x,y}(P_N)|}{2} - 1)$ when $\sigma(P_N) > 0$ and n is even; and $\Omega_{x,y}(P_N) = a_{2 \cdot n} + \frac{1-a_{2 \cdot n}}{n-1} \cdot (|n_{x,y}(P_N)| - 1)$ when n is odd. Now, let the

cohesiveness measure M_5 be such that for all situations (N, K) and all preference profiles P_N^K ,

$$M_5(P_N^K) = \frac{2}{k(k-1)} \sum_{\{x,y\} \in \bar{K}} \Omega_{x,y}(P_N).$$

This cohesiveness measure satisfies REP, NEU, MON, and RAN. The following example shows that it is not consistent.

Let $N = \{1, 2, 3, 4\}$ and $K = \{x, y\}$. Suppose that the preference profile P_N is such that xP_1y , $P_2 = P_1$, yP_3x , and $P_4 = P_3$. Also, let the preference profiles $\bar{P}_N^K$ and $\tilde{P}_N^K$ be such that $\bar{P}_N^K$ is (x, y) -different from P_N^K for individual 3 and $\tilde{P}_N^K$ is (x, y) -different from $\bar{P}_N^K$ for individual 4. Then, $M_5(P_N^K) = 0$, $M_5(\bar{P}_N^K) = \frac{3}{5}$, and $M_5(\tilde{P}_N^K) = 1$. So, $M_5(\tilde{P}_N^K) - M_5(\bar{P}_N^K) = \frac{2}{5}$ and $M_5(\bar{P}_N^K) - M_5(P_N^K) = \frac{3}{5}$. CON would imply that $M_5(\tilde{P}_N^K) - M_5(\bar{P}_N^K) = M_5(\bar{P}_N^K) - M_5(P_N^K)$.

References

- [1] Alcalde-Unzu, J. and Vorsatz, M. (2008). The Measurement of Consensus: An Axiomatic Analysis. FEDEA WP 2008-28.
- [2] Alesina, A. and Rodrik, D. (1994). Distributive Politics and Economic Growth. *Quarterly Journal of Economics* 109, 465–490.
- [3] Atkinson, A. (1970). On the Measurement of Inequality. *Journal of Economic Theory* 2, 244-263.
- [4] Baldiga, K. and Green, J. (2009). Choice-Based Measures of Conflict in Preferences. Harvard University.
- [5] Cowell, F. (2008). Measuring Inequality. *LSE Perspectives in Economic Analysis*, Oxford University Press.
- [6] Cook, W. and Seiford, L. (1978). Priority Ranking and Consensus Formation. *Management Science* 24, 1721–1732.

- [7] Cook, W. and Seiford, L. (1982). On the Borda–Kendall Consensus Method for Priority Ranking Problems. *Management Science* 28, 621–637.
- [8] Dalton, H. (1920). The Measurement of the Inequality of Incomes. *Economic Journal* 30, 348–461.
- [9] Duclos, J. , Esteban, J. and Ray, D. (2006). Polarization: Concepts, Measurement, Estimation. *Econometrica* 74, 1737–1772.
- [10] Echenique, F. and Fryer, R. (2007). A Measure of Segregation Based on Social Interactions. *Quarterly Journal of Economics* 122, 1241–285.
- [11] Eisenberg, J. (2007). Group Cohesiveness. In: Encyclopedia of Social Psychology (eds. F. Baumeister and K. Vohs), Thousands Oaks, Sage, 386–388.
- [12] Esteban, J. and Ray, D. (1994). On the Measurement of Polarization. *Econometrica* 62, 819–852.
- [13] García-Lapresta, J. and Pérez-Roman, D. (2008). Some consensus measures and their applications in group decision making. In: Computational Intelligence in Decision and Control. (eds. D. Ruan, J. Montero, J. Lu, L. Martínez, P. D’Hondt, E. Kerre). World Scientific, Singapore, 611–616.
- [14] Gini, C. (1921). Measurement of Inequality and Incomes. *Economic Journal* 31, 124–126.
- [15] Hays, W. (1960). A Note on Average Tau as a Measure of Concordance. *Journal of the American Statistical Association* 55, 331–341.
- [16] Hutchens, R. (2004). One Measure of Segregation. *International Economic Review* 45, 555–578.
- [17] Kendall, M. (1962). Rank Correlation Methods, 3rd edition, Hafner Publishing Company, New York.

- [18] Kemeny, J. (1959). Mathematics Without Numbers. *Daedalus* 88, 577–591.
- [19] Klamler, C. (2008). A Distance Measure for Choice Functions. *Social Choice and Welfare* 30, 419–425, 2008.
- [20] Kranich, L. (1996). Equitable Opportunities: An Axiomatic Approach. *Journal of Economic Theory* 71, 131–147.
- [21] Montalvo, D. and Reynal-Querol, M. (2005). Ethnic Polarization, Political Conflict, and Civil War. *American Economic Review* 95, 796–816.
- [22] Persson, P. and Tabellini, G. (1994). Is Inequality Harmful for Growth? *American Economic Review* 90, 869–887.
- [23] Roemer, J. (1998). Equality of Opportunity. Harvard University Press.

ÚLTIMOS DOCUMENTOS DE TRABAJO

- 2010-23: Do we agree? Measuring the cohesiveness of preferences”, **Jorge Alcalde-Unzu y Marc Vorsatz.**
- 2010-22: “The Weight of the Crisis: Evidence From Newborns in Argentina”, **Carlos Bozzoli y Climent Quintana-Domeque.**
- 2010-21: “Exclusive Content and the Next Generation Networks”-, **Juan José Ganuza and María Fernanda Vicens.**
- 2010-20: “The Determinants of Success in Primary Education in Spain”, **Brindusa Anghel y Antonio Cabrales.**
- 2010-19: “Explaining the fall of the skill wage premium in Spain”, **Florentino Felgueroso, Manuel Hidalgo y Sergi Jiménez-Martín.**
- 2010-18: “Some Students are Bigger than Others, Some Students’ Peers are Bigger than Other Students’ Peers”, **Toni Mora y Joan Gil.**
- 2010-17: “Electricity generation cost in isolated system: the complementarities of natural gas and renewables in the Canary Islands”, **Gustavo A. Marrero y Francisco Javier Ramos-Real.**
- 2010-16: “Killing by lung cancer or by diabetes? The trade-off between smoking and obesity”, **Federico A. Todeschini, José María Labeaga y Sergi Jiménez-Martín.**
- 2010-15: “Does gender matter for academic promotion? Evidence from a randomized natural experiment”, **Natalia Zinovyeva y Manuel F. Bagues.**
- 2010-14: “Spain, Japan, and the Dangers of Early Fiscal Tightening”, **R. Anton Braun y Javier Díaz-Giménez.**
- 2010-13: “Competition and horizontal integration in maritime freight transport”, **Pedro Cantos-Sanchez, Rafael Moner-Colonques, Jose Sempere-Monerris y Óscar Alvarez.**
- 2010-12: “Eppur si Muove! Spain: Growing without a Model”, **Michele Boldrin, J. Ignacio Conde-Ruiz y Javier Díaz Gimenez.**
- 2010-11: “The Spanish Business Bankruptcy Puzzle and the Crisis”, **Marco Celentani, Miguel García-Posada y Fernando Gómez.**
- 2010-10: “Promoting Employment of Disabled Women in Spain; Evaluating a Policy”, **Judit Vall Castello.**
- 2010-09: “The Role of Construction in the Housing Boom and Bust in Spain”, **Carlos Garriga.**
- 2010-08: “Did Good Cajas Extend Bad Loans? Governance, Human Capital and Loan Portfolios”, **Vicente Cuñat y Luis Garicano.**
- 2010-07: “Unemployment and Temporary Jobs in the Crisis: Comparing France and Spain”, **Samuel Bentolila, Pierre Cahuc, Juan J. Dolado y Thomas Le Barbanchon.**
- 2010-06: “La Subida del Impuesto Sobre el Valor Añadido en España: Demasiado Cara y Demasiado Pronto”, **Juan Carlos Conesa , Javier Díaz-Giménez, Julián Díaz-Saavedra y Josep Pijoan-Mas.**
- 2010-05: “Off-the-peak preferences over government size”, **Francisco Martínez-Mora y M. Socorro Puy.**
- 2010-04: “Green Shoots? Where, when and how?”, **Maximo Camacho, Gabriel Perez-Quiros y Pilar Poncela.**
- 2010-03: “The Spanish Crisis from a Global Perspective”, **Jesús Fernández-Villaverde y Lee Ohanian.**
- 2010-02: “Fiscal Centralization and the Political Process”, **Facundo Albornoz y Antonio Cabrales.**
- 2010-01: “The Evolution of Adult Height Across Spanish Regions 1950-1980: A New Source of Data”, **Mariano Bosch, Carlos Bozzoli y Climent Quintana-Domeque.**
- 2009-40: “Demographic Change and Pension Reform in Spain: An Assessment in a Two-Earner, OLG Model”, **Alfonso R. Sánchez Martín y Virginia Sánchez Marcos.**
- 2009-39: “The new growth model: How and with Whom?”, **Florentino Felgueroso y Sergi Jiménez-Martín.**
- 2009-38: “Rain and the Democratic Window of Opportunity”, **Markus Brückner and Antonio Ciccone.**
- 2009-37: “International Commodity Prices, Growth, and the Outbreak of Civil War in Sub-Saharan Africa”, **Markus Brückner and Antonio Ciccone.**
- 2009-36: “Determinants of Economic Growth: Will Data Tell?”, **Antonio Ciccone y Marek Jarocinski.**
- 2009-35: “The Timing of Work and Work-Family Conflicts in Spain: Who Has a Split Work Schedule and Why?”, **Catalina Amuedo-Dorantes y Sara de la Rica.**
- 2009-34: “Fatter Attraction: Anthropometric and Socioeconomic Characteristics in the Marriage Market”, **Pierre-André Chiappori, Sonia Oreffice y Climent Quintana-Domeque.**
- 2009-33: “Infant disease, economic conditions at birth and adult stature in Brazil”, **Víctor Hugo de Oliveira Silva y Climent Quintana-Domeque.**
- 2009-32: “Are Drinkers Prone to Engage in Risky Sexual Behaviors?”, **Ana I. Gil Lacruz, Marta Gil Lacruz y Juan Oliva Moreno.**
- 2009-31: “Factors Explaining Charges in European Airports: Competition, Market Size, Private Ownership and Regulation”, **Germà Bel y Xavier Fageda.**