Soc Choice Welfare (1992) 9: 99–116

Social Choice and Welfare

Two-stage discrete aggregation: the Ostrogorski paradox and related phenomena

Manfred Nermuth*

Institut für Wirtschaftswissenschaften, Universität Wien, Lueger-Ring 1, A-1010 Wien, Austria

Received May 22, 1991 / Accepted November 13, 1991

Abstract. Motivated by certain paradoxa that have been discussed in the literature (Ostrogorski paradox), we prove an impossibility theorem for *two-stage aggregation procedures for discrete data.* We consider aggregation procedures of the following form: The whole population is partitioned into subgroups. First we aggregate over each subgroup, and in a second step we aggregate the subgroup aggregates to obtain a total aggregate. The data are either dichotomous $(1-0;$ yes-no) or take values in a finite ordered set of possible attributes (e.g. exam grades $A, B, \ldots F$). Examples are given by multistage voting procedures (indirect democracy, federalism), or by the forming of partial grades and overall grades in academic examinations and similar evaluation problems (sports competitions, consumer reports). It is well known from standard examples that the result of such a two-stage aggregation procedure depends, in general, not only on the distribution of attributes in the whole population, but also on *how* the attributes are distributed across the various subgroups (in other words: how the subgroups are defined). This dependence leads to certain "paradoxa". The *main result* of the present paper is that these paradoxa are not due to the special aggregation rules employed in the examples, but are *unavoidable in principle*, provided the aggregators satisfy certain natural assumptions. More precisely: the only aggregator functions for which the result of a two-stage (a fortiori: multi-stage) aggregation does not depend on the partitioning are "degenerate" aggregators of the following form: there exists a partial order ("dominance") on the set of possible attributes such that the aggregate over any collection of data is always equal to the supremum (w.r.t. dominance) of the attributes occurring in the data, regardless of the relative frequencies of these occurrences. In the voting context, degeneracy corresponds to the unanimity principle. Our theorem is true for arbitrary partitionings of arbitrary (finite) sets and generalizes the results of

* This paper was completed during a visit to the University of Bielefeld. I am much indebted to the Faculty of Economics there for its hospitality; in particular I should like to thank Gerhard Schwödiauer and Walter Trockel for their support.

Deb & Kelsey (for the matrix case with dichotomous variables and majority voting) to general two-stage aggregation procedures for attributes belonging to a finite ordered set. The general result is illustrated by some examples.

1. Introduction

The research reported in this paper was motivated by the following examples.

Example 1.1.a. Ostrogorski Paradox (Daudt and Rae 1976). There are five *types of voters*, $n = A, B, C, D, E$, each type representing 20% of the electorate; three *issues*, $m = \alpha, \beta, \gamma$, and two political parties, 1 (Red) and 0 (Black), whose platforms differ on all three issues. The voters' opinions are shown in Table 1.

Table 1. The Ostrogorski Paradox

	A	B	C	D	E	
α	0	1	1	0	0	0
β	1	0	1	0	0	0
γ	1	1	0	0	0	0
	1	1	1	0	0	

An entry 1 (resp. 0) in row m, column n of Table 1 means that a voter of type n agrees with party Red (resp. Black) on issue m. Each voter votes for the party with which he agrees on most issues; i.e. types A, B, C vote for the Red party, and types D, E vote for the Black party. This gives the Reds a 60% majority in parliament, a Red government is formed, and all issues are decided according to the Red ideology. However, if referenda were held on each of the issues separately, then, on every single issue a 60% majority of the voters would support the Black platform! Thus, the democratically elected Red government decides against the will of a majority of the population in all issues.

Example 1.1.b. Apportionment Problem: how by redrawing constituency boundaries one can change the composition of parliament. This is simply a re-interpretation of Table 1 (Deb and Kelsey 1987, p 169). Let the cells in the matrix represent individuals and the entries in the cells denote the party they support. Then the Black party controls all the seats if there are three (row) constitutencies α, β, γ; but it becomes a minority party in the parliament if there are five constituencies A, B, C, D, E.

Example 1.1.c. Federalism vs. Centralism: let the cells in the matrix represent individuals as in Example 1.1.b, but let the entries in the matrix denote their position on a *single* issue, say a proposed constitutional amendment (1 = pro, 0 = contra). The nation is partitioned into five *states* A, B, C, D, E. State n supports the amendment if the majority of voters in state n supports it. Then under a "federalist" constitution the amendment is passed because a majority of the states (A, B, C) supports it; but in a nationwide ("centralist") referendum it is defeated by a 9:6 majority. Obviously the same structure can also be interpreted in terms of "indirect" vs. "direct" democracy.

Example 1.2. Exam Paradox. A student (an Economics major) has to take courses in five different *subjects* (Economics, Business Administration, Mathematics, Statistics, Law); and in each subject there are four exams (say one per year). There are five possible *grades:* 1, 2, 3, 4, 5 (where 1 is the best and 5 is the worst (failing) grade). At the end of the four years, a *subject grade* is determined for each subject by forming the average of the four corresponding course grades and then rounding to the nearest integer (ties are broken in the student's favour, i.e. 1.5 gives 1, etc.); and an *overall grade* is determined by forming the average of the five subject grades and rounding again, in the same way as before. Consider two students, one male and one female. His grades are shown in Table 2m, and her grades are shown in Table 2f.

Table 2m. His grades

Subject	Course grades				Average	Subject grade
Economics	1	1	2	2	1.5	1
Business	1	1	2	2	1.5	1
Mathematics	1	1	2	2	1.5	1
Statistics	2	2	3	3	2.5	2
Law	2	2	3	3	2.5	2

Table 2f. Her grades

Subject	Course grades				Average	Subject grade
Economics	1	1	1	1	1.0	1
Business	1	1	1	1	1.0	1
Mathematics	1	1	2	3	1.75	2
Statistics	1	1	2	3	1.75	2
Law	1	2	2	2	1.75	2

His average of subject grades is 1.4, resulting in overall grade 1, and her average of subject grades is 1.6, resulting in overall grade 2. If we look at the 20 basic grades in the various courses, he has six 1's, ten 2's, and four 3's, whereas she has thirteen 1's, five 2's, and only two 3's, i.e. her course grades are *much* better than his. If we formed the average of all course grades directly and then rounded to obtain an alternative overall grade, he would get only 2, but she would get 1, exactly the opposite of the former procedure.

Clearly the situation described in the Exam Paradox is not limited to academic examinations, but will occur whenever "candidates" are to be ranked according to several different criteria ("subjects") and an overall ranking is also desired. Examples of this sort can be found in sports competitions (athletes competing in several disciplines each with several contests), intelligence tests (various "dimensions" of intelligence, each measured by several tasks), opinion polls, etc. The same problem arises in consumer reports (e.g. in Germany those published by "Stiftung Warentest") where similar products of different firms (cars, TV sets,...) are ranked by grades ("very good", "good", "satisfactory", etc.) and these (overall) grades are based on (partial) grades for things like technical standard, reliability, design,..., each of which is in turn an aggregate of more basic properties (e.g. the technical standard of a car depends on the technical standard of the engine, the brakes, the body, etc.).

Observe also that the basic data (course grades) need not really reflect the "candidate's" performance in different "courses", but may equally well reflect his performance in the same "course", *as judged by different observers* (jurors, voters). This interpretation takes us back to the Theory of Social Choice or Voting Theory discussed in Example 1.1, where now the entries in the matrix are no longer restricted to be dichotomous (0 or 1), but can take values in an arbitrary ordered set (e.g. $1, 0, -1$ for pro, abstain, contra; cf. Example 3.1).

It is the purpose of this paper to exhibit the general abstract structure underlying these (and many similar) examples and to prove that "paradoxa" of the type just shown are not due to special features of the examples, but are *unavoidable in principle*, except in trivial or degenerate (in a sense made precise below) cases. To see the general structure more clearly, note that in all our examples we have a set of data (the entries of a matrix) which we *aggregate in two stages*: the entries are partitioned into subsets, and we aggregate first the elements in each subset, and then we aggregate again the subset aggregates. In fact, in each example, we consider two different partitionings: in the *Ostrogorski Paradox*, if a voter decides for which party he is going to vote, he aggregates over the issues (i.e. the entries in a column of Table 1); the resulting distribution of seats in parliament (the bottom row in Table 1) is then aggregated again to form a (Red) government. Alternatively, if a referendum is held on an issue, we aggregate over voters first (i.e. the entries in a row of Table 1); the resulting entries in the rightmost column of Table 1 could then be aggregated again to obtain the "majority ideology" (Black). Similarly, in the *Exam Paradox*, we aggregate first the grades within each subject (i.e. the partitioning is the row partitioning in Table 2), and then we aggregate these subject grades. Alternatively, we could choose the trivial partitioning and aggregate all 20 course grades in one step. (Of course we could also consider the column partitioning of Table 2: this would give four "yearly grades" in the intermediate step, rather than five course grades).

The general feature illustrated by these examples is that the overall aggregate obtained by such a *two-stage aggregation procedure (TSAP)* depends, in general, on the partitioning chosen. That this is *possible* has of course been well known for a long time from standard examples (e.g. the apportionment problem). We ask if it is *unavoidable* for all TSAPs, no matter which partitioning and which types of aggregation rules are employed. This question seems to be of some interest, in view of the importance and widespread use of two- (or multi-)stage procedures in various fields, especially in Voting.

Motivated by the examples given above, we will study two-stage aggregation procedures in which the data are "qualitative" in the sense that they are either dichotomous (as in Example 1.1) or the possible attributes[2] belong to a finite ordered set (as in Example 1.2). Our main result is a complete characterization of the aggregation rules (satisfying certain more or less obvious assumptions), for which the overall aggregate formed by a TSAP does not depend on the partitioning. These aggregation rules are of the following "degenerate" form: There exists a partial ordering ("dominance") of the set of possible attributes

[2] Note that the "attributes" here are the possible *values* (1, 2,...5) of the variable ("academic ability") with which we are concerned. This terminology follows Wilson (1975) (cf. Sect. 2). It should not be confused with the interpretation, also commonly found in the literature, of the word "attribute" as a *variable* ("sex", "age"...), not the value of a variable. This was pointed out by a referee.

such that the aggregate over any collection of data is equal to the supremum (with respect to dominance) of the various attributes occurring in the data, regardless of the distribution (relative frequencies) of the attributes. In particular, in the dichotomous case (which is formally a special case of the finite ordered set case), one of the two possible attributes must be dominant. This means that a person with this attribute has veto power (i.e. in the voting context, degeneracy corresponds to the unanimity principle).

This result is not restricted to the matrix case with row and column partitionings, but covers arbitrary partitionings of arbitrary sets. The present paper generalized (although the setup is slightly different) the results obtained by Deb and Kelsey (1987) for the matrix case with dichotomous variables and majority rules to general two-stage aggregation procedures for finite totally ordered attribute spaces. Although we treat only two-stage procedures, it should be obvious that our results are true for multi-stage procedures as well.

Remark. Our focus on "qualitative" or "discrete" data is justified by the observation that for "quantitative" or "continuous" data the paradoxa illustrated above can often be avoided without difficulty. For example, if the set of possible grades is an interval of the real numbers, then simply forming averages at all stages resolves the exam paradox[3]. (Loosely speaking, if "averaging" is legitimate, then, by Fubini's Theorem on the interchangeability of the order of integration, it does not matter whether one integrates "first over rows and then over columns" or vice versa).

Literature. The Ostrogorski paradox was given its name by Daudt and Rae (1976) because Moise Ostrogorski had already discussed related problems at the turn of the century[4]. A similar example (with an (11×11)-matrix) was presented independently by Anscombe (1976). Further discussions of the Ostrogorski paradox can be found in Bezembinder and Van Acker (1985); see also Wagner (1983, 1984); Kelly (1989). The Exam paradox is analyzed in Nermuth (1989). The problem of aggregating n rows (of "course grades") in an $n\times m$ matrix into a summary row (of "subject grades"), where the entries are *quantitative* data in the sense that they belong to an algebraic field, is considered in Rubinstein and Fishburn (1986). The connection between problems of aggregation and social choice is already emphasized in Wilson (1975). Aggregation over different subgroups also plays a role in certain statistical paradoxa like Simpson's paradox, and voting paradoxa (Bös 1979). The formal structure of multi-stage majority decision is analyzed (from a different viewpoint) in Murakami (1966), Fishburn (1971). Although our results apply only to totally ordered attribute spaces, it may also be noted that the classical problems of aggregating complex objects like preference relations or equivalence relations (classifications) can also be treated

---

[3] In our Example 1.2, the five grades are of course *ordinal* and could equally well have been denoted by A, B, C, D, F. Our forming of "averages" was merely a convenient way to describe the rule by which compound or aggregate grades are determined (1.5 is *not* a grade). Alternatively, we could have stated that "$AABB$ gives A", "$AABC$ gives B", etc.

[4] "Le Parlement ... reflète très insuffisamment l'opinion du pays, les élections ... sont plutôt des votes de confiance en faveur du parti au pouvoir ... que des consultations sur les grands problèmes du jour. Ceux-ci sont toujours confondus dans les programmes, de sorte qu'il est impossible de démêler le vote: certains électeurs se décident en raison d'une mesure législative, d'autres en raison d'une autre mesure,... et l'on ne sait jamais si telle ou telle mesure a véritablement réuni la majorité des suffrages et quelle majorité" (Ostrogorski 1912, p 682).

by order theoretic methods, due to the observation that such attribute spaces can themselves be endowed with some kind of partial (not total) order (making them into lattices or semilattices); cf. Barthélemy et al. (1986), Monjardet (1990) and references given there.

In Sect. 2 and Sect. 3 we develop the necessary concepts and state our results. All proofs are collected in Sect. 4. Section 5 contains concluding remarks.

2. Aggregators

Following the notation of Wilson (1975), let A be a finite set of *attributes*, containing at least two elements, and let P be a finite nonempty set, called the *population*. Such a pair (A, P) is said to be a *domain*. Let $X = A^P$ denote the set of all functions from P into A. Each element $x \in X$ is an *assignment* of attributes to the members of the population; we write $x = (x_v)_{v \in P}$ where $x_v = x(v) \in A$ is the attribute assigned by x to person $v \in P$. For a subset $k \subset P$, we denote by x_k the restriction of x to k, i.e. $x_k = (x_v)_{v \in k} \in A^k$. An *aggregator* for the domain (A, P) is a function $f: X \to A$, which for each assignment $x = (x_v)$ of attributes to the members of the population prescribes a summary attribute or *aggregate* $f(x)$ for the population as a whole. We assume that the attribute space A is *totally ordered* by an order[5] $\geqq$, with associated strict order $>$. Instead of $a \geqq b$ or $a > b$ we write also $b \leqq a$ or $b < a$ (read: "a is greater than or equal to b", etc.).

We wish to describe certain properties which a "reasonable" aggregator f should have. To this end, we define a partial order on X, also denoted by $\geqq$, as follows $(x, x' \in X)$:

$$x \geqq x' \Leftrightarrow x_v \geqq x'_v \; \forall v \in P \; .$$

Moreover, denote by Π the group of all permutations of the elements of P. For $\pi \in \Pi$, $x \in X$ we define $\pi(x) \in X$ by

$$(\pi(x))_v := x_{\pi(v)} \; \forall v \in P \; .$$

Definition 2.1. An aggregator f for the domain (A, P) is called

(i) Paretian if for $\forall a \in A$, $x \in X$: $(x_v = a \; \forall v \in P) \Rightarrow f(x) = a$
(ii) symmetric if for $\forall x \in X$, $\pi \in \Pi$: $f(\pi(x)) = f(x)$
(iii) monotone if for $\forall x, x' \in X$: $x \geqq x' \Rightarrow f(x) \geqq f(x')$.

If f is Paretian, symmetric, and monotone, then f is called admissible.

For $a, b \in A$ we denote by $[a, b] := \{c \in A \mid a \leqq c \leqq b \text{ or } b \leqq c \leqq a\}$ the *interval* of attributes between a and b (inclusive). We define the *span* of an assignment $x \in X$ as the interval between the smallest and the largest value of x, i.e.

[5] A *binary relation* R on a set B is a subset $R \subset B \times B$. We write a Rb iff $(a, b) \in R$. A binary relation R on B is a *partial order* if for all $a, b, c \in B$: *(i)* aRa, *(ii)* aRb and $bRa \Rightarrow a = b$, *(iii)* aRb and $bRc \Rightarrow aRc$ (transitivity). A partial order R is a *(total) order* if for all $a, b \in B$ *(iv)* either aRb or bRa. Let R be a partial order on B. The associated *strict partial order* $\bar{R}$ on B is defined by *(v)* $a\bar{R}b \Leftrightarrow aRb$ and not bRa. An element $d \in C \subset B$ is called R-maximum [resp. R-minimum] of the subset $C \subset B$ iff dRc for $\forall c \in C$ [resp. cRd for $\forall c \in C$]. By *(ii)*, a set C has at most one R-maximum [resp. R-minimum]; it is denoted by $d = \max_R(C)$ [resp. $d = \min_R(C)$], if it exists. If $\mathscr{R}$ is a collection of partial orders on B, then the intersection $\bigcap_{R \in \mathscr{R}} R \subset B \times B$ is also a partial order on B.

$\text{span}(x) := [\min_{\geqq}(x(P)),\ \max_{\geqq}(x(P))]$. Clearly, if f is admissible, then $f(x) \in \text{span}(x)\, \forall x \in X$.

Definition 2.2. An aggregator f for the domain (A, P) is called degenerate if there exists a partial order R on A such that

$$f(x) = \max_R(\text{span}(x)) \quad \text{for } \forall x \in X = A^P\ . \tag{2.1}$$

If (2.1) holds, we say that f is degenerate with R.

Intuitively, a degenerate aggregator f works as follow. First, order the attributes according to some criterion R ("importance") in such a way that every interval $[a, b]$ contains an R-maximal ("most important") element[6]. Then, for every assignment $x = (x_v)$, simply pick the "most important" attribute in the span of x and assign it to the population as a whole. In particular, the aggregate attribute $f(x)$ depends only on the span of x, but not at all on the distribution of the individual attributes x_v within this span[7]. This is clearly an undesirable feature.

We conclude this section with a characterization of degenerate aggregators. First we have:

Lemma 2.1. *If f is degenerate, then f is admissible.*

Let f be an admissible aggregator for a domain (A, P). We say that f (or P) is *trivial* iff $|P| = 1$ (the population contains only one member). Clearly, a trivial aggregator is degenerate with every partial order R on A. Assume therefore for the rest of this section that f is *nontrivial*, i.e. $|P| \geqq 2$. Let $[a, b] \subset A$ be an interval, $d \in [a, b]$. We say that d is an *f-dominant attribute* over $[a, b]$ if

$$f(x) = d \quad \text{for } \forall x \in X \quad \text{with} \quad d \in \text{span}(x) \subset [a, b]\ . \tag{2.2}$$

Clearly there exists at most one f-dominant attribute over any given interval $[a, b]$, it is denoted by $d = d_f(a, b)$, if it exists. We define a binary relation dom_f on A, called the *dominance relation* associated with f, by the condition $(a, b \in A)$:

$$a\, \text{dom}_f\, b \Leftrightarrow a = d_f(a, b)\ . \tag{2.3}$$

If $a\, \text{dom}_f\, b$, we say that *a f-dominates b*. This is equivalent to the following statement (cf. Lemma 4.2 (i)): if only one person has attribute a, and all others have attribute b, then the aggregate is a. We have the following characterization of degeneracy:

---

[6] Apart from this requirement, the (partial) "importance" ordering R need have nothing to do with the given fixed total order $\geqq$ on A. For Example (cf. Example 3.1), if $A = \{$*pro, abstain, contra*$\}$, with the natural order $\geqq$ given by *pro* > *abstain* > *contra*, a degenerate aggregator could be defined by the partial order R which makes *abstain* more important than both *pro* and *contra*. Then the society as a whole would be *pro* [resp. *contra*] iff the vote is unanimously *pro* [resp. unanimously *contra*], and in all other cases society would *abstain* (i.e. make no decision). Alternatively, we could choose R' which makes *contra* more important than *pro* and *pro* more important than *abstain*. Then the society as a whole would be *contra* whenever at least one person votes *contra*, and otherwise it would be *pro* (unless everybody *abstains*, of course).

[7] It is possible, though for an aggregator f to depend only on the span of x without being degenerate. For example, take $P = A = \{1, 2, 3, 4\}$, and for $x \neq (a, a, a, a)$ let $f(x) = 3$ iff $4 \in \text{span}(x)$, and $f(x) = 2$ otherwise. Then $f(1, 1, 1, 4) = 3$, but $f(2, 2, 3, 3) = 2$, i.e. f is not degenerate (if it were, we should have both $3R2$ and $2R3$, which is impossible since $2 \neq 3$).

Lemma 2.2. *Let f be an admissible nontrivial aggregator. Then f is degenerate if and only if every interval contains an f-dominant attribute. More precisely,*

(i) if f is degenerate with the partial order R, then $\max_R([a,b])=d_f(a,b)$ for every interval $[a,b]\subset A$. *Conversely,*
(ii) if every interval $[a,b]$ contains an f-dominant attribute $d_f(a,b)$, then dom_f *is a partial order and f is degenerate with* dom_f.

Remark 2.1. If f is degenerate with R, then $R\supset \mathrm{dom}_f$ (as subsets of $A\times A$); i.e. dom_f is the "smallest" partial order with which f is degenerate. In particular, if f is degenerate, dom_f could also be defined as the intersection of all partial orders with which f is degenerate.

Remark 2.2. In the dichotomous case, where A contains only two elements, say $A=A_2:=\{yes, no\}$, w.l.o.g. *yes* > *no*, monotonicity means that the aggregator f is a *majority rule* of the following form: there exists a number p_f, $1\leqq p_f\leqq |P|$, such that $f(x)=yes$ if and only if the vector x contains at least p_f "*yes*" components. Degeneracy means that $p_f=|P|$ (if *no* dom_f *yes*) or $p_f=1$ (if *yes* dom_f *no*), i.e. a proposal P_1 (resp. its negation $P_0=\neg P_1$) is adopted if and only if the vote is unanimous.

3. Two-stage aggregation procedures

We now introduce *two-stage aggregation procedures* of the following form. Let $(A,\geqq)$ be a fixed totally ordered attribute space, as in Sect. 2. Let N be a nontrivial population, referred to as the *whole population*, and let I be a partitioning of N. Each cell $i\in I$ is a subset of N, and is called a *subpopulation* or *subgroup*. An assignment $x=(x_\nu)_{\nu\in N}\in X=A^N$ for the whole population N defines an assignment $x_i:=(x_\nu)_{\nu\in i}\in A^i$ for every subpopulation $i\in I$; we write also $x=(x_i)_{i\in I}$. Now instead of aggregating the attributes x_ν over the whole population N all at once, we aggregate first over each subpopulation i, thus obtaining a summary attribute y_i for each subpopulation i; and only in a second step we aggregate the y_i's into a summary attribute for the whole population.

Definition 3.1. A two-stage aggregation procedure (TSAP) F for the domain (A,N) is given by a list $F:=(I,(f_i)_{i\in I},f_I)$ such that I is a partitioning of the whole population N, not all $i\in I$ are singletons, and for $\alpha\in I_0:=I\cup\{I\}$ the function $f_\alpha: A^\alpha\to A$ is an admissible aggregator for the domain (A,α)[8].

We define the function $\bar f=(f_i)_{i\in I}: A^N\to A^I$ by $\bar f(x)=(f_i(x_i))_{i\in I}$. Let $x=(x_i)_{i\in I}\in A^N$ be an assignment. For each $i\in I$, the *first-stage aggregator* f_i associates with $x_i=(x_\nu)_{\nu\in i}\in A^i$ an attribute $y_i=f_i(x_i)\in A$. This gives a vector of subpopulation aggregates $y=(y_i)_{i\in I}=\bar f(x)\in A^I$. The *second-stage aggregator* f_I associates with $y\in A^I$ a summary attribute for the whole population, denoted by $f_0(x):=f_I(y)=f_I(\bar f(x))\in A$. Of course the function $f_0=f_I\circ\bar f: A^N\to A$ thus defined is an aggregator for the domain (A,N). It is called the aggregator *induced* by the two-stage aggregation procedure F. The functions f_α, $\alpha\in I_0$, are called the

[8] More explicitly, if $\alpha=i\in I$, then $f_\alpha=f_i$ is an aggregator for the domain (A,i); and if $\alpha=I$, then $f_\alpha=f_I$ is an aggregator for the domain (A,I).

intermediate aggregators of F. A TSAP $F=(I,(f_i)_{i\in I},f_I)$ is called *trivial* if the partitioning I is trivial, i.e. if $I=\{N\}$[9].

When is the induced aggregator f_0 admissible? First we note that if F is trivial, then f_0 coincides with the (only) first-stage aggregator f_N, the second-stage aggregator $f_I=\mathrm{id}_A$ being trivial. In this case, the procedure F really has only one stage and f_0 is always admissible. If F is not trivial, we have

Theorem 3.1. *The aggregator f_0 induced by a nontrivial two-stage aggregation procedure $F=(I,(f_i)_{i\in I},f_I)$ for a domain (A,N) is admissible if and only if there exists a partial order R on A such that all the intermediate aggregators of F are degenerate with R. In this case, f_0 is also degeneate with R.*

Thus, except in trivial or degenerate cases, the aggregator f_0 induced by a two-stage aggregation procedure is not admissible. Since by definition f_0 is Paretian and monotone, and also symmetric with respect to permutations which leave the subgroups $i\in I$ invariant, it must be the case that f_0 is not symmetric with respect to some permutation of attributes *across* subgroups. In other words, the aggregate value $f_0(x)$ depends not only on the distribution of attributes in the whole population, but it depends also on how the attributes are distributed across the various subgroups. This phenomenon is of course well known for many familiar examples (e.g. two-stage voting procedures with simple or qualified majority rules like in the Ostrogorski paradox). Theorem 3.1 states that *all* nondegenerate two-stage procedures suffer from this defect. Note also that we make no assumptions on the sizes of the various subgroups.

Example 3.1. Representative majority decision (direct vs. indirect democracy). Adapting the terminology of Murakami (1966) to our framework, let $A=A_3:=\{1,0,-1\}$, where $1,0,-1$ are interpreted as *pro, abstain, contra*, respectively (cf. footnote 6). An aggregator f for a domain (A_3,P) is called a *majority voting-operation* if

$$f(x)=\mathrm{sign}\left(\sum_{\nu\in P}x_\nu\right)\quad\text{for }\forall x\in X\ . \tag{3.1}$$

The aggregator f_0 induced by a nontrivial TSAP $F=(I,(f_i)_{i\in I},f_I)$ for a domain (A_3,N) is called a *representative system* if all the intermediate aggregators of F are majority voting operations. Since majority voting operations are not degenerate, Theorem 3.1 implies:

Corollary 3.1. *A representative system cannot be symmetric.*

Next consider a TSAP $F=(I,(f_i)_{i\in I},f_I)$ with induced aggregator f_0 as before, and let $G=(J,(g_j)_{j\in J},g_J)$ be another TSAP for the same domain (A,N), with partitioning J, subgroups $j\in J$, intermediate aggregators $g_j:A^j\to A$, $j\in J$, $g_J:A^J\to A$; and put $\bar g=(g_j)_{j\in J}$, $g_0=g_J\circ\bar g$.

Definition 3.2. We say that F and G are consistent if they induce the same aggregator, i.e. if $f_0(x)=g_0(x)$ for $\forall x\in X=A^N$.

[9] The case $I=\{\{\nu\}\mid\nu\in N\}$, which we have excluded by definition, is obviously isomorphic to the case $I=\{N\}$ and need not be considered separately.

Intuitively, the two partitionings I and J correspond to different criteria for classifying the whole population into subgroups. Examples are given in the Introduction. It is possible in general (though none of our examples has this feature) that a subgroup $i \in I$ formed according to one criterion coincides with a subgroup $j \in J$ formed according to the other criterion, i.e. that $i = j \in I \cap J$. Such a subgroup will be called a *common* subgroup. If $I = J$, then all subgroups are common. The next theorem states, roughly, that consistency prevails if and only if the aggregators for common subgroups coincide, and all other intermediate aggregators are degenerate with the same partial order R. We denote the collection of common subgroups by $C := I \cap J$, and define also $I_0 := I \cup \{I\}$, $J_0 := J \cup \{J\}$, $C_0 := I_0 \cap J_0$.

Theorem 3.2. *Two two-stage aggregation procedures* $F = (I, (f_i)_{i \in I}, f_I)$ *and* $G = (J, (g_j)_{j \in J}, g_J)$ *for the same domain* (A, N) *are consistent if and only if*

(i) $f_\alpha = g_\alpha$ *for all* $\alpha \in C_0$*, and*
(ii) there exists a partial order R on A such that all other intermediate aggregators f_α, g_β, $\alpha \in I_0 \setminus C_0$, $\beta \in J_0 \setminus C_0$, *are degenerate with R.*

It is well known from examples that the outcome of a two-stage aggregation procedure can depend on "how the subgroups are defined" (e.g. if the boundary between two voting districts is changed, this may affect the outcome of an election). Theorem 3.2 says that this dependence is not caused by special aggregation rules, but exists for every nondegenerate two-stage procedure. Note also that Theorem 3.1 is formally a special case of Theorem 3.2 (choose F nontrivial and G trivial in Theorem 3.2).

4. Proofs

4.1. Notation

Let (A, P) be a domain, $X = A^P$. We define an equivalence relation $\sim$ on X by

$$x \sim x' \Leftrightarrow \exists\, \pi \in \Pi \quad \text{with} \quad x' = \pi(x) \ .$$

If $x \sim x'$ we say that x and x' *differ only by a permutation.* The equivalence class of $x \in X$ is denoted by $\langle x \rangle := \{x' \in X \mid x' \sim x\}$, and the set of all equivalence classes, the quotient set, by $\langle X \rangle := X/\sim$. We define a partial order $\geqq$ on $\langle X \rangle$ by

$$\langle x \rangle \leqq \langle x' \rangle \Leftrightarrow \exists\, x'' \in \langle x' \rangle \quad \text{with} \quad x \leqq x'' \text{ (in } X) \ .$$

If f is an admissible aggregator for (A, P), then, by symmetry, $f(x) = f(x')$ whenever $x \sim x'$, i.e. we can consider f also as a function defined on $\langle X \rangle$. Clearly this function is also monotone.

Let $a, b, c \in A$ be attributes, and let $r \geqq 0$, $s \geqq 0$ be numbers with $r + s \leqq |P|$. Then $\langle a_r b_s c \rangle \in \langle X \rangle$ denotes the equivalence class of all assignments in which exactly r persons have attribute a, exactly s persons have attribute b, and all the others (if any are left) have attribute c. Similarly, $\langle a_r b \rangle$ denotes the equivalence class of all assignments in which r persons have attribute a and the rest have attribute b ($0 \leqq r \leqq |P|$); and $\langle a \rangle$ denotes the equivalence class of the assignment in which all persons have attribute a. By abuse of language, but without danger of confusion, we will henceforth omit the angle brackets and identify X with $\langle X\rangle$, whenever convenient. Thus "$x=a_r b$" means that x is an assignment with $\langle x\rangle=\langle a_r b\rangle$, "$x=a$" means that all components of x are equal to a, etc.

Let $F=(I,(f_i)_{i\in I}, f_I)$, $G=(J,(g_j)_{j\in J}, g_J)$ be two TSAPs for the same domain (A,N). Throughout this section. we will reserve certain symbols to denote elements of certain sets, viz. $a,b,c,d,e\in A$, $\pi\in\Pi$, where Π is the group of permutations of N, $x\in X=A^N$, $x_i\in A^i$ $(i\in I)$, $x_j\in A^j$ $(j\in J)$, $y\in Y:=A^I$, $z\in Z:=A^J$, and similarly for a^k, e', $\bar{y}$ etc. Thus if we write $y=a_1^k b$ it will be understood that y is a vector of subpopulation aggregates with one component equal to a^k and $|I|-1$ components equal to b; or if we write $x_i'=c$ it will be understood that x_i' is a vector with $|i|$ components, all equal to $c\in A$.

Let $k,l\in I$ be two subgroups. A permutation π of N is called a *kl-switch* if it interchanges only one person in k with one person in l, i.e. if $\exists\, v_1\in k$, $v_2\in l$ such that $\pi(v_1)=v_2$, $\pi(v_2)=v_1$, $\pi(v)=v$ otherwise.

We say that $(c,d)\in A\times A$ can be *kl-switched* to $(c',d')\in A\times A$, written $(c,d)\overset{kl}{\longleftrightarrow}(c',d')$, if there exists an assignment x and a *kl-switch* π, such that for $x'=\pi(x)$:

$$(f_k(x_k), f_l(x_l))=(c,d)\ ,\quad (f_k(x_k'), f_l(x_l'))=(c',d') \tag{4.1.1}$$

(i.e. by switching two persons $v_1\in k$, $v_2\in l$ the subpopulation aggregates for the two subpopulations k and l change from (c,d) to (c',d')).

Obviously $(c,d)\overset{kl}{\longleftrightarrow}(c',d')$ iff $(c',d')\overset{kl}{\longleftrightarrow}(c,d)$. We say that $y\in Y=A^I$ can be *directly kl-switched* to $y'\in Y$, written $y\overset{kld}{\longleftrightarrow}y'$, if there exists an assignment x and a kl-switch π such that

$$y\sim\bar{f}(x)\ ,\quad y'\sim\bar{f}(\pi(x)) \tag{4.1.2}$$

(i.e., after suitable rearrangement of their components, the two vectors y, y' differ only in the two components corresponding to the subgroups k, l; and this difference is due to a kl-switch).

Finally, we say that $y\in Y$ can be *kl-switched* to $y'\in Y$, written $y\overset{kl}{\longleftrightarrow}y'$, if there exists a finite sequence $(y^t)_{t=0}^T$, $T\geqq 1$, in Y such that $y=y^0\overset{kld}{\longleftrightarrow}y^1\overset{kld}{\longleftrightarrow}y^2\overset{kld}{\longleftrightarrow}\cdots\overset{kld}{\longleftrightarrow}y^T=y'$ (i.e. y can be transformed into y' by a finite sequence of direct kl-switches). Obviously the relation $\overset{kl}{\longleftrightarrow}$ is an equivalence relation on Y.

4.2. *Degeneracy criterion*

Let f be an admissible nontrivial aggregator for a domain (A,P) and let dom_f be the associated dominance relation.

Lemma 4.1. *For* $\forall\, a,b\in A$:

(i) $a\,\mathrm{dom}_f\,a$

(ii) $a\,\mathrm{dom}_f\,b$ *and* $b\,\mathrm{dom}_f\,a\Rightarrow a=b$

(iii) assume $c\in[a,b]$. *Then* $a\,\mathrm{dom}_f\,b\Rightarrow a\,\mathrm{dom}_f\,c$.

Proof. Obvious from the definitions.

Lemma 4.2. *For* $\forall a, b, c \in A$:
(i) $a \operatorname{dom}_f b \Leftrightarrow f(a_1 b) = a$
(ii) $d = d_f(a, b) \Leftrightarrow d \operatorname{dom}_f c$ *for* $\forall c \in [a, b]$.

Proof. Obvious from monotonicity and Lemma 4.1.

Proof of Lemma 2.1. Let f be degenerate with R. It is obvious from (2.1) that f is Paretian and symmetric. To show that f is monotone, choose $x \leqq x'$ and write $\operatorname{span}(x) = [a, b]$, $\operatorname{span}(x') = [a', b']$, where $a \leqq a'$, $b \leqq b'$. Then $d: = f(x) = \max_R([a, b])$, $d' := f(x') = \max_R([a', b'])$; and if $d \geqq d'$, then d, d' must both belong to $[a, b] \cap [a', b']$, hence $d = d'$ (since both are R-maxima). Therefore $d \leqq d'$, and f is monotone. Q.E.D.

Proof of Lemma 2.2. Necessity is obvious. To prove sufficiency, we must show first that dom_f is a partial order. By Lemma 4.1, it suffices to show transitivity. Let $a \operatorname{dom}_f b$ and $b \operatorname{dom}_f c$ (∗). If $b \notin [a, c]$, then either $a \in [b, c]$ or $c \in [a, b]$. By Lemma 4.1, (∗) implies: if $a \in [b, c]$, then $b \operatorname{dom}_f a$, hence $b = a$, hence $a \operatorname{dom}_f c$; and if $c \in [a, b]$, then also $a \operatorname{dom}_{f}, c$. If $b \in [a, c]$, then $[a, c] = [a, b] \cup [b, c]$. Let $d := d_f(a, c)$. If $d \in [b, c]$, then $d = d_f(b, c) = b$ by (∗), hence $a \operatorname{dom}_f d$, hence $d = a$. If $d \in [a, b]$, then $d = d_f(a, b) = a$. In either case, $a \operatorname{dom}_{f}, c$. This proves that dom_f is a partial order.

It remains to show that f is degenerate with dom_f, i.e. $d_f(a, b) = \max_{\operatorname{dom}_f}([a, b])$. Choose x with $\operatorname{span}(x) = [a, b]$ (this is possible because f is nontrivial) and let $f(x) = d_f(a, b) =: d$. By Lemma 4.2 (ii), $d \operatorname{dom}_f c$ for $\forall c \in [a, b]$, i.e. (2.1) is satisfied. Q.E.D.

4.3. *Proof of Theorem 3.1*

Let $F = (I, (f_i)_{i \in I}, f_I)$ be a nontrivial TSAP for a domain (A, N), with $\bar{f} = (f_i)_{i \in I}$, and induced aggregator f_0.

Sufficiency. It is obvious from the definitions that if all intermediate aggregators are degenerate with the same partial order R, then f_0 also has this property. By Lemma 2.1, f_0 is admissible.

Necessity. Assume that the induced aggregator $f_0 = f_I \circ \bar{f}$ is admissible. This implies that switching does not change the total aggregate $f_0(x) = f_I(y)$:

Claim 1: Let $k, l \in I$. Then

$$y \stackrel{kl}{\longleftrightarrow} y' \Rightarrow f_I(y) = f_I(y') \ . \tag{4.3.1}$$

Proof of Claim 1: Passing from y to y' involves only two kinds of operations: permutations of y (which do not change $f_I(y)$ because f_I is symmetric) and permutations of x, which do not change the aggregate because f_0 is symmetric by assumption. This proves Claim 1. We have to show that there exists a partial order R on A such that all intermediate aggregators $f_\alpha, \alpha \in I_0 = I \cup \{I\}$, are degenerate with R. Assume, indirectly, that this is not the case. Then we can find a nontrivial first-stage aggregator $f_k, k \in I$, such that f_k and f_I are not both degenerate with the same R. By Lemma 2.2, this implies that there exists an interval $[a, b]$ such that

$$f_k \text{ and } f_I \text{ do not have the same dominant attribute over } [a, b] \ . \tag{4.3.2}$$

Note that, by assumption, f_l is nontrivial. Choose an interval $[a,b]$ of *minimal* length $|[a,b]|$ such that (4.3.2) holds. Obviously $|[a,b]| \geqq 2$, w.l.o.g. $a<b$. Then f_k and f_I have the same dominant attribute over every shorter interval, and we can define:

$$d_{f_k}(a,b-1)=d_{f_I}(a,b-1)=:\bar{a} \ , \tag{4.3.3a}$$

$$d_{f_k}(a+1,b)=d_{f_I}(a+1,b)=:\bar{b} \ , \tag{4.3.3b}$$

where $a+1$ [resp. $b-1$] denotes the smallest attribute greater than a [resp. the largest attribute smaller than b].

Claim 2: Either $a=\bar{a}$ or $b=\bar{b}$. Therefore $a \leqq \bar{a} < \bar{b} \leqq b$.

Proof. Assume, indirectly, that $a<\bar{a}$, $\bar{b}<b$. Then for $\alpha=k, I$: $\bar{a}\in[a+1,b]$, hence $b \operatorname{dom}_{f_\alpha} \bar{a}$, and also $\bar{b}\in[a,b-1]$, hence $\bar{a} \operatorname{dom}_{f_\alpha} \bar{b}$. By Lemma 4.1, this implies $\bar{a}=\bar{b}$, hence by Lemma 4.2 $\bar{a} \operatorname{dom}_{f_\alpha} c$ for all $c\in[a,b]$, $\alpha=k, I$, contradicting (4.3.2). This proves Claim 2.

Claim 3: Let $\alpha=k$ or $\alpha=I$ and assume that an f_α-dominant attribute $d_\alpha=d_{f_\alpha}(a,b)$ exists. Then $d_\alpha=a=\bar{a}$ or $d_\alpha=b=\bar{b}$.

Proof. Fix α and write $d=d_\alpha$. If $a<d<b$, then by Lemma 4.2 (ii) $d=\bar{a}$ and $d=\bar{b}$, hence $\bar{a}=\bar{b}$, contradicting Claim 2. Therefore $d=a$ or $d=b$. Again by Lemma 4.2 (ii), in the first case $d=\bar{a}$ and in the second case $d=\bar{b}$. This proves Claim 3.

Now choose a subpopulation $l\in I$, $l\neq k$ (such an l exists because F is nontrivial). Let us write

$$f_k(a_1 b)=:a^k \ , \quad f_k(b_1 a)=:b^k \tag{4.3.4}$$

This implies

$$\begin{aligned}&a\leqq b^k\leqq a^k\leqq b \quad \text{and either} \quad a<a^k \quad \text{or}\\ &b^k<b \ , \quad \text{w.l.o.g. } a<a^k \ .\end{aligned} \tag{4.3.5}$$

Since f_l is Paretian, there exist numbers r, s, $0\leqq r$, $s\leqq |l|-1$, such that

$$f_l(a_{r+1} b)=a<f_l(a_r b)=:a^l \ , \quad f_l(b_{s+1} a)=b>f_l(b_s a)=:b^l \ . \tag{4.3.6}$$

We claim that the following kl-switches are possble in $A\times A$:

$$(a,b) \overset{kl}{\longleftrightarrow} (b^k,b^l) \tag{4.3.7a}$$

$$(b,a) \overset{kl}{\longleftrightarrow} (a^k,a^l) \ . \tag{4.3.7b}$$

To prove (4.3.7a), choose $x_k=a$, $x_l=b_{s+1}a$; $x_k'=b_1 a$, $x_l'=b_s a$. Then by (4.3.4), (4.3.6) $(f_k(x_k), f_l(x_l))=(a,b)$, and $(f_k(x_k'), f_l(x_l'))=(b^k,b^l)$, i.e. (4.1.1) is satisfied. Moreover, (x_k,x_l) and (x_k',x_l') differ only by a switch (a person $v_1\in k$ with attribute a has switched with a person $v_2\in l$ with attribute b). In a similar way we prove (4.3.7b), choosing $x_k=b$, $x_l=a_{r+1}b$; $x_k'=a_1 b$, $x_l'=a_r b$.

Now we distinguish three cases (we will use (4.3.4), (4.3.5), (4.3.6) repeatedly without explicit reference).

Case 1. If $b^k = b$, then $a^k = b$, and $b \operatorname{dom}_{f_k} a$ by (4.3.4), hence by (4.3.2) it is not true that $b \operatorname{dom}_{f_I} a$, i.e. $f_I(b_1 a) \neq b$. Moreover, by Claim 3, $b = \bar{b}$. By (4.3.7b) we can kl-switch $(b, a) \overset{kl}{\longleftrightarrow} (b, a^l)$ in $A \times A$, hence, by repeatedly applying this switch, $y = b_1 a \overset{kl}{\longleftrightarrow} y' = b_1 a^l$. But $a^l \in [a+1, b]$, hence $f_I(y') = \bar{b} = b \neq f_I(y)$, contradicting (4.3.1).

Case 2. If $a < b^k < b$, we switch (using 4.3) $y = a_1 b_1 b^k \overset{kl}{\longleftrightarrow} y' = b_1^k b_1^l b^k$ and also $y = a_1 b_1 b^k \overset{kl}{\longleftrightarrow} y'' = a_1^k a_1^l b^k$, hence $y' \overset{kl}{\longleftrightarrow} y''$. Then $\operatorname{span}(y') \subset [a, b-1]$, hence $f_I(y') = \bar{a}$ by (4.3.3a); and $\operatorname{span}(y'') \subset [a+1, b]$, hence $f_I(y'') = \bar{b} \neq \bar{a}$ by (4.3.3b) and Claim 2, contradicting (4.3.1).

Case 3. If $a = b^k$, we can switch $(a, b) \overset{kl}{\longleftrightarrow} (a, b^l)$ by (4.3.7a), hence $y = a_1 b \overset{kl}{\longleftrightarrow} y' = a_1 b^l$ with $f_I(y') = \bar{a}$; and by (4.3.7b) we have $y = a_1 b_1 b \overset{kl}{\longleftrightarrow} y'' = a_1^k a_1^l b$ with $f_I(y'') = \bar{b}$, again contradicting (4.3.1). Q.E.D.

4.4. *Proof of Theorem 3.2*

Let $F = (I, (f_i)_{i \in I}, f_I)$, $G = (J, (g_j)_{j \in J}, g_J)$ be two TSAPs for the same domain (A, N).

Sufficiency. If *(i)* and *(ii)* of Theorem 3.2 hold, then obviously $f_0(x) = g_0(x)$ for all $x \in X$, i.e. F and G are consistent.

Necessity. Assume that F and G are consistent, i.e. $f_0(x) = g_0(x)$, $\forall x \in X$. We have to show that *(i)* and *(ii)* of Theorem 3.2 hold.

If F is trivial, then $I = \{N\}$, $I_0 = \{N, I\}$, and the only two intermediate aggregators of F are $f_N = f_0 : A^N \to A$, and $f_I = \mathrm{id}_A : A \to A$ (the identity function). If G is also trivial, then $J = I$, $C_0 = J_0 = I_0$, i.e. *(ii)* is satisfied vacuously, and *(i)* is satisfied because $g_N = g_0 = f_0 = f_N$ by consistency, and $g_J = \mathrm{id}_A = f_I$. If F is trivial, but G is not trivial, then $C_0 = \emptyset$, i.e. *(i)* is satisfied vacuously, and *(ii)* is satisfied by Theorem 3.1, because $g_0 = f_0 = f_N$ is admissible. Assume therefore from now on that both F and G are nontrivial.

Proof of Theorem 3.2 (i). Let $k \in C = I \cap J$ be a common subgroup. Assume indirectly that $f_k \neq g_k$. Then there exists $\bar{x}_k \in A^k$ such that $a := f_k(\bar{x}_k) \neq g_k(\bar{x}_k) =: b$. Since g_J is Paretian, there exists a number r, $0 \leqq r \leqq |J| - 1$, such that

$$g_J(a_r b) \neq a = g_J(a_{r+1} b) \ . \tag{4.4.1}$$

Choose a set $J_r \subset J \setminus \{k\}$ with $|J_r| = r$, and define x, x' by: $x_k = \bar{x}_k$, $x_j = a$ for $j \in J_r$, $x_j = b$ otherwise; and $x'_k = a$, $x'_j = x_j$ otherwise. Then x and x' coincide outside the common subgroup k, and $f_k(x_k) = f_k(x'_k) = a$. This implies $f_0(x) = f_0(x')$. On the other hand, $\bar{g}(x) = a_r b$, $\bar{g}(x') = a_{r+1} b$, hence, by (4.4.1) $g_0(x) = g_J(a_r b) \neq g_J(a_{r+1} b) = g_0(x')$, contradicting consistency. This proves $f_k = g_k$ for $k \in I \cap J$.

If $I=J$, we must also show that $f_I=g_I$. Choose an arbitrary $y\in A^I$, and define $x\in A^N$ by $x_v=y_i$ for $v\in i$, $i\in I$. Then $\bar{f}(x)=\bar{g}(x)=y$, and by consistency: $f_I(y)=f_I(\bar{f}(x))=f_0(x)=g_0(x)=g_I(\bar{g}(x))=g_I(y)$. This proves Theorem 3.2 (i).

Proof of Theorem 3.2 (ii). Assume $I\neq J$ (otherwise there is nothing to prove). Then $C_0=I\cap J=C$, $I_0\setminus C_0=(I\setminus C)\cup\{I\}$, $J_0\setminus C_0=(J\setminus C)\cup\{J\}$, and $I\setminus C\neq\emptyset\neq J\setminus C$. Let $D:=\{k\in(I\cup J)\setminus C \mid |k|\geqq 2\}$ be the collection of all nontrivial noncommon subgroups. We have

$$N=\bigcup_{k\in C\cup D} k \tag{4.4.2}$$

because every $v\in N$ belongs to exactly one cell $i\in I$ and also to one $j\in J$; if $i\neq j$, then one of them must be nontrivial. We distinguish two cases.

Case 1. There are two sets $h,k\in D$ with a nonempty intersection (this is the "normal" case, satisfied in all our examples): $h\cap k\neq\emptyset$, $|h|\geqq 2$, $|k|\geqq 2$. W.l.o.g. we can assume that $k\in I\setminus C$, $h\in J\setminus C$, and h is not a subset of k, i.e. $\exists\, l\in I\setminus C$, $l\neq k$, such that $h\cap l\neq\emptyset$. Then kl-switches do not change the total aggregate:

Claim 1: $y \overset{kl}{\longleftrightarrow} y' \Rightarrow f_I(y)=f_I(y')$.

Proof. Pick $v_1\in h\cap k$, $v_2\in h\cap l$. Switches between v_1 and v_2 do not change $\bar{g}(x)$, hence do not change $f_0(x)=g_0(x)=g_J(\bar{g}(x))$, because v_1, v_2 both belong to the same J-cell h, and g_h is symmetric. The assertion follows as in the proof of (4.3.1). This proves Claim 1. By exactly the same argument as in the proof of Theorem 3.1, Claim 1 implies that

$$\begin{aligned}&f_k \text{ and}\\ &f_I \text{ are both degenerate with the same partial order, say } R\,.\end{aligned} \tag{4.4.3}$$

By Lemma 2.2, we can assume $R=\mathrm{dom}_{f_k}=\mathrm{dom}_{f_I}$. To prove that g_h and g_J also have property (4.4.3), it suffices to show that aRb implies

$$g_h(a_1 b)=g_J(a_1 b)=a\,. \tag{4.4.4}$$

Assume aRb. Pick $v_1\in h\cap k$ and define x by $x_{v_1}=a$, $x_v=b$ otherwise. Put $g_h(a_1 b)=:c$, so that $\bar{g}(x)=c_1 b$. By (4.4.3) and consistency, $f_0(x)=f_I(a_1 b)=a=g_0(x)=g_J(c_1 b)\Rightarrow c=a$. This proves (4.4.4).

Now consider an arbitrary nontrivial subgroup $h'\in D$, w.l.o.g. $h'\in J$. If h' intersects another nontrivial subgroup in D, the above argument shows that $g_{h'}$ is degenerate with R. If h' intersects no nontrivial subgroup in D, it must contain a singleton $k'=\{v'\}\in I (v'\in N)$. Assume aRb and define x by $x_{v'}=a$, $x_v=b$ otherwise. Put $g_{h'}(a_1 b)=:c$, so that $\bar{g}(x)=c_1 b$. By (4.4.3) and consistency, $f_0(x)=f_I(a_1 b)=a=g_0(x)=g_I(c_1 b)\Rightarrow c=a\Rightarrow g_{h'}(a_1 b)=a$, i.e. h' is degenerate with R. This proves Theorem 3.2 in Case 1.

Case 2. The sets in D are pairwise disjoint. Then $C\cup D$ is a partitioning of N, by (4.4.2). Choose $h\in D$, w.l.o.g. $h\in I\setminus C$. Then h is the union of r singletons in J, where $2\leqq r=|h|\leqq |J|-1$. Since h is arbitrary, it suffices to show that the three aggregators f_h, f_I, g_J are all degenerate with the same partial order R. By Lemma 2.2, we have to show that f_h, f_I, g_J have the same dominant attribute over every interval.

Choose $k \in C \cup D$, $k \neq h$, $k \in J$. Then k is the union of s I-cells, where $1 \leqq s \leqq |J| - 1$ (if $k \in C, s = 1$; if $k \in D, s \geqq 2$).

Denote by $I_1 := \{i \in I \mid i \cap h = \emptyset = i \cap k\}$ and $J_1 := \{j \in J \mid j \cap h = \emptyset = j \cap k\}$ the subgroups "outside" h and k. We use the following variant of the "switching" technique. Let $a, b \in A$ be two attributes, w.l.o.g. $a \leqq b$. Consider two assignments x, x' such that $x_h = a, x_k = b$; $x'_h = b$, $x'_k = a$; and $x_l = x'_l$ otherwise ($l \in (C \cup D) \setminus \{h, k\}$). Write $y_{I_1} := (f_i(x_i))_{i \in I_1} \in A^{I_1}$, $z_{J_1} := (g_j(x_j))_{j \in J_1} \in A^{J_1}$. Then, in obvious notation,

$$y := \bar{f}(x) = a_1 b_s y_{I_1} \; , \quad y' := \bar{f}(x') = b_1 a_s y_{I_1} \tag{4.4.5a}$$

$$z := \bar{g}(x) = a_r b_1 z_{J_1} \; , \quad z' := \bar{g}(x') = b_r a_1 z_{J_1} \tag{4.4.5b}$$

and (because $a \leqq b$):

$$y \geqq y' \; , \quad z \leqq z' \; . \tag{4.4.6}$$

Consistency and monotonicity imply

$$f_I(y) = g_J(z) \leqq g_J(z') = f_I(y') \leqq f_I(y) \tag{4.4.7}$$

i.e. all these terms are equal. In particular, we have shown: If $z = a_r b_1 z_{J_1} \in Z = A^J$ is any vector containing (at least) one b and (at least) r a's, then we can replace $r-1$ of these a's by b's, obtaining $z' = a_1 b_r z_{J_1}$, without changing the total aggregate. Starting from $z = b_1 a$, we perform such replacements, repeatedly if necessary, until fewer than r a's are left. This shows that $g_J(b_1 a) = g_J(a_t b)$, where t satisfies $1 \leqq t \leqq r-1$, $|J| - 1 = T(r-1) + t$ for some integer $T \geqq 1$. An analogous argument proves that also $g_J(a_1 b) = g_J(b_t a)$. Moreover, $b_1 a \leqq a_1 b$ and $a_t b \geqq b_t a$ in Z (because $2t \leqq |J|$), hence by monotonicity

$$g_J(a_1 b) = g_J(b_1 a) =: d \; . \tag{4.4.8a}$$

If we choose $y_{I_1} = b, z_{J_1} = b$ in (4.4.5) (this is possible: choose $x_v = b \,\forall\, v \notin h \cup k$), we see from (4.4.7) that $f_I(a_1 b) = g_J(a_1 b)$; similarly, we have also $f_I(b_1 a) = g_J(b_1 a)$; hence, by (4.4.8a),

$$f_I(a_1 b) = f_I(b_1 a) = d \; . \tag{4.4.8b}$$

This implies that if x is any assignment with $\text{span}(\bar{g}(x)) = [a, b]$ or $\text{span}(\bar{f}(x)) = [a, b]$, then $g_0(x) = g_J(\bar{g}(x)) = d = f_I(\bar{f}(x)) = f_0(x)$ (again by monotonicity, because e.g. $a_1 b \geqq \bar{g}(x) \geqq b_1 a$). Now choose $\bar{x}_h \in A^h$ with $\text{span}(\bar{x}_h) = [a, b]$ and define x, x' by $x_h = \bar{x}_h$, $x_i = a$ otherwise; $x'_h = \bar{x}_h, x'_i = b$ otherwise. Then $\text{span}(\bar{g}(x)) = \text{span}(\bar{g}(x')) = [a, b]$, hence $f_0(x) = f_0(x') = d$. Put $c := f_h(\bar{x}_h) \in [a, b]$. Then $\bar{f}(x) = c_1 a$, $\bar{f}(x') = c_1 b$, hence $f_I(c_1 a) = f_I(c_1 b) = d$, hence

$$c = d = d_{f_I}(a, b) \; . \tag{4.4.9a}$$

Next define x, x' by $x_k = d$, $x_j = a$ otherwise; $x'_k = d$, $x'_j = b$ otherwise. Then $\bar{f}(x) = d_s a$, $\bar{f}(x') = d_s b$, $\bar{g}(x) = d_1 a$, $\bar{g}(x') = d_1 b$. This implies, by (4.4.9a) and consistency: $f_0(x) = f_0(x') = d = g_0(x) = g_0(x') = g_J(d_1 a) = g_J(d_1 b)$, hence

$$d = d_{g_J}(a, b) \; . \tag{4.4.9b}$$

Finally, if d is not also f_h-dominant over $[a, b]$, then either $f_h(d_1 a) \neq d$ or

$fh(d_1 b) \neq d$; w.l.o.g. $f_h(d_1 a) =: e \neq d$. Choose a (singleton) $l \subset h$, $l \in J$, and define x by $x_l = d$, $x_j = a$ for $j \subset h, j \neq l$, and $x_j = e$ otherwise. Then $\bar{f}(x) = e$, hence $f_0(x) = e$; and $\bar{g}(x) = d_1 a_{r-1} e$, hence $g_0(x) = g_J(d_1 a_{r-1} e) = d$. This contradicts consistency, and proves that also

$$d = d_{f_h}(a, b) \ . \tag{4.4.9c}$$

By Lemma 2.2, (4.4.9a), (4.4.9b), (4.4.9c) prove the Theorem in Case 2, since $[a, b]$ was arbitrary. Q.E.D.

5. Concluding remarks

Throughout this paper we have assumed that the attribute space A is a priori *totally ordered* by on order $\geqq$, and that admissible aggregators are *monotone* with respect to this order. One might wonder to which extent such an assumption is necessary for our results. That it cannot be dispensed with altogether, can be seen as follows.

Let A be a set *without* an a priori order structure $\geqq$. Call an aggregator f for a domain (A, P) *admissible* if it is Paretian and symmetric (i.e. satisfies conditions (i) and (ii) in Def. 2.1), and call f *degenerate* if it satisfies Def. 2.2, where (2.1) is replaced by

$$f(x) = \sup_R(x(P)) \quad \text{for } \forall x \in X \ . \tag{5.1}$$

With these definitions, Theorem 3.1 and Theorem 3.2 are no longer true, as is shown by the following odd example.

Example 5.1. Let $A = A_2 = \{0, 1\}$. Call an aggregator f for a domain (A_2, P) *odd* if $|P|$ is odd, and $f(x) = 1$ iff an odd number of persons have attribute 1 (i.e. iff $|\{v \in P \mid x_v = 1\}|$ is odd). It is easily seen that an odd aggregator f is admissible and, unless it is trivial, not degenerate. Now let $F = (I, (f_i)_{i \in I}, f_I)$ be a nontrivial TSAP for a domain (A_2, P) such that all intermediate aggregators are odd. Then it is not hard to verify that the induced aggregator f_0 is also odd, hence admissible *and* nondegenerate.

References

Anscombe GEM (1976) On frustration of the majority by fulfillment of the majority's will. Anal 36: 161–168

Barthélemy JP, Leclerc B, Monjardet B (1986) On the use of ordered sets in problems of comparison and consensus of classifications. J Classification 3: 187–224

Bezembinder Th, Acker P Van (1985) The Ostrogorski paradox and its relation to nontransitive choice. J Math Sociol 11: 131–158

Bös D (1979) A voting paradox of fiscal federalism. J Publ Econ 11: 369–382

Daudt H, Rae DW (1976) The Ostrogorski paradox: a peculiarity of compound majority decision. Eur J Polit Res 4: 391–398

Deb R, Kelsey D (1987) On constructing a generalized Ostrogorski paradox: necessary and sufficient conditions. Math Soc Sci 14: 161–174

Fishburn PC (1971) The theory of representative majority decision. Econometrica 39: 273–284

Kelly JS (1989) The Ostrogorski paradox. Soc Choice Welfare 6: 71–76

Monjardet B (1990) Arrowian characterization of latticial federation consensus functions. Working paper P.052, Centre d'Analyse et de Mathématique Sociales, 54 bd Raspail, Paris

Murakami Y (1966) Formal structure of majority decision. Econometrica 34: 709–718
Nermuth M (1989) Aggregation of qualitative data: the exam paradox. Discussion Paper, Fakultät für Wirtschaftswissenschaften, Universität Bielefeld
Ostrogorski M (1912) La démocratie et les partis politiques. Calmann-Lévy, Paris
Rubinstein A, Fishburn PC (1986) Algebraic aggregation theory. J Econ Theory 38: 63–77
Wagner C (1983) Anscombe's paradox and the rule of three-fourths. Theory Decis 15: 303–308
Wagner C (1984) Avoiding Anscombe's paradox. Theory Decis 16: 223–235
Wilson R (1975) On the theory of aggregation. J Econ Theory 10: 89–99