ALLOTMENT METHODS: PROCEDURES FOR PROPORTIONAL DISTRIBUTION OF INDIVISIBLE ENTITIES[2]

by

Aanund Hylland[1]

Working Paper 1990/11

Copyright Aanund Hylland
Published by Handelshøyskolen BI
Norwegian School of Management
P.box 580
N-1301 Sandvika

The Research Center
April 1990

Stiftelsen
Bedriftsøkonomisk Institutt
Norwegian School of Management

ARBEIDSNOTAT / WORKING PAPER ☒
FORSKNINGSRAPPORT / RESEARCH REPORT ☐

Tittel	Nummer
Allotment Methods: Procedures for Proportional Distribution of Indivisible Entities	1990/11
	Antall sider
	141 s.

Forfatter

Aanund Hylland

Fagseksjon	Prosjekt
Institutt for økonomi	

Sammendrag

In various contexts, it is necessary to divide a set of indivisible entities among given units in proportion to a certain criterion. Perfectly proportional distribution cannot be guaranteed, since the units are indivisible. One must look for criteria for making approximately proportional distribution. Allotment methods, which are studied in this paper, are procedures for making such division.

Allotment methods are used in connection with political elections. In party-list systems of proportional representation, such methods are used to distribute the seats among the parties, on the basis of their votes. In any electoral system, similar methods can be used to apportion the seats in an elected assembly among geographical districts, according to their population.

Stikkord (norsk)		Stikkord (engelsk)	
	Beslutningsteori		Decision Theory
	Valgordninger		Electoral Systems

CONTENTS

PREFACE

This paper was written many years ago, when I was a doctoral student at the John F. Kennedy School of Government, Harvard University. It has existed only as occasional photocopies. I have found it convenient to give the paper a more systematic circulation.

Although I have continued working on the issue, I have decided to print the paper exactly as it was written in 1978. Further results are published elsewhere. Reference no. 16 was never finished as a paper with that title, but I have discussed the issues in a number of papers, most of them, however, written in Norwegian.

Sandvika, April 1990

Aanund Hylland

2. INTRODUCTION

In many countries, political elections are based on party-list systems of proportional representation. This is the case in most of Western Europe. The most important aspect of the election result in such a system is the distribution of seats among the parties. Therefore, the most important part of the election procedure is the rules by which this distribution is determined.[3]

This distribution of seats shall be "proportional," although actual electoral systems usually represent a compromise between proportionality and other considerations. If representatives were perfectly divisible, there would be no problems in distributing them proportionally. But representatives are not divided, although one can imagine systems which rely on fractional representatives. Therefore, a method must be sought which approximates proportional distribution.

The same problem arises in a related context. The seats in an elected assembly are often apportioned among geographical districts in proportion to population. Again, exact proportionality is impossible, and an approximation method must be used. (This, of course, holds whether or not the election is conducted by a proportional system.)

A variety of different systems for distributing the seats in proportional elections are in actual use, and many more have been proposed. All of them have shortcomings; for any system, one can construct situations in which that system produces a seemingly unreasonable result.[4] Therefore, pointing out that a method has a certain undesirable property does not necessarily provide a decisive argument against the method. My impression is that political discussion of electoral systems, of which there has been a lot, has a tendency to concentrate on specific examples and the effects of the proposed methods on these examples. I do not claim that such examples are irrelevant, but the debate should not solely be concerned with them.

There are a number of general principles which one can reasonably require that a method for proportional representation satisfy. The discussion should, at least in part, be concerned with such principles. That is, one should ask questions like: Which are the requirements a method ideally should satisfy? Which principles are more important and which must yield in case of a conflict?

In order for such a discussion to be fruitful, one must have the various requirements precisely defined and know something about their consequences. In particular, for a given set of requirements one must know whether there exist methods which satisfy the requirements, and if such methods exist, one would want a description of them. These are questions which lend themselves to investigation by formal and mathematical method. Such investigation is the topic of this paper.

In a companion paper [16], I discuss the political significance of the various requirements and point out consequences of several of the results in this paper. The present paper, however, is mainly formal and mathematical. Definitions and results are explained to some extent, but a more thorough discussion is left to [16].

As mentioned earlier, the kind of methods discussed here are not only of interest in connection with proportional elections. Much of the same analysis is relevant for the problem of how to divide the seats in an elected assembly among geographical districts. Indeed, a significant part of the literature in this area has been concerned with the question of how to apportion the seats in the House of Representatives of the USA among the states.

In general, issues of the type discussed in this paper arise whenever indivisible entities shall be divided "in proportion" to something. The use of words and phrases below will correspond to the case of proportional elections but can easily be "translated" to fit other applications. The relevance and relative importance of the various requirements and criteria will, of course, depend on the circumstances. (This is true even within the framework of proportional representation.)

There is a considerable descriptive literature on methods for proportional representation and their properties. Relatively few authors have adopted the approach of this paper, namely to formulate principles and criteria and study their consequences. But there is some literature in the area; see references below.[5]

There is no commonly accepted terminology in this field. Several of the concepts and methods discussed here do not have established names, or they have more than one name. When appropriate, I have adopted names which are in use elsewhere, but sometimes this is not possible or desirable. I try to avoid naming concepts and methods after persons.

As mentioned above, this paper is mainly formal and mathematical; results are stated and proved. Mathematical symbolism is used throughout, but the proofs are quite elementary and require little or no prior knowledge of mathematics. The proofs are written out in fairly great detail, and several of them are therefore relatively long. In a couple of places, standard results are introduced and used without being proved, but these cases are clearly pointed out and should not cause any problems.[6] For the sake of completeness, the discussion in this paper is sometimes carried further than the corresponding discussion in [16], and, it may be argued, further than is likely to be of practical interest. Whenever an example is needed, no attempt is made to make it realistic in the proportional representation framework; rather, formal simplicity is sought.

3. BASIC FORMULATION

Suppose that the total number of seats and the vote obtained by each party are given. The distribution of seats shall be determined. An allotment method is a procedure for doing this.[7]

Formally, for any integers $n \geq 0$ and $k \geq 2$, define

$$T_{k,n} = \{(r_1, \ldots, r_k) | r_1, \ldots, r_k \text{ are non-negative integers and } \sum_{i=1}^{k} r_i = n\}.$$

Hence, $T_{k,n}$ is the set of possible allocations when there are k parties and a total of n seats.

A vector of the form $(n; x_1, \ldots, x_k)$, where $k \geq 2$, n is a non-negative integer, and $x_1, \ldots, x_k$ are positive rational numbers, will be called a <u>situation</u>. A situation comprises all the given information; k is the number of parties, n is the total number of seats to be allotted, and $x_1, \ldots, x_k$ are the votes of the parties.

<u>Definition 1</u>

An <u>allotment method</u> is a set-valued function F, defined on all situations and satisfying $\phi \neq F(n; x_1, \ldots, x_k) \subseteq T_{k,n}$ for all situations $(n; x_1, \ldots, x_k)$.

Hence, an allotment method F is a function defined on vectors of varying length. An element of the set $F(n; x_1, \ldots, x_k)$ is itself a vector, the length of which must be consistent with that of the argument.

$(r_1, \ldots, r_k) \in F(n; x_1, \ldots, x_k)$ will be written

$$F(n; x_1, \ldots, x_k) \rightarrow (r_1, \ldots, r_k).$$

If $F(n; x_1, \ldots, x_k)$ has only one element,

$$F(n; x_1, \ldots, x_k) = (r_1, \ldots, r_k)$$

will be used as an abbreviation for

$$F(n; x_1, \ldots, x_k) = \{(r_1, \ldots, r_k)\}.$$

Individual voters' votes are not supposed to be available to the allotment method; it only knows how many votes are cast for each party. Hence the method will necessarily have a property often referred to as <u>anonymity</u>; it treats all voters equally.[8] It is possible, however, to incorporate unequal treatment in the model, by counting certain persons' votes twice or more. More generally, there can be assigned to each voter a positive rational number, representing the weight attached to that person's vote. In the specification of a situation $(n; x_1, \ldots, x_k)$, x_i shall then be the sum of the weights of the persons who voted for party i.

The letters x, y and z, with or without subscripts, superscripts, primes or the like, will be used to represent the vote of a party. Whenever such a symbol occurs, it is understood to denote a positive rational number. Similarly, r, s and t will represent the number of seats allotted to a party; these letters will always denote non-negative integers. Sometimes the symbols $\overline{x}$, $\overline{r}$ etc. will be used instead of $(x_1, \ldots, x_k)$, $(r_1, \ldots, r_k)$, etc.; the length of the vector will be clear from the context. k and n (again possibly with subscripts etc.) are the number of parties and the total number of seats, respectively; they are integers with $k \geq 2$ and $n \geq 0$.

If F and G are allotment methods and $F(n; \overline{x}) \subseteq G(n; \overline{x})$ for all situations $(n; \overline{x})$, F is said to be a <u>submethod</u> of G. This is written $F \subseteq G$. Of course, $F \subseteq F$ for any F.

F_T will denote the trivial allotment method defined by

$$F_T(n; x_1 \ldots, x_k) = T_{k,n}$$

for all situations $(n; x_1, \ldots, x_k)$.

For most of the methods which will be considered, $F(n; \overline{x})$ will have only one element except on a "thin" set of situations.[9] (F_T is an example of a method for which this is not true.) This means that whenever there is a "tie" (that is, if $F(n; \overline{x})$ has two or more elements), the tie can be broken by making arbitrarily small changes in the votes. The possibility of ties is, however, crucial; no reasonable method can completely avoid them. Consider, for example, the trivial situation when one seat shall be distributed and there are two parties with exactly equal votes. If the method treats the parties equally, both the possible allotments (1, 0) and (0, 1) must be elements of $F(1; x,x)$; hence there is a tie.[10]

An element of $F(n; \overline{x})$ shall be thought of as a <u>possible</u> allocation in the situation $(n; \overline{x})$. Of course, there must be at least one such possibility in every situation. When the method F is used in practice, there must be some procedure for breaking ties. That is, somehow one final outcome must be chosen when $F(n; \bar{x})$ has two or more elements. Presumably, this will be done by lottery; for example, each element of $F(n; \bar{x})$ can be chosen with equal probability. But such a lottery procedure is not a part of the present formulation. The elements of $F(n; \bar{x})$ are merely possible allocations; the choice among them is unspecified.[11]

If $n = 0$, everything is trivial. The case is included for technical convenience. It is assumed that the vote of a party is a strictly positive number. This assumption is made to avoid some technical problems which can arise if $x_1 = \ldots = x_k = 0$, and it should not make any difference for the relevance of the results. If it should happen that a party does not get any votes at all, it can simply be eliminated before the allotment method is used.[12]

An allotment method shall be defined for situations involving any number of parties. For some of the results below, this is essential; the proof breaks down and the statement is wrong if the number of parties is limited.[13] Other results hold even if there is a specific upper limit on the number of parties. In practical applications, it may be possible to establish such an upper limit.[14] Therefore, the difference between these two types of results can have some significance. Below, it is always pointed out whether a proof depends on there being an unlimited number of parties. When the phrase "the number of parties is limited to K" is used (with integer $K \geq 2$), it is understood that allotment methods shall only be defined on situations $(n; x_1, \ldots, x_k)$ with $k \leq K$, and similar restrictions shall be read into the definitions of all requirements, conditions etc.[15]

In the same way, one could require that there be at least a certain number of parties. But this restriction does not seem to make much difference for the results.

Usually, it is essential that n can be any non-negative integer. Many of the proofs will break down if allotment methods are only supposed to be defined for certain values of n.

The votes of the parties are allowed to be any positive <u>rational</u> numbers. This may seem unnecessarily general; in any actual election, the votes are presumably <u>integers</u>. But sometimes the votes are given as fractions of the total, or as percentages, or the like, and perhaps one wants to apply an allotment method directly to that kind of data. Therefore, there is good reason to permit non-integer rational votes.[16] There does not seem to be any reason for allowing arbitrary positive real numbers as votes.[17]

4. EXACT PROPORTIONALITY

4.1 Definitions

Let a situation $(n; x_1, \ldots, x_k)$ be given. For $i = 1, \ldots, k$, $\hat{x}_i$ is defined by

$$\hat{x}_i = n \cdot \frac{x_i}{\sum_{j=1}^{k} x_j}$$

Then $\hat{x}_i$ can be viewed as the "exact number of seats" to which party i is entitled. Since $\hat{x}_i$ is generally not an integer, the party cannot actually get $\hat{x}_i$ seats. It is clear that

$$\sum_{i=1}^{k} \hat{x}_i = n.$$

When $n > 0$, the symbol $V_{\bar{x}}$ will be used to denote the average number of votes behind each seat, that is

$$V_{\bar{x}} = \frac{\sum_{i=1}^{k} x_i}{n} .$$

Then it follows that

$$\hat{x}_i = \frac{x_i}{V_{\bar{x}}} .$$

For any real number a, $\lfloor a \rfloor$ and $\lceil a \rceil$ are defined as follows:

$\lfloor a \rfloor$ = the greatest integer less than or equal to a, and

$\lceil a \rceil$ = the smallest integer greater than or equal to a.

Hence $\lfloor a \rfloor$ is a "rounded downwards," and $\lceil a \rceil$ is a "rounded upwards." If a is an integer, $\lfloor a \rfloor = \lceil a \rceil = a$; otherwise $\lfloor a \rfloor < a < \lceil a \rceil$ and $\lfloor a \rfloor + 1 = \lceil a \rceil$.

An important class of allotment methods are those for which the allotment of seats only depends on n and $\hat{x}_1, \ldots, \hat{x}_k$. This is equivalent to saying that the allocation does not change if all the votes are scaled up or down by a constant.

Definition 2

An allotment method F is scale independent if

(4.1) $F(n; x_1, \ldots, x_k) = F(n; ax_1, \ldots, ax_k)$

for all situations $(n; x_1, \ldots, x_k)$ and all positive rational numbers a.[18]

4.2 The method of the largest remainder

Even though it is impossible in general to give a party its "exact number of seats," one can try to come close. "Closeness" can be defined in various ways, but it turns out that many of these definitions lead to the same result.

Definition 3

For any positive real number p, define the allotment method F_{ℓ^p} by

$$F_{\ell^p}(n;\ x_1,\ \ldots,\ x_k) = \{(r_1,\ \ldots,\ r_k) \in T_{k,n} \mid \sum_{i=1}^{k} |\hat{x}_i - r_i|^p \leq \sum_{i=1}^{k} |\hat{x}_i - s_i|^p$$
$$\text{for all } (s_1,\ \ldots,\ s_k) \in T_{k,n}\}.$$

Moreover, define F_{ℓ^∞} by

$$F_{\ell^\infty}(n;\ x_1,\ \ldots,\ x_k) = \{(r_1,\ \ldots,\ r_k) \in T_{k,n} \mid \max_{1\leq i\leq k} |\hat{x}_i - r_i| \leq \max_{1\leq i\leq k} |\hat{x}_i - s_i|$$
$$\text{for all } (s_1,\ \ldots,\ s_k) \in T_{k,n}\}.$$

Hence F_{ℓ^p} and F_{ℓ^∞} choose the allocations $(r_1, \ldots, r_k)$ which are closest to $(\hat{x}_1, \ldots, \hat{x}_k)$ in terms of the measures $\sum_{i=1}^{k} |\hat{x}_i - r_i|^p$ and $\max_{1\leq i\leq k} |\hat{x}_i - r_i|$, respectively.[19] Since $T_{k,n}$ is a finite set, there will always be an allocation for which the measure is minimized. Therefore, F_{ℓ^p} and F_{ℓ^∞} are well-defined allotment methods. $|\hat{x}_i - r_i|$ is, in a sense, the error in the allotment for party i. F_{ℓ^1} minimizes the sum of these errors, F_{ℓ^2} minimizes the sum of the errors squared,[20] etc. The higher p is, the more weight is given to large errors. In the limit, F_{ℓ^∞} only pays attention to the "worst" error. In spite of this apparent difference, all the methods F_{ℓ^p}, for $p \geq 1$, are equal. F_{ℓ^∞} is also essentially the same method.

Definition 4

The (complete) method of the largest remainder,[21] F_{LR}, is the allotment method defined by the following algorithm:

> Let n and $x_1, \ldots, x_k$ be given, compute $\hat{x}_1, \ldots, \hat{x}_k$ and define $m = n - \sum_{i=1}^{k} \lfloor \hat{x}_i \rfloor$. Order the parties according to the size of the numbers $\hat{x}_i - \lfloor \hat{x}_i \rfloor$, breaking ties in an arbitrary way.
>
> Give, for each i, $\lfloor \hat{x}_i \rfloor$ seats to party i. Give one of the remaining m seats to each of the m first parties in the ordering constructed above.
>
> $F_{LR}(n; x_1, \ldots, x_k)$ consists of all allotments that can be constructed as described above, for different choice of the ordering of the parties.

$F_{LR}(n; x_1, \ldots, x_k)$ can only have more than one element when there is more than one possible ordering, that is, when $\hat{x}_i - \lfloor \hat{x}_i \rfloor = \hat{x}_j - \lfloor \hat{x}_j \rfloor$ for some different parties i and j. It is easy to see that $0 \leq m \leq n - 1$, which implies that the algorithm is well defined. Also, it is straightforward to prove that whenever party i gets $\lfloor \hat{x}_i \rfloor + 1$ seats (and hence gets a seat in the second phase of the algorithm), then $\hat{x}_i > \lfloor \hat{x}_i \rfloor$.

F_{LR} is obviously scale independent.

Theorem 1[22]

(a) For all $p \geq 1$, $F_{\ell^p} = F_{LR}$.

(b) $F_{LR} \subseteq F_{\ell^\infty}$. Whenever $F_{\ell^\infty}(n; \overline{x})$ has one element, $F_{LR}(n; \overline{x}) = F_{\ell^\infty}(n; \overline{x})$.

Proof

$F_{\ell^p} \subseteq F_{LR}$ will be proved first. Suppose that this is not true. Then there exists a situation $(n; \overline{x})$ and an allocation $\overline{r}$ with $\overline{r} \in F_{\ell^p}(n; \overline{x})$ and $\overline{r} \notin F_{LR}(n; \overline{x})$. There are three ways in which we can have $\overline{r} \notin F_{LR}(n; \overline{x})$:

(A) There is a party which does not get as many seats as it should get in the first phase of the definition of F_{LR}. That is, there is an i such that $r_i \le \lfloor \hat{x}_i \rfloor - 1 \le \hat{x}_i - 1$. Then there must exist a j with $r_j > \hat{x}_j$, since $\sum_{i=1}^{k} r_i = n = \sum_{i=1}^{k} \hat{x}_i$. If one seat is tranferred from party j to party i, the criteria which defines F_{ℓ^p} is reduced; $|\hat{x}_i - r_i|^p$ is reduced by at least 1 (here the condition $p \ge 1$ is needed), and $|\hat{x}_j - r_j|^p$ is increased by at most 1 or reduced. Hence we cannot have $\overline{r} \in F_{\ell^p}(n; \overline{x})$.

(B) A party gets more seats than it can possibly get by the definition of F_{LR}; that is, there is an i with $r_i \ge \lfloor \hat{x}_i \rfloor + 2 \ge \hat{x}_i + 1$. Then there must exist a j with $r_j < \hat{x}_j$, and a contradition can be derived as in (A).

(C) For all i, $r_i = \lfloor \hat{x}_i \rfloor$ or $r_i = \lfloor \hat{x}_i \rfloor + 1$, but the parties for which the latter is true are not those prescribed by the algorithm for F_{LR}. Hence there exist i and j with $r_i = \lfloor \hat{x}_i \rfloor + 1$, $r_j = \lfloor \hat{x}_j \rfloor$, and $\hat{x}_i - \lfloor \hat{x}_i \rfloor < \hat{x}_j - \lfloor \hat{x}_j \rfloor$. It can be shown that transferring a seat from party i to party j reduces the criterion which defines F_{ℓ^p}. (This follows immediately if one writes out the terms in this criterion which correspond to parties i and j before and after the transfer, and compares them.) Again, $\overline{r} \in F_{\ell^p}(n; \overline{x})$ is contradicted.

Now let $\overline{r} \in F_{\ell^p}(n; \overline{x})$. Hence $\overline{r} \in F_{LR}(n; \overline{x})$. Suppose that $\overline{s}$ is another element of $F_{LR}(n; \overline{x})$. It follows from the definition of F_{LR} that $\hat{x}_1 - r_1, \ldots, \hat{x}_k - r_k$ are exactly the same k numbers as $\hat{x}_1 - s_1, \ldots, \hat{x}_k - s_k$, although possibly in a different order. Therefore, $\overline{r}$ and $\overline{s}$ score equally on the criterion which defines F_{ℓ^p}, and $\overline{s} \in F_{\ell^p}(n; \overline{x})$. Hence, $F_{\ell^p} = F_{LR}$.

When $F_{LR}(n; \overline{x}) \to \overline{r}$, no transfer of seats can reduce the value of the criterion by which F_{ℓ^∞} is defined. (This is easily seen from the definition of F_{LR}.) Hence $F_{LR} \subsetneq F_{\ell^\infty}$. Since $F_{LR}(n; \overline{x})$ is never empty, the second part of (b) is trivial.

□

Part (b) implies that F_{LR} and F_{ℓ^∞} are essentially equal. In particular, if the numbers $\hat{x}_1 - \lfloor \hat{x}_1 \rfloor, \ldots, \hat{x}_k - \lfloor \hat{x}_k \rfloor$ are all different, then $F_{\ell^\infty}(n; \overline{x})$ will have only one element and the last statement will apply. This is the case "almost always" if the number of votes is large compared to k and n. It is also possible that $F_{LR}(n; \overline{x})$ and $F_{\ell^\infty}(n; \overline{x})$ are equal in situations where they have more than one element. But the two methods are not always equal. Consider the situation (2; 44, 39, 39, 39, 39), with a total of 200 votes. The criterion which defines F_{ℓ^∞} is minimized at any allocation in which the two seats are given to different parties; its value is 0.61. But the allocation (0, 1, 1, 0, 0) cannot be obtained if F_{LR} is used. There is a natural modification of F_{ℓ^∞} which makes it equal to F_{LR}: Compare two allocations according to the maximum error in each; if these are equal, compare the next largest error in each situation (which may be equal to the largest for one or both situations); etc.[23] This can be called the "lexicographic version" of F_{ℓ^∞}.

F_{ℓ^p} is also a well-defined allotment method when $0 < p < 1$, but it does not seem to be a reasonable one. The corresponding criterion places relatively little weight on "large errors," that is, large values of $|\hat{x}_i - r_i|$. In particular, if $0 < p < 1$ there exists a k (depending on p) with the following property: Define $x_1 = x_2 = 2$ and $x_i = 1$ for $i = 3, \ldots, k$. Then $F_{\ell^p}(2; \overline{x}) = \{(2, 0, 0, \ldots, 0), (0, 2, 0, \ldots, 0)\}$, hence F_{ℓ^p} must give both seats to the same party. Of course, $F_{LR}(2; \overline{x}) = (1, 1, 0, \ldots, 0)$.

4.3 Quota methods

An alternative way of describing the method of the largest remainder is the following: Give out the seats one by one, each time giving the seat to the party which has the highest remaining vote. Each time a party gets a seat, $V_{\overline{x}}$ is deducted from its vote. Ties are broken arbitrarily.[24] This procedure makes it clear that $V_{\overline{x}}$ is the "cost" per seat, often called the "quota."

There are other methods which can be described in the same way, but with the quota defined differently.

Definition 5

A quota criterion is a real-valued function V, defined on all pairs (n, y) where n is an integer with n > 1 and y is a positive rational number, such that, for all such n and y,

$$\frac{y}{n+1} < V(n, y) \tag{4.2}$$

and

$$V(n, y) < \frac{y}{n-1} . \tag{4.3}$$

An allotment method F is a quota method[25] if there exists a quota criterion V such that $F(n; x_1, \ldots, x_k)$, for any situation $(n; x_1, \ldots, x_k)$ with n > 1, consists of the allocations which can be obtained as follows:

> Compute $y = \sum_{i=1}^{k} x_i$. Distribute the n seats one by one, each time giving the seat to the party with the highest remaining vote and deducting V(n, y) from that party's vote. Ties are broken arbirarily.

If n = 1, the seat is given to the largest party, or any of the largest parties, if there is a tie.

V(n, y) is only supposed to be defined for n > 1, since its value for n = 0 and 1 would have no influence on the distribution of seats.

It should be noted that the quota criterion depends only on the number of seats and the total vote of all parties.[26] If it is allowed to depend on all the votes, a wider class of methods emerges.[27]

Even if conditions (4.2) and (4.3) do not hold, an allotment method can be defined by the algorithm of Definition 5. But if (4.2) is not satisfied, it may be impossible to give the parties one seat for each time the quota divides their votes. If (4.3) does not hold, one will sometimes have to award seats on the basis of non-positive "remaining votes."

When (4.2) and (4.3) are satisfied, the quota method will give a party one seat for each time the quota divides its vote. The remaining seats, if any, are given to the parties with the largest "remainders."

The method of the largest remainder clearly is a quota method; the quota criterion is $V(n, y) = \frac{y}{n}$. One "extreme" in the class of quota methods is obtained by letting (4.2) barely be satisfied, that is, by letting V(n, y) be slightly greater than $\frac{y}{n+1}$. This method is also in practical use.[28] To the extent both these methods have their merits and are candidates for actual use, other quota methods should also be considered.

In Section 8.4, a certain class of quota methods is discussed and compared to divisor methods (see Definition 13).

4.4 Lower and upper bound

As will be seen in Section 5, the method of the largest remainder has some undesirable properties. If it is given up, one has to abandon the idea that the allocation of seats shall be as close as possible to each party's "exact representation," at least if closeness is measured by the criteria of Definition 3. But still one can require that the deviations not be too large. This requirement is most conveniently divided into two parts.

Definition 6

Let F be an allotment method. Then

(a) F satisfies the <u>lower bound condition</u> if

$$r_i \geq \lfloor \hat{x}_i \rfloor \text{ for } i = 1, \ldots, k,$$

whenever $F(n; x_1, \ldots, x_k) \rightarrow (r_1, \ldots, r_k)$.

(b) F satisfies the <u>upper bound condition</u> if

$$r_i \leq \lceil \hat{x}_i \rceil \text{ for } i = 1, \ldots, k,$$

whenever $F(n; x_1, \ldots, x_k) \rightarrow (r_1, \ldots, r_k)$.[29]

F_{LR} satisfies both conditions.

The conditions can be weakened slightly by allowing $r_i = \hat{x}_i - 1$ in the lower bound condition, and $r_i = \hat{x}_i + 1$ in the upper bound condition. Only when $\hat{x}_i$ is an integer does this represent any weakening at all, and very rarely does it make any difference for the results.[30]

Conceivably, one could strengthen the lower bound condition by requiring not only that party i get at least $\lfloor \hat{x}_i \rfloor$ seats, but also that a similar condition hold for any coalition of parties. Formally, F satisfies this "extended lower bound condition" if $F(n; x_1, \ldots, x_k) \rightarrow (r_1, \ldots, r_k)$ implies

$\sum_{i \in I} r_i \geq \lfloor \sum_{i \in I} \hat{x}_i \rfloor$ for any $I \subseteq \{1, \ldots, k\}$. When the number of parties is limited to 3, F_{LR} satisfies this condition.[31] For larger numbers of parties, however, the condition is inconsistent. Consider the situation (3; 1, 1, 1, 1, 1), and let $(r_1, \ldots, r_5)$ be a corresponding allotment. Two of the parties must be without seats; assume, for example, $r_4 = r_5 = 0$. This contradicts the condition, since $\lfloor \hat{x}_4 + \hat{x}_5 \rfloor = \lfloor 0.6 + 0.6 \rfloor = 1$.[32]

An "extended upper bound condition" can be defined in a similar way, and similar results are obtained. (The situation (2; 1, 1, 1, 1, 1) is an example in which the condition cannot be satisfied.)

4.5 Correct rounding

There is yet another conceivable approach to making the number of seats close to each party's "exact representation": one could attempt to compute r_i by "correct rounding" of $\hat{x}_i$. That is, r_i shall be $\lfloor \hat{x}_i \rfloor$ if $\hat{x}_i < \lfloor \hat{x}_i \rfloor + \frac{1}{2}$, and $\lfloor \hat{x}_i \rfloor + 1$ if $\hat{x}_i > \lfloor \hat{x}_i \rfloor + \frac{1}{2}$ (and either of these if $\hat{x}_i = \lfloor \hat{x}_i \rfloor + \frac{1}{2}$). Unfortunately, this does not always lead to the correct total number of seats being allotted. A simple example is the situation (1; 1, 1, 1), where $\hat{x}_1 = \hat{x}_2 = \hat{x}_3 = \frac{1}{3}$. The best one can hope for is to achieve correct rounding when any two parties are considered and only the votes and seats of these parties are taken into account.

<u>Definition 7</u>

An allotment method F is <u>pairwise fair</u>[33] if $F(n; \bar{x}) \to \bar{r}$ implies

$$\left| r_i - (r_i + r_j) \frac{x_i}{x_i + x_j} \right| \leq \frac{1}{2}$$

for all i and j.

$(r_i + r_j) \frac{x_i}{x_i + x_j}$ is the "exact representation" of party i, when only the votes and seats of parties i and j are considered. The condition says that r_i shall be equal to the correct rounding of this number. (The similar condition for party j follows and need not be included.) It is not immediately obvious that any pairwise fair methods exist, but the existence is proved in Section 8.3 (see Theorem 7).

5. CONSISTENCY AND MEMBERSHIP MONOTONICITY

5.1 Consistency

Pairwise fairness says something about the relative representation of two parties. The condition depends only on the votes of these two parties. This represents a special case of a general idea.

Let a situation be given, and consider two parties, i and j. What will happen to these parties' representation if the votes of other parties change, but i's vote and j's vote stay the same? Obviously, anything can happen; for example, if other parties grow, there is nothing wrong in i or j (or both) losing seats. But suppose that the changes do not lead to any change in the total representation of i and j. Then it seems reasonable to require that the division of these seats between i and j not be affected either. When the votes of i and j and their total representation are given, the other votes can be seen as irrelevant to this division and should not influence it.

A similar argument can be made even if the number of parties and the total number of seats change, provided, as before, that i's and j's vote and their total representation are not affected.

Because of the possibility of ties, it is not possible to require that i's and j's representation never change under these circumstances, but the idea is captured by the following definition.

Definition 8

An allotment method F is consistent[34] if the following implication is true for all possible values of the variables:

If

$F(n_1; x_1, \ldots, x_{k_1}) \rightarrow (r_1, \ldots, r_{k_1})$,

$F(n_2; y_1, \ldots, y_{k_2}) \rightarrow (s_1, \ldots, s_{k_2})$,

$x_{i_1} = y_{i_2}$,

$x_{j_1} = y_{j_2}$, and

$r_{i_1} + r_{j_1} = s_{i_2} + s_{j_2}$, then there exist $t_1, \ldots, t_{k_2}$ such that

$$F(n_2; y_1, \ldots, y_{k_2}) \to (t_1, \ldots, t_{k_2}),$$

$$t_{i_2} = r_{i_1},$$

$$t_{j_2} = r_{j_1}, \text{ and}$$

$$t_i = s_i \text{ for all } i \notin \{i_2, j_2\}.$$

The definition relates in the following way to the discussion above: Let a situation $(n_1; \bar{x})$ and a corresponding allotment $\bar{r}$ be given, and consider parties i_1 and j_1. Another situation $(n_2; \bar{y})$ and an allotment $\bar{s}$ are also given. The vote of the designated parties and their total representation are unchanged, but everything else can have changed, including the numbering of the parties. Because of the possibility of ties, it is not possible to require that the two parties' representation has not changed. But if it has changed, it must at least be possible to transfer seats between the two parties, so as to restore the original division of their $r_{i_1} + r_{j_1}$ seats without affecting any other parties. This transfer is represented by $\bar{t}$. (If there are no ties, $\bar{t}$ must be equal to $\bar{s}$, and the representation of the two parties was unchanged.)

When consistency is used in proofs below, k_1 or k_2 will normally be equal to 2. In fact, the concept obtained by requiring that the implication in Definition 8 hold only when $k_1 = 2$ or $k_2 = 2$, is no weaker than consistency; the latter can be proved by two applications of the former.

It is possible to generalize Definition 8 by considering more than two parties. That is, one can require that whenever the votes and total representation of a group of parties remain unchanged, then each party in the group gets the same number of seats as before. (Here ties are ignored.) Definition 8 represents the special case that the "group" has two members. This generalization will follow from consistency and membership monotonicity, see Section 5.3 below.

The method of the largest remainder is not consistent, as the following example shows: In the situation (2; 31, 9), "exact representation" is 1.55

and 0.45, and hence F_{LR} (2; 31, 9) = (2, 0). The situation (2; 31, 9, 5) gives 1.38, 0.4 and 0.22 as the "exact representation," and F_{LR} (2; 31, 9, 5) = (1, 1, 0). Hence the emergence of a third party, which is nowhere near getting a seat itself, changes the distribution of seats between the first two parties.

5.2 Membership monotonicity

What will happen to the allocation of seats when the membership of the body to be elected changes while the votes remain unchanged? It seems reasonable to require that the number of seats won by a party change in the same direction as the total membership, or at least not in the opposite direction.

Definition 9

An allotment method is membership monotone[35] if the implication (a) holds for all situations $(n; x_1, \ldots, x_k)$, and (b) holds whenver $n \geq 1$:

(a) $F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$ implies that there exist $s_1, \ldots, s_k$ with $F(n+1; x_1, \ldots, x_k) \to (s_1, \ldots, s_k)$ and $s_i \geq r_i$ for all i.

(b) $F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$ implies that there exist $s_1, \ldots, s_k$ with $F(n-1; x_1, \ldots, x_k) \to (s_1, \ldots, s_k)$ and $s_i \leq r_i$ for all i.[36]

This means that if a situation and a corresponding allocation of seats are given and the total number of seats later is increased by one, it is possible to give the extra seat to one of the parties and leave the others with unchanged representation. Similarly, if the total number of seats is decreased by one, it is possible to reduce one party's representation by one and make no other change.

A stronger form of membership monotonicity is conceivable: one can require that $F(n; \overline{x}) \to \overline{r}$ and $F(n+1, \overline{x}) \to \overline{s}$ imply $s_i \geq r_i$ for all i. This would mean that each party's representation must increase or stay the same when n increases. But this condition is not a reasonable one; for example, it implies that $F(1; x, x, x) \to (1, 0, 0)$ and $F(2; x, x, x) \to (0, 1, 1)$ cannot both hold. It also contradicts consistency.

Both conditions (a) and (b) are needed later; methods can be constructed which satisfy the one and not the other. As was the case for consistency, the somewhat complicated formulation of Definition 9 has to do with the possibility of ties. As long as ties do not occur, one can let the condition read: $F(n; \overline{x}) = \overline{r}$ and $F(n+1; \overline{x}) = \overline{s}$ imply $r_i \leq s_i$ for all i. Indeed, this is equivalent to Definition 9 in situations without ties.

Membership monotonicity follows from consistency in the presence of a fairly weak extra condition, see Theorem 3 below. But there exist methods which are consistent but not membership monotone and vice versa; examples are given in section 11.2.

The method of the largest remainder is not membership monotone: $F_{LR}(3; 13, 13, 4) = (1, 1, 1)$, since the exact representation is 1.3, 1.3 and 0.4. $F_{LR}(4; 13, 13, 4) = (2, 2, 0)$; here exact representation is 1.73, 1.73 and 0.53.

5.3 A technical lemma

The following technical lemma will be useful later.

Lemma 1

Let F be an allotment method which satisfies the condition of membership monotonicity in the case k = 2. Then (a) below holds. If F is also consistent, (b), (c) and (d) hold.

(a) If $F(n; x_1, x_2) \to (r_1, r_2)$, then there exist non-decreasing, infinite sequences $r_1^{(0)}$, $r_1^{(1)}$, ... and $r_2^{(0)}$, $r_2^{(1)}$, ..., such that $r_1^{(0)} = r_2^{(0)} = 0$, $r_1^{(n)} = r_1$, $r_2^{(n)} = r_2$, and $F(m; x_1, x_2) \to (r_1^{(m)}, r_2^{(m)})$ for all m.

(b) Let k_1 and k_2 be given, and let σ be a function defined on a subset I of $\{1, \ldots, k_1\}$ with values in $\{1, \ldots, k_2\}$, such that $i \neq j$ implies $\sigma(i) \neq \sigma(j)$. Let $x_1, \ldots x_{k_1}$ and $y_1, \ldots, y_{k_2}$

satisfy $y_{\sigma(i)} = x_i$ for all $i \in I$. Assume

$F(n_1; x_1, \ldots, x_{k_1}) \to (r_1, \ldots, r_{k_1})$, and

$F(n_2; y_1, \ldots, y_{k_2}) \to (s_1, \ldots, s_{k_2})$.

Then there exist $t_1, \ldots, t_{k_2}$ such that

(i) $F(n_2; y_1, \ldots, y_{k_2}) \to (t_1, \ldots, t_{k_2})$,

(ii) $t_j = s_j$ for all j not in the range of σ, and

(iii) either $t_{\sigma(i)} \leq r_i$ for all $i \in I$, or $t_{\sigma(i)} \geq r_i$ for all $i \in I$.

(c) F is membership monotone.

(d) If G is consistent, and $G(n; x_1, x_2) \subseteq F(n; x_1, x_2)$ for all n, x_1 and x_2, then $G \subseteq F$.

Proof

(a) If $F(n; x_1, x_2) \to (r_1, r_2)$ and $n > 0$, part (b) of the definition of membership monotonicity can be used to find $r_1^{(n-1)}$ and $r_2^{(n-1)}$ with $r_1^{(n-1)} \leq r_1$, $r_2^{(n-1)} \leq r_2$ and $F(n-1; x_1, x_2) \to (r_1^{(n-1)}, r_2^{(n-1)})$. If $n-1 > 0$, $r_1^{(n-2)}$, $r_2^{(n-2)}$ can then be found in the same way, and so on down to $r_1^{(0)}$, $r_2^{(0)}$. Similarly, (a) of the definition can be used to find $r_1^{(n+1)}$, $r_2^{(n+1)}$, $r_1^{(n+2)}$, $r_2^{(n+2)}$, etc., satisfying the requirements of the lemma.

(b) Let $F(n_1; \overline{x}) \to \overline{r}$ and $F(n_2; \overline{y}) \to \overline{s}$, and choose $\overline{t}$ such that

(i) $F(n_2, \overline{y}) \to \overline{t}$;

(ii) $t_j = s_j$ for all j not in the range of σ, and

(iii') $\sum_{i \in I} |t_{\sigma(i)} - r_i|$ is minimized, subject to (i) and (ii).

The set defined by (i) and (ii) is finite and non-empty. (It contains $\overline{s}$.) Hence $\overline{t}$ is well-defined. If $\overline{t}$ does not satisfy (iii), there must exist

i and j with $t_{\sigma(i)} > r_i$ and $t_{\sigma(j)} < r_j$. By consistency, $F(r_i+r_j;\ x_i,\ x_j) \to (r_i,\ r_j)$.[37] Let $m = t_{\sigma(i)} + t_{\sigma(j)}$, and apply part (a) to find $r_i^{(m)}$ and $r_j^{(m)}$ with $F(m;\ x_i,\ x_j) \to (r_i^{(m)},\ r_j^{(m)})$, and either $r_i^{(m)} \leq r_i$ and $r_j^{(m)} \leq r_j$ (if $m \leq r_i + r_j$), or $r_i^{(m)} \geq r_i$ and $r_j^{(m)} \geq r_j$. Construct $\bar{t}'$ from $\bar{t}$ by $t'_{\sigma(i)} = r_i^{(m)}$ and $t'_{\sigma(j)} = r_j^{(m)}$, leaving the rest of $\bar{t}$ unchanged. Since $r_i^{(m)} + r_j^{(m)} = m$, consistency implies that $\bar{t}'$ satisfies (i), and it satisfies (ii) since $\bar{t}$ does. Moreover, it is clear that $\sum_{i \in I} |t'_{\sigma(i)} - r_i| < \sum_{i \in I} |t_{\sigma(i)} - r_i|$, since the sum of the terms corresponding to i and j is reduced when going from $\bar{t}$ to $\bar{t}'$. This contradicts the choice of $\bar{t}$. Hence $\bar{t}$ must satisfy (iii), and the proof is complete.

(c) This is a corollary of part (b): Let $k_1 = k_2$ and let $\sigma_{(i)} = i$ for $i = 1, \ldots, k_1$. With $n_1 = n$ and $n_2 = n + 1$, (a) of Definition 9 follows. ($\bar{t}$ of part (b) is $\bar{s}$ of Definition 9.) With $n_1 = n$ and $n_2 = n - 1$, (b) of Definition 9 follows.

(d) Let $G(n;\ x_1,\ \ldots,\ x_k) \to (r_1,\ \ldots,\ r_k)$. Choose $s_1, \ldots, s_k$ such that

(i) $F(n;\ x_1,\ \ldots,\ x_k) \to (s_1,\ \ldots,\ s_k)$; and

(ii) $\sum_{i=1}^{k} |s_i - r_i|$ is minimized, subject to (i).

If $\bar{s} \neq \bar{r}$, there exist i and j with $s_i > r_i$ and $s_j < r_j$. Since G is consistent, $G(r_i+r_j;\ x_i,\ x_j) \to (r_i,\ r_j)$. By assumption, $F(r_i+r_j;\ x_i,\ x_j) \to (r_i,\ r_j)$. Part (b) can then be used to tranfer seats between parties i and j to achieve either $s_i \geq r_i$ and $s_j \geq r_j$, or $s_i \leq r_i$ and $s_j \leq r_j$. This reduces the sum in (ii), and contradicts the choice of $\bar{s}$. Hence $\bar{s} = \bar{r}$. This completes the proof.

□

Nowhere in the proof is there reference to situations which have a larger number of parties than the situations about which statements are made. Therefore, the lemma can be used when the number of parties is limited.

Lemma 1(b) includes the generalization of consistency to the case of more than two parties, mentioned at the end of Section 5.1. Suppose, in addition to the assumptions of part (b), that $\sum_{i\in I} r_i = \sum_{i\in I} s_{\sigma(i)}$. Then $t_{\sigma(i)} = r_i$ for all $i \in I$. This is analogous to the conclusion of Definition 8.

When part (b) is applied below, which will often be done, the set I usually consists of two elements. Typically, $I = \{1, 2\}$, $\sigma(1) = i$ and $\sigma(2) = j$. Then, if $s_i < r_1$ and $s_j > r_2$, the substitution of $\bar{t}$ for $\bar{s}$ represents a transfer of one or more seats from party j to party i. Applications of Lemma 1(b) will usually be formulated this way, as has already been done in the proof of part (d).

5.4 Equal treatment of the equal

In most cases, there is no reason why the order in which the parties are mentioned shall have any influence on the allotment of seats.

Definition 10

An allotment method is neutral[38] if, for all n, $\bar{x}$, and $\bar{r}$ and all permutations σ of $\{1, \ldots, k\}$,

$$F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$$

if and only if

$$F(n; x_{\sigma(1)}, \ldots, x_{\sigma(k)}) \to (r_{\sigma(1)}, \ldots, r_{\sigma(k)}).$$

Theorem 2

If an allotment method F is consistent and membership monotone, then F is neutral.

Proof

This is immediate from Lemma 1(b): Let $k_1 = k_2 = k$, $n_1 = n_2$ and $\bar{y} = (x_{\sigma(1)}, \ldots, x_{\sigma(k)})$, and assume that $F(n_1; \bar{x}) \to \bar{r}$. Then the $\bar{t}$ which exists by the lemma must be equal to $(r_{\sigma(1)}, \ldots, r_{\sigma(k)})$. Thus $F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$ implies $F(n; x_{\sigma(1)}, \ldots, x_{\sigma(k)}) \to (r_{\sigma(1)}, \ldots, r_{\sigma(k)})$. The converse implication is obtained by starting with

$(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$ and using the permutation $\tau = \sigma^{-1}$ (defined by $\tau(i) = j$ if and only if $\sigma(j) = i$) as the σ of the lemma.

□

Does neutrality follow from consistency alone? Perhaps one should expect this to be the case, but I have not been able to prove it.[39] Neither have I found any counterexamples. In most of the results below, it is not necessary to assume neutrality; but if necessary, the condition will be imposed without hesitation.

In the definition of neutrality, two different situations and corresponding allotments are compared. Another aspect of the ideal that the equal shall be treated equally has to do with situations in which two parties have the same number of votes. Because of the possibility of ties, one cannot require that two such parties always get the same number of seats; then F(1; 1, 1) would have to be empty. But their representation should not differ by more than one. For later reference, two versions of this condition are formulated. There is a strong version in which the representation of the two parties shall never differ by more than one, and a weak one in which it is only required that their representation can be brought within a difference of one without changing the representation of other parties.

Definition 11

An allotment method is

(a) strongly balanced[40] if $F(n; \overline{x}) \rightarrow \overline{r}$ and $x_i = x_j$ imply $|r_i - r_j| \leq 1$;

(b) weakly balanced if $F(n; \overline{x}) \rightarrow \overline{r}$ and $x_i = x_j$ imply that there exists an $\overline{s}$ with $F(n; \overline{x}) \rightarrow \overline{s}$, $|s_i - s_j| \leq 1$ and $s_{i'} = r_{i'}$ for $i' \notin \{i, j\}$.

If F is strongly balanced, F is obviously weakly balanced. ($\overline{s}$ can be chosen equal to $\overline{r}$.) The two conditions are really different; F_T is weakly but not strongly balanced. If $F \subseteq G$ and G is strongly balanced, F also has the property. The same is not true for the weak condition.

Theorem 3

Let F be a consistent and weakly balanced allotment method. Then F is membership monotone.

Proof

By Lemma 1(c), it is only necessary to prove membership monotonicity for situations with two parties. Assume that $F(n; x_1, x_2) \to (r_1, r_2)$, and choose $\bar{s} = (s_1^{(1)}, s_1^{(2)}, s_2^{(1)}, s_2^{(2)})$ such that

(i) $F(2n+1; x_1, x_1, x_2, x_2) \to \bar{s}$; and

(ii) $\sum_{i=1}^{2} \sum_{j=1}^{2} |s_i^{(j)} - r_i|$ is minimized subject to (i).

If $|s_1^{(1)} - s_1^{(2)}| \geq 2$, the fact that F is weakly balanced can be used to bring $s_1^{(1)}$ and $s_1^{(2)}$ within a distance of at most one, without changing $s_2^{(1)}$ and $s_2^{(2)}$. This change cannot increase the sum in (ii); hence $|s_1^{(1)} - s_1^{(2)}| \leq 1$ can be assumed in the first place. Similarly, one can assume $|s_2^{(1)} - s_2^{(2)}| \leq 1$. Since the total number of seats is odd, one of these differences must be equal to zero and the other equal to one. Assume $s_1^{(1)} = s_1^{(2)}$ and $s_2^{(1)} + 1 = s_2^{(2)}$. (There is no loss of generality here; all possible cases can be treated similarly.) Then $s_1^{(1)} + s_2^{(1)} = n = r_1 + r_2$. If $s_1^{(1)} < r_1$ and $s_2^{(1)} > r_2$ or vice versa, $s_1^{(1)}$ and $s_2^{(1)}$ can be replaced by r_1 and r_2, $s_1^{(2)}$ and $s_2^{(2)}$ remaining unchanged. (This is possible by consistency.) But this change decreases the sum in (ii) and contradicts the choice of $\bar{s}$. Hence $s_1^{(1)} = r_1$ and $s_2^{(1)} = r_2$, and

$$F(2n+1; x_1, x_1, x_2, x_2) \to (r_1, r_1, r_2, r_2+1).$$

Consistency then gives $F(n+1; x_1, x_2) \to (r_1, r_2+1)$, and part (a) of Definition 9 is proved.

The proof of (b) uses $F(2n-1; x_1, x_1, x_2, x_2)$; otherwise it is similar.

□

If K $\geq$ 4 and the number of parties is limited to K, the theorem is true and the proof can be applied unchanged. But situations with four parties are used. If the number of parties is limited to two, there exist consistent and balanced methods which are not membership monotone.[41]

6. VOTE MONOTONICITY

Membership monotonicity requires that a party shall not lose seats when the total number of seats increases. At least equally compelling is a condition that a party shall not lose when its own vote increases while everything else remains unchanged. Similarly, in any one situation, a larger party should not get fewer seats than a smaller one. These conditions are closely related.

Definition 12

An allotment method F is

(a) externally vote monotone if $F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$, $F(n; y_1, \ldots, y_k) \to (s_1, \ldots, s_k)$, $y_i > x_i$ and $y_j = x_j$ for all $j \neq i$, imply $s_i \geq r_i$;

(b) internally vote monotone if $F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$ and $x_i > x_j$ imply $r_i \geq r_j$.

Part (a) does not require $s_j \leq r_j$ for $j \neq i$, hence nothing is said about the representation of the parties whose votes stay the same. But when external vote monotonicity is used, $s_j \leq r_j$ for $j \neq i$, or something slightly weaker, will usually follow from other assumptions. In particular, whenever Lemma 1(b) applies, there will exist a $\bar{t}$ in $F(n; \bar{y})$ with $t_i = s_i \geq r_i$ and $t_j \leq r_j$ for $j \neq i$.

One could try to strengthen the conditions by requiring strict inequality in the conclusions ($s_i > r_i$ in (a), $r_i > r_j$ in (b)). But this is not reasonable; there cannot be anything wrong (or "unfair") in a small increase in a party's vote, or a small difference between two parties, not making any difference for the representation. Moreover, each of the two conditions emerging from such a strengthening is easily seen to be inconsistent.

Another way of making the conditions stronger could be to weaken the premise $y_i > x_i$ of (a) to $y_i \geq x_i$. (In (b), this corresponds to replacing $x_i > x_j$ by $x_i \geq x_j$.) But this is equally unreasonable. Condition (a) would then contradict neutrality, since $F(1; 1, 1) = \{(1, 0), (0, 1)\}$ would be impossible. Condition (b) would become inconsistent; $F(1; 1, 1)$ could include neither $(1, 0)$ nor $(0, 1)$.

On the other hand, (a) is not the weakest form of external vote monotonicity one can imagine. When $F(n; \overline{x}) \to \overline{r}$ and $\overline{y}$ differs from $\overline{x}$ only by $y_i > x_i$, the condition does not only require the existence of an $\overline{s}$ in $F(n; \overline{y})$ with $s_i \geq r_i$; it requires $s_i \geq r_i$ for all $\overline{s}$ in $F(n; \overline{y})$. Therefore, when the vote of a party increases, it shall not only be possible to keep constant or increase its number of seats; its representation shall necessarily be kept constant or increased. (This contrasts with the definition of membership monotonicity.) It turns out that this strong a condition is necessary in order to draw any meaningful conclusions. (F_T will, for example, satisfy the weaker version suggested above.)

External vote monotonicty will, in effect, rule out the possibility of "thick" areas on which there is a tie.[42] If there is a tie at the situation $(n; \overline{x})$, so that party i can get r_i seats while some smaller number is also possible, an arbitrarily small increase in x_i will break this tie and make any representation below r_i impossible.

Remarks similar to those in the last two paragraphs also apply to internal vote monotonicity.

The two types of vote monotonicity are related, as shown by the following theorem.

Theorem 4

Let F be a consistent allotment method.

(a) If F is externally vote monotone, then F is internally vote monotone.

(b) If F is membership monotone and internally vote monotone, then F is externally vote monotone.

Proof

(a) Assume that F is externally but not internally vote monotone. By applying consistency to a case where internal vote monotonicity fails, one can find n, x_1, x_2, r_1 and r_2, such that $x_1 > x_2$, $r_1 < r_2$ and

(6.1) $$F(n; x_1, x_2) \to (r_1, r_2).$$

Another application of consistency gives

(6.2) $$F(n;\ x_2,\ x_1) \rightarrow (r_2,\ r_1).$$

There must exist s_1 and s_2 such that

(6.3) $$F(n;\ x_1,\ x_1) \rightarrow (s_1,\ s_2).$$

External vote monotonicity applied to (6.1) and (6.3) gives $s_2 \geq r_2$. Applied to (6.2) and (6.3) this condition gives $s_1 \geq r_2$, which implies $s_1 > r_1$. Hence $s_1 + s_2 > r_1 + r_2$. But $s_1 + s_2 = n = r_1 + r_2$, and this contradiction proves (a).

(b) Assume that F is consistent and membership monotone, and not externally vote monotone. Then there exist n, $\overline{x}$, $\overline{y}$, $\overline{r}$ and $\overline{s}$ such that $y_i > x_i$, $y_j = x_j$ for $j \neq i$, $F(n;\ \overline{x}) \rightarrow \overline{r}$, $F(n;\ \overline{y}) \rightarrow \overline{s}$ and $s_i < r_i$. Then $s_j > r_j$ for some j. By changing the names of several variables and applying consistency twice, it follows that there exist x_1, x_2, x_3, r_1, r_2, s_2 and s_3 (these being the old x_i, x_j, y_i, r_i, r_j, s_j and s_i, respectively), such that

(6.4) $$x_1 < x_3$$

(6.5) $$r_1 > s_3$$

(6.6) $$r_2 < s_2$$

(6.7) $$F(r_1+r_2;\ x_1,\ x_2) \rightarrow (r_1,\ r_2)$$

(6.8) $$F(s_2+s_3;\ x_2,\ x_3) \rightarrow (s_2,\ s_3)$$

For each n, consider the possibility of finding $t_1 \leq r_1$, $t_2 \leq r_2$ and $t_3 \leq s_3$ such that

(6.9) $$F(n;\ x_1,\ x_2,\ x_3) \rightarrow (t_1,\ t_2,\ t_3).$$

For n = 0, it is obviously possible to find such numbers t_1, t_2 and t_3; for $n > r_1 + r_2 + s_3$, it is clearly not possible. Hence there must exist a greatest n for which it is possible. Let this n and corresponding values of t_1, t_2 and t_3 be given. By membership monotonicity, there exist numbers t_1', t_2', t_3' and i_0 such that

(6.10) $F(n+1; x_1, x_2, x_3) \to (t'_1, t'_2, t'_3)$, $t'_{i_0} = t_{i_0} + 1$, and
$t'_i = t_i$ for $i \neq i_0$.

There are now three cases to consider:

(A) $i_0 = 3$. Then $t'_3 \leq s_3 + 1$ and $t'_2 \leq r_2 < s_2$; hence $t'_2 + t'_3 \leq s_2 + s_3$. By (6.8) and Lemma 1(b), there exist t''_2 and t''_3 with $t''_i \leq s_i$ for $i = 2, 3$ and $F(n+1; x_1, x_2, x_3) \to (t'_1, t''_2, t''_3)$. $t''_2 \leq r_2$ would now contradict the choice of n, since $t'_1 = t_1 \leq r_1$. Hence $t''_2 > r_2$. Then t''_2 and t''_3 can be substituted for t'_2 and t'_3 in (6.10), and the proof can proceed as in (B).

(B) $t'_1 \leq r_1$, $t'_2 > r_2$, $t'_3 \leq s_3$. (This includes the case $i_0 = 2$, since $i_0 = 2$ and $t'_2 \leq r_2$ will contradict the choice of n. The case also can emerge from the adjustment made in (A), after renaming t''_2 and t''_3.) By (6.7) and Lemma 1(b), there exist t''_1 and t''_2 such that $F(n+1; x_1, x_2, x_3) \to (t''_1, t''_2, t'_3)$, with either $t''_i \leq r_i$ for $i = 1, 2$, or $t''_i \geq r_i$ for $i = 1, 2$. The first possibility contradicts the choice of n; in the second case, t''_1 and t''_2 can be substituted for t'_1 and t'_2 in (6.10), and the proof proceeds to (C).

(C) $t'_1 \geq r_1$, $t'_3 \leq s_3$. (This includes the case $i_0 = 1$, since $i_0 = 1$ and $t'_1 \leq r_1$ contradicts the choice of n. It can also emerge from (B) after renaming.) By (6.5), $t'_1 > t'_3$. (6.4) and (6.10) now imply that F is not internally vote monotone. This completes the proof. □

The assumption that F is consistent is essential; in the absence of consistency there exist methods which are internally but not externally vote monotone, and vice versa.[43] I do not know whether part (b) is true if membership monotonicity is dropped from the premise. If F is balanced, membership monotonicity will follow from consistency by Theorem 3 and need not be assumed. The conditions of being internally vote monotone and balanced are based on similar ideas; if the former is imposed as a requirement, a case can be made for also requiring the latter.

Part (a) and its proof holds if the number of parties is limited to K for any $K \geq 2$. If $K \geq 3$, the same is true of part (b), but the use of situations with three parties in the proof of (b) is essential.[44]

7. DIVISOR METHODS

An important class of allotment methods is the class of divisor methods. These methods are defined in this section, and their properties are discussed here and later, particularly in Sections 8, 11 and 12. The generality and importance of the class is established by Theorem 10; see the discussion of that theorem in Section 11.1.

Definition 13

(a) An allotment method F is a (complete) divisor method[45] if there exist positive real numbers $d_1, d_2, \ldots,$ with $d_\alpha \leq d_{\alpha+1}$ for all positive integers α, such that

$$F(n; x_1, \ldots, x_k) = \{(r_1, \ldots, r_k) \in T_{k,n} \mid \frac{x_i}{d_{r_i+1}} \leq \frac{x_j}{d_{r_j}},$$

for all $i, j \in \{1, \ldots, k\}$ with $r_j > 0\}$.

If $d_\alpha < d_{\alpha+1}$ for all α, F is a strict divisor method.

(b) An allotment method F is a partial divisor method[46] if there exists a complete divisor method G such that $F \subseteq G$.

If F is a divisor method, the allotments in $F(n; x_1, \ldots, x_k)$ can be found in this way: Compute the quotients $\frac{x_i}{d_\alpha}$ for $i = 1, \ldots, k$ and $\alpha = 1, \ldots, n$.[47] Pick out the n largest of these quotients. Each of the selected quotients will "belong to" a certain party, that is, it comes from dividing a certain party's vote by some d_α. Give each party as many seats as it has quotients among the selected ones.[48]

The computation of F can also be described as follows: Distribute the seats one by one. Assume that the allotment is $(r_1, \ldots, r_k)$ after the distribution of the first m seats. Find the i for which $\frac{x_i}{d_{r_i+1}}$ is largest. (If there is a tie, choose any of the possible i's.) Then give the next seat, seat no. m+1, to party i. (In effect, part i is at this stage competing for its seat no. r_i+1. This competition is decided by the size of the quotient $\frac{x_i}{d_{r_i+1}}$.) $F(n; \bar{x})$ shall consist of all allotments that can be obtained by using this procedure, with m taking on the values 0, ..., n-1.

It should be obvious that the descriptions in the last two paragraphs are equivalent to Definition 13(a). It follows immediately that any complete divisor method is scale independent and membership monotone. The methods are also easily seen to be consistent.

It is possible to find partial divisor methods which fail to satisfy consistency, scale independence, or membership monotonicity.[49] On the other hand, there do exist methods which are proper submethods of divisor methods and satisfy these three conditions.[50]

Any complete or partial divisor method is vote monotone, both in the external and the internal sense.

Moreover, any complete divisor method is weakly balanced. It is strongly balanced if and only if it is strict.[51]

In Definition 13, it is required that $d_\alpha > 0$ for all α. In Section 11.3, the definition is generalized in a way which effectively includes the possibility $d_\alpha = 0$ for some α. A slightly different approach is taken in Definition 14(c), where $d_1 = 0$ and $d_\alpha > 0$ for $\alpha \geq 2$.

Let a divisor method F and its sequence $d_1, d_2, \ldots$ of divisors be given. If all the divisors are multiplied by the same positive real number, the method remains the same. On the other hand, assume that the divisor methods F and F' have divisors $d_1, d_2, \ldots$ and $d'_1, d'_2, \ldots$, respectively, and assume that there is no positive real number a such that $d'_\alpha = ad_\alpha$ for all α. Then F and F' are different. This is proved as follows: There must exist an α such that $\frac{d_{\alpha+1}}{d_\alpha} \neq \frac{d'_{\alpha+1}}{d'_\alpha}$; otherwise there would exist a number a with $d'_\alpha = ad_\alpha$ for all α. Assume, without loss of generality, that $\frac{d_{\alpha+1}}{d_\alpha} > \frac{d'_{\alpha+1}}{d'_\alpha}$. Then it is possible to find positive integers x and y such that $y\frac{d_{\alpha+1}}{d_\alpha} > x > y\frac{d'_{\alpha+1}}{d'_\alpha}$, which implies $\frac{x}{d_{\alpha+1}} < \frac{y}{d_\alpha}$ and $\frac{x}{d'_{\alpha+1}} > \frac{y}{d'_\alpha}$. Hence $F(2\alpha; x, y) \rightarrow (r_1, r_2)$ implies $r_1 \leq \alpha$, while $F'(2\alpha; x, y) \rightarrow (s_1, s_2)$ implies $s_1 \geq \alpha + 1$. Therefore, the sets $F(2\alpha; x, y)$ and $F'(2\alpha, x, y)$ are disjoint and $F \neq F'$.[52] The argument also shows that neither $F \subseteq F'$ nor $F' \subseteq F$ can hold; therefore, a complete divisor method cannot be a submethod of another complete divisor method. Moreover, it follows that no method can be a submethod of two different complete divisor methods; hence, for a given partial divisor method F, G of Definition 13(b) is unique.

8. SPECIAL DIVISOR METHODS

8.1 Definitions

Some divisor methods are particularly important or interesting, because they are in actual use, and because they satisfy certain conditions or are uniquely characterized by natural sets of conditions.

Definition 14

(a) The method of the highest average,[53] F_{HA}, is the divisor method defined by the sequence of divisors $d_\alpha = \alpha$.

(b) The method of major fractions,[54] F_{MF}, is the divisor method defined by the sequence of divisors $d_\alpha = \alpha - \frac{1}{2}$.

(c) The method of the smallest divisor,[55] F_{SD}, is the allotment method obtained by setting $d_\alpha = \alpha - 1$ in the definition of divisor methods; assuming, for the purpose of this definition, that $\frac{a}{0} > b$ and $\frac{a}{0} = \frac{b}{0}$ for all positive real numbers a and b.

The method of the smallest divisor is not a divisor method in the sense of Section 7, since $d_1 = 0$. The ad hoc assumptions concerning expressions of the form $\frac{a}{0}$ imply that in a situation $(n; x_1, \ldots, x_k)$ where $n < k$, no party shall get more than one seat, but the n seats can be given to any of the parties. For $n \geq k$, each party gets one seat and the remaining n-k seats are allotted in the obvious way.[56]

8.2 Characterization of F_{HA} and F_{SD}

Consistency and the lower bound condition essentially characterize the method of the highest average, as shown by the following theorem.

Theorem 5

(a) F_{HA} is consistent and satisfies the lower bound condition.

(b) Let F be a consistent allotment method which satisfies the lower bound condition. Then $F \subseteq F_{HA}$.[57]

(c) Let $K \geq 2$ and let the number of parties be limited to K. Then there exist infinitely many different complete divisor methods which are consistent and satisfy the lower bound condition.

Proof[58]

Part (c) is proved first: Let $d_1 = 1$, and choose d_2, d_3 ... such that

$$\alpha - \frac{1}{K} \leq d_\alpha \leq \alpha \quad \text{for } \alpha = 2, 3, \ldots . \tag{8.1}$$

Clearly, the sequence d_α satisfies the requirements of Definition 13(a), and it therefore defines a divisor method F. It is also clear that an infinity of different sequences can be chosen, and since d_1 is fixed, each of these defines a different divisor method.[59]

F is consistent, since it is a divisor method. Let $k \leq K$ and $F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$, define

$$a = \max_{1 \leq i \leq k} \frac{x_i}{d_{r_i+1}}$$

and let j be a party for which this maximum is achieved. (8.1) implies

$$a \leq \frac{x_j}{r_j + 1 - \frac{1}{K}}$$

or

$$a \cdot (r_j + 1 - \frac{1}{K}) \leq x_j . \tag{8.2}$$

For all i with $r_i > 0$, Definition 13(a) and (8.1) give

$$a \leq \frac{x_i}{d_{r_i}} \leq \frac{x_i}{r_i - \frac{1}{K}}$$

which implies

$$a \cdot (r_i - \frac{1}{K}) \leq x_i . \tag{8.3}$$

If $r_i = 0$, (8.3) is trivially true. Summing (8.3) for all $i \neq j$ and adding (8.2) gives

$$a \cdot (n + 1 - \frac{k}{K}) \leq \sum_{i=1}^{k} x_i .$$

Since $k \leq K$, this implies

(8.4) $$a \cdot n \leq \sum_{i=1}^{k} x_i \ .$$

(8.1) and the definition of a give, for each i,

(8.5) $$\frac{x_i}{r_i + 1} \leq \frac{x_i}{d_{r_i+1}} \leq a,$$

and (8.4) and (8.5) imply

(8.6) $$\hat{x}_i = n \cdot \frac{x_i}{\sum_{i=1}^{k} x_i} \leq r_i + 1.$$

Consider an arbitrary but fixed party i. If the inequality (8.6) is strict, it follows that $\lfloor \hat{x}_i \rfloor \leq r_i$, which is the lower bound condition. If (8.5) is not a strict inequality, then the inequality (8.2) or (8.3) which involves i must be strict; hence (8.4) is strict. Therefore, (8.6) is a strict inequality in any case, and the proof of (c) is complete.

This argument also proves part (a), since the divisors of F_{HA} satisfy (8.1) for all K.

Part (b) remains. Let F be consistent and satisfy the lower bound condition. First it will be shown that F is (strongly) balanced. If it is not, consistency can be used to find x, r and s with $r \geq s + 2$ and $F(r+s;\ x,\ x) \rightarrow (r,\ s)$. But the exact representation of each party is $\frac{r+s}{2} \geq s + 1$, hence this contradicts the lower bound condition.[60]

It now follows from Theorem 3 that F is membership monotone. Since F_{HA} is consistent and membership monotone, Lemma 1(d) applies, and in order to prove $F \subseteq F_{HA}$, it is sufficient to prove $F(n;\ x_1,\ x_2) \subseteq F_{HA}(n;\ x_1,\ x_2)$ for all n, x_1 and x_2. Assume that this is not true. Then there must exist x_1, x_2, r_1 and r_2 with $r_1 > 0$,

(8.7) $$\frac{x_1}{r_1} < \frac{x_2}{r_2 + 1}$$

and

(8.8) $$F(r_1+r_2;\ x_1,\ x_2) \rightarrow (r_1,\ r_2)\ .$$

Let k be a (large) positive integer, and consider the situation $(kr_1+r_2;\ y_1,\ \ldots,\ y_{k+1})$, where $y_i = x_1$ for $1 \leq i \leq k$, and $y_{k+1} = x_2$. The exact representation of the last party will be

$$\hat{y}_{k+1} = (kr_1 + r_2)\,\frac{x_2}{kx_1 + x_2}\,.$$

The limit of this expression, as k approaches infinity, is $r_1 \cdot \frac{x_2}{x_1}$, which is strictly greater than r_2+1 by (8.7). For large enough k, it then follows that

$$\hat{y}_{k+1} > r_2 + 1. \tag{8.9}$$

In particular, a little algebra shows that (8.9) holds whenever

$$k > \frac{x_2}{r_1x_2 - (r_2+1)\,x_1}\,.$$

This is a well defined expression; the denominator is positive by (8.7).

Now choose $s_1, \ldots, s_{k+1}$ such that

$$F(kr_1+r_2;\, y_1, \ldots, y_{k+1}) \rightarrow (s_1, \ldots, s_{k+1}),$$

and such that s_{k+1} is minimized. If $s_{k+1} > r_2$, there must exist some i with $1 \leq i \leq k$ and $s_i < r_1$. By (8.8) and Lemma 1(b), it is possible to transfer seats from party k+1 to party i, until either $s_i \geq r_1$ or $s_{k+1} \leq r_2$. This contradicts the choice of s_{k+1}. Hence $s_{k+1} \leq r_2$, but that contradicts (8.9) and the lower bound condition. The proof is complete.

□

Part (c) makes it clear that the result in part (b) depends essentially on there being an unlimited number of parties. But the proof of (b) places a bound on how much an allotment produced by F can deviate from the definition of F_{HA} when the number of parties is limited. When (8.7) and (8.8) hold,

$$\delta = \frac{r_1x_2 - (r_2+1)x_1}{x_2}$$

can be used as a measure of this deviation. (This measure is not the only one possible, but it has the advantage that it does not change if all votes are multiplied by a positive constant.) If F is consistent and satisfies the lower bound condition, and the number of parties is limited to K for some $K \geq 4$, every example satisfying (8.7) and (8.8) will have $\delta \leq \frac{1}{K-1}$. If δ is larger, the proof of part (b) can be carried through and a contradiction derived. ($K \geq 4$ is necessary to apply Theorem 3.) The largest possible deviation goes to 0 as K goes to infinity.

Parts (a) and (b) provide a characterization of all consistent methods which satisfy the lower bound condition. The result is not quite as good a characterization as one ideally would want, since there exist submethods of F_{HA} which are not consistent. (All submethods will, of course, satisfy the lower bound condition.) On the other hand, the possibility of submethods must be included; the method described in note 50 is a consistent, proper submethod of F_{HA}. But the result is not too bad either; consistency and the lower bound condition will uniquely determine the allocation of seats in all situations except the ones in which there is a tie according to F_{HA}, and such ties are very rare.[61]

Several of the results below will provide characterizations of the same kind.

A close analogy of Theorem 5 exists for the method of the smallest divisor.

<u>Theorem 6</u>

(a) F_{SD} is consistent and satisfies the upper bound condition.

(b) Let F be a consistent allotment method satisfying the upper bound condition. Then $F \subseteq F_{SD}$.[62]

(c) Let $K \geq 2$ and let the number of parties be limited to K. Then there exist infinitely many methods which are consistent and satisfy the upper bound condition, and which are different in the strong sense that none of them is a submethod of any of the others.[63]

<u>Proof</u>

The proof is similar to the proof of Theorem 5. In (c), the infinitely many method will be "divisor methods" with $d_1 = 0$; division by 0 being taken care of as in Definition 14(c). The methods are obtained by letting $d_1 = 0$

and $d_2 = 1$, and choosing d_α such that $\alpha - 1 \leq d_\alpha \leq \alpha - 1 + \frac{1}{K}$ for $\alpha = 3, 4, \ldots$ This gives infinitely many methods, and they are different by the argument in the last paragraph of Section 7.

Let F be such a method and assume that $F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$. If $n \leq k$, each party gets at most one seat and the upper bound condition is trivial.[64] Otherwise, $r_i \geq 1$ for all i and $r_j > 1$ for at least one j. It then makes sense to define

$$a = \min_{\substack{1 \leq i \leq k \\ r_i > 1}} \frac{x_i}{d_{r_i}} .$$

Exactly as above, it can be proved that

$$\sum_{i=1}^{k} x_i \leq a \cdot n$$

and, for all i with $r_i > 1$,

$$a \leq \frac{x_i}{r_i - 1} .$$

Either the first inequality is strict, or the second is strict for all i. The upper bound condition follows.

Part (a) follows immediately.

To prove (b), let F be consistent and satisfy the upper bound condition. As in the proof of Theorem 5(b), F is (strongly) balanced and membership monotone.[65] Suppose that F is not a submethod of F_{SD}. Lemma 1(d) applies; hence there exist x_1, x_2, r_1 and r_2 with

$$\frac{x_1}{r_1 - 1} < \frac{x_2}{r_2}$$

and

$$F(r_1+r_2; x_1, x_2) \to (r_1, r_2).$$

Division by 0 is defined as before. Hence $r_1 > 1$. $r_2 = 0$ is possible, but this causes no problems.

Consider the situation $(r_1+kr_2;\ y_1, \ldots, y_{k+1})$, where $y_1 = x_1$ and $y_i = x_2$ for $2 \leq i \leq k+1$. For sufficiently large k, the exact representation of the first party is

$$\hat{y}_1 = (r_1 + kr_2)\ \frac{x_1}{x_1 + kx_2} < r_1 - 1 \ .$$

Find $s_1, \ldots, s_{k+1}$ such that $F(r_1+kr_2;\ y_1, \ldots, y_{k+1}) \rightarrow (s_1, \ldots, s_{k+1})$, with s_1 maximized. If $s_1 < r_1$, then $s_i > r_2$ for some i with $2 \leq i \leq k$, and Lemma 1(b) can be used to transfer seats to the first party, contrary to the choice of s_1. Hence $s_1 \geq r_1$, contradicting the upper bound condition. This completes the proof.

□

8.3 Characterizations of F_{MF}

A couple of characterizations of the method of major fractions and its submethods can be given. The first of these exactly characterizes the set of methods F for which $F \subseteq F_{MF}$, hence the result does not have the weakness mentioned in the comments to Theorem 5.

Theorem 7

An allotment method F is pairwise fair if and only if $F \subseteq F_{MF}$.

Proof

Assume that F is not pairwise fair. Then there exist n, $\overline{x}$ and $\overline{r}$ with $F(n;\ \overline{x}) \rightarrow \overline{r}$, such that there exist i and j for which

$$\left| r_i - (r_i + r_j)\ \frac{x_i}{x_i + x_j} \right| > \frac{1}{2} \ .$$

Assume

(8.10) $$r_i - (r_i + r_j)\ \frac{x_i}{x_i + x_j} > \frac{1}{2} \ ,$$

which implies $r_i > 0$ and

(8.11) $$r_j - (r_i + r_j)\ \frac{x_j}{x_i + x_j} < -\frac{1}{2} \ .$$

(The symmetry of (8.10) and (8.11) shows that the assumption (8.10) does not represent any loss of generality.) This gives

$$\text{(8.12)} \qquad \frac{x_i}{r_i - \frac{1}{2}} < \frac{x_i + x_j}{r_i + r_j} < \frac{x_j}{r_j + \frac{1}{2}} .$$

By the definition of F_{MF}, $\bar{r}$ is not an element of $F_{MF}(n; \bar{x})$. Therefore, F is not a submethod of F_{MF}.

Conversely, assume that F is not a submethod of F_{MF}. Then there exist n, $\bar{x}$ and $\bar{r}$ with $F(n; \bar{x}) \to \bar{r}$, such that there exist i and j with $r_i > 0$ and

$$\frac{x_i}{r_i - \frac{1}{2}} < \frac{x_j}{r_j + \frac{1}{2}} .$$

This implies (8.12), from which (8.10) and (8.11) follow. Hence F is not pairwise fair, and the proof is complete.

□

In [3], the method of major fractions is characterized in a different but related way, given here as Theorem 8.

Definition 15

An allotment method F is relatively well-rounded[66] if there do not exist any n, $\bar{x}$ and $\bar{r}$ such that $F(n; \bar{x}) \to \bar{r}$ and both $r_i > \hat{x}_i + \frac{1}{2}$ for some i and $r_j < \hat{x}_j - \frac{1}{2}$ for some j.

If $r_i > \hat{x}_i + \frac{1}{2}$, the "exact representation" of party i has been rounded upwards, although its fractional part is less than one half; it has been "over-rounded."[67] Sometimes over-rounding is unavoidable, for example, in the situation (1; 1, 1, 1). Similarly, $r_j < \hat{x}_j - \frac{1}{2}$ represents "under-rounding," which also sometimes is unavoidable. But it is not necessary to have both over-rounding and under-rounding in the same allotment; if $r_i > \hat{x}_i + \frac{1}{2}$ and $r_j < \hat{x}_j - \frac{1}{2}$, one or more seats can be transferred to party j from party i so as to correct at least one of the "wrong" roundings. This is the idea behind the concept relative well-roundedness.

The method of the largest remainder is relatively well-rounded, as is obvious from its definition. But the method of major fractions is essentially the only consistent and relatively well-rounded method.

Theorem 8

(a) F_{MF} is consistent and relatively well-rounded. If $F \subsetneq F_{MF}$, F is relatively well-rounded.

(b) Let F be a consistent and relatively well-rounded allotment method. Then $F \subseteq F_{MF}$.[68]

Proof

(a) F_{MF} is consistent. Let $F \subseteq F_{MF}$. If F is not relatively well-rounded, there exist n, $\bar{x}$, $\bar{r}$, i and j such that $F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$,

(8.13) $$r_i > \hat{x}_i + \frac{1}{2}, \text{ and}$$

(8.14) $$r_j < \hat{x}_j - \frac{1}{2} .$$

n must be positive, hence $V_{\bar{x}}$ is defined; see Section 4.1. By definition, $\hat{x}_i = \frac{x_i}{V_{\bar{x}}}$ and similarly for j. (8.13) and (8.14) then give

$$\frac{x_i}{r_i - \frac{1}{2}} < V_{\bar{x}} < \frac{x_j}{r_j + \frac{1}{2}} .$$

By the definition of F_{MF}, $\bar{r}$ cannot belong to $F_{MF}(n; \bar{x})$. This contradicts $F \subseteq F_{MF}$, and (a) is proved.[69]

(b) Let F be consistent, and assume that F is not a submethod of F_{MF}. Then there exist n, $\bar{x}$ and $\bar{r}$ as in the latter part of the proof of Theorem 7. Hence (8.10) and (8.11) follow. By consistency, $F(r_i+r_j; x_i, x_j) \to (r_i, r_j)$. But in the situation $(r_i+r_j; x_i, x_j)$, (8.10) and (8.11) say exactly what is needed to contradict relatively well-roundedness. The proof is complete.

□

8.4 Quota criteria which depend on all the votes

A quota method, in the sense of Section 4.3, shall be defined by a quota criterion which depends only on the number of seats and the total vote; see Definition 5. The methods of Definition 14 can be described in a related way, but the quota may depend on all the parties' votes. When the quota is appropriately chosen for each method, the determination of the parties which shall get a seat on the basis of the "remainder" (in addition to getting one seat for each time the quota divides the party's vote), becomes particularly simple.

The method of major fractions will be considered first. When $F(n; \overline{x}) \to \overline{r}$, Definitions 13(a) and 14(b) imply that one can find a positive real number V_{MF} satisfying

(8.15) $$\max_{1 \leq i \leq k} \frac{x_i}{r_i + \frac{1}{2}} \leq V_{MF} \leq \min_{\substack{1 \leq i \leq k \\ r_i > 0}} \frac{x_i}{r_i - \frac{1}{2}} .$$

This gives, for all i,

(8.16) $$\frac{x_i}{V_{MF}} - \frac{1}{2} \leq r_i \leq \frac{x_i}{V_{MF}} + \frac{1}{2} .$$

(When $r_i = 0$, (8.16) is trivial. Otherwise, it follows from (8.15)).

This gives the following characterization of the method: Let a "price per seat" or quota, V_{MF}, be given. Divide each party's vote by V_{MF}. The quotient is generally not an integer; round it upwards if it contains a "major fraction" (that is, if its fractional part exceeds one half), and round it downwards if its fractional part is less than one half. If V_{MF} is chosen arbitrarily, these rounded quotients need not sum to the total number of seats and can therefore not be used as an allotment of seats. But it follows from (8.16) that it is possible to find a V_{MF} for which the total comes out correctly.[70] When V_{MF} satisfies (8.15), the resulting allotment of seats will agree with F_{MF}.[71] On the other hand, it is easily seen that if V_{MF} is not chosen within the range given by (8.15), too few or too many seats will be distributed.[72] Hence, for all n and $\overline{x}$,

(8.17)
$$F_{MF}(n;\, x_1, \ldots, x_k) = \{(r_1, \ldots, r_k) \in T_{k,n} \mid$$
there exists a number $V_{MF} > 0$ such that
$$\frac{x_i}{V_{MF}} - \frac{1}{2} \leq r_i \leq \frac{x_i}{V_{MF}} + \frac{1}{2} \text{ for all } i\}.$$ [73]

Generally, V_{MF} of (8.17) will not be unique. (It is unique if and only if there is a tie, that is, if $F_{MF}(n; \overline{x})$ has two or more elements.) Although V_{MF} serves the purpose of a quota or number of votes per seat, it cannot always be chosen equal to the average number of votes per seat, $V_{\overline{x}}$. Neither can it always be chosen to satisfy (4.2) or (4.3) of Definition 5 (with $y = \sum_{i=1}^{k} x_i$).

By exactly similar reasoning, the method of the highest average can be characterized in the following way:

$$F_{HA}(n;\, x_1, \ldots, x_k) = \{(r_1, \ldots, r_k) \in T_{k,n} \mid$$
there exists a number $V_{HA} > 0$ such that
$$\frac{x_i}{V_{HA}} - 1 \leq r_i \leq \frac{x_i}{V_{HA}} \text{ for all } i\}.$$

Except in the case of a tie, this means $r_i = \left\lfloor \frac{x_i}{V_{HA}} \right\rfloor$. Hence the party gets a seat for each time its vote divides V_{HA}, fractions being disregarded. Going the other way, one can start with a divisor V_{HA} and give party i $\left\lfloor \frac{x_i}{V_{HA}} \right\rfloor$ seats. If V_{HA} is chosen as large as possible, provided that all the seats shall be distributed, one gets the method of the highest average.[74]

Comments similar to the ones made above about V_{MF} will apply to V_{HA} as well. In particular, a formula analogous to (8.15) is easily derived. V_{HA} will, however, always satisfy (4.3) of Definition 5; in fact, the stronger statement $V_{HA} \leq V_{\overline{x}}$ is true.

In the same way, when $n \geq k$,

$$F_{SD}(n;\, x_1, \ldots, x_k) = \{(r_1, \ldots, r_k) \in T_{k,n} \mid$$
there exists a number $V_{SD} > 0$ such that
$$\frac{x_i}{V_{SD}} \leq r_i \leq \frac{x_i}{V_{SD}} + 1 \text{ for all } i\}.$$

(If $V = \infty$ is allowed and $\frac{x_i}{\infty}$ is set equal to 0, the formula holds for all n.) Here $r_i = \left\lceil \frac{x_i}{V_{SD}} \right\rceil$, and a party gets one seat for each time its vote divides V_{SD}, plus one seat for the remainder. (The possibility of ties is again ignored.) Conversely, if V_{SD} is given and party i is awarded $\left\lceil \frac{x_i}{V_{SD}} \right\rceil$ seats, V_{SD} being chosen as small as possible provided that not too many seats be distributed, the allotment will agree with the method of the smallest divisor.[75]

$V_{SD} \geq V_{\bar{x}}$ in all situations. Otherwise, comments similar to the ones made about V_{MF} apply.

8.5 Quota methods and divisor methods

If there are only two parties, F_{MF} and F_{LR} are equal. Also, F_{HA} will be equal to the quota method obtained by going to the limit of (4.2) in Definition 5.[76] Similarly, F_{SD} is equal to the method obtained by going to the limit of (4.3) in Definition 5.[77]

More generally, let there be two parties and let a real number a satisfying $-1 < a < 1$ be given. Then the quota method give by

$$V(n, y) = \frac{y}{n + a} \tag{8.18}$$

is equal to the divisor method given by

$$d_\alpha = \alpha + \frac{a}{2} - \frac{1}{2}\,. \tag{8.19}$$ [78]

$a = 0$ corresponds to F_{LR} and F_{MF}.

For more than two parties, the two methods given by (8.18) and (8.19) are not equal. In particular, the quota method is not consistent; this can be proved, for any a, by an example similar to the one used for F_{LR} in Section 5.1. But the two methods will be almost equal and will coincide on a large majority of all situations.[79]

The two classes of methods defined by letting a vary from -1 to 1 in the descriptions above allow a great deal of flexibility. Various sets of objectives can be accommodated by methods from the classes, and it is a question whether there is ever a need for considering more general classes of methods. See further discussion in [16].

8.6 Minimizing inequality

A series of criteria for evaluating allotment methods can be based on the idea that the parties as far as possible shall be treated equally; the unavoidable inequalities shall be minimized. But the problem is that this idea can be formalized in many ways.

For example, let a situation $(n, \bar{x})$ and a corresponding allotment $\bar{r}$ be given. $\frac{x_i}{r_i}$ will represent the average number of votes behind each seat won by party i. In a proportional system, one would like these numbers to be as equal as possible; differences represent "inequalities." Therefore, if it is possible to reduce the difference $\left|\frac{x_i}{r_i} - \frac{x_j}{r_j}\right|$ by transferring seats between parties i and j, this should be done. It can be proved that for all situations there exists an allotment for which no bilateral transfer of seats can reduce the corresponding difference. Hence an allotment method can be defined which chooses exactly the allotment for which no such reduction is possible.[80] If one regards this criterion as the "correct" measure of fairness or proportionality, this method is therefore the ideal one.

Another possibility is to consider the numbers $\frac{r_i}{x_i}$, that is, the average number of seats earned by each vote cast for party i. An argument similar to the one above will lead to a different method.[81]

Instead of looking at the ordinary differences between $\frac{x_i}{r_i}$ and $\frac{x_j}{r_j}$, one can consider their relative difference; the relative difference between numbers a and b with $0 < a < b$ being defined by $\frac{b-a}{a}$. Thus a third method is defined.[82]

Still another possibility is to choose the allotment which minimizes

$$\sum_{i<j} \left|\frac{x_i}{r_i} - \frac{x_j}{r_j}\right| .$$

As one can imagine, a large number of criteria can be constructed by combining these possibilities in different ways. In [13], more than 60 such criteria are listed and discussed.[83] In my opinion, none of these stands out as the "correct" criterion of proportionality, and neither can I find any one of them clearly more compelling than the others. The various criteria can give some information relevant to the question of which allotment method should be chosen in a given connection, but they can in no way decide the issue.[84]

9. CONSISTENCY AND LOWER AND UPPER BOUND

No consistent allotment method can satisfy both the lower and the upper bound condition. This follows from Theorem 5(b), by the following argument: The theorem implies that such a method must be a submethod of F_{HA}. But it is easily shown by examples that no submethod of F_{HA} can satisfy the upper bound condition. For example, $F_{HA}(4; 5, 1, 1) = (4, 0, 0)$; the "exact representation" of the first party is $\frac{20}{7} < 3$, and upper bound is contradicted.

The argument in the last paragraph depends on the number of parties being unlimited, since the proof of Theorem 5(b) depends on this.[85] A much more elementary and explicit proof can be given, which also allows the number of parties to be limited.

Theorem 9

(a) No consistent allotment method satisfies the lower and the upper bound condition.

(b) Let $K \geq 4$ and let the number of parties be limited to K. Then there does not exist any consistent allotment method which satisfies the lower and the upper bound condition.

(c) If the number of parties is limited to three, F_{MF} is consistent and satisfies the lower and the upper bound condition.

Proof

Assume that F is a counterexample to (a) or (b). By the lower and the upper bound condition, F is (strongly) balanced.[86] Theorem 3, the proof of which never uses situations with more than four parties, can be applied to conclude that F is membership monotone. Consider the situation (3; 5, 1). The exact representation for the first party is 2.5, and therefore at least one of the following two cases must occur:

(A) $F(3; 5, 1) \to (2, 1)$. By membership monotonicity, $F(4; 5, 1) \to (3, 1)$ or $F(4; 5, 1) \to (2, 2)$. The latter violates the upper bound condition, hence

(9.1) $$F(4; 5, 1) \to (3, 1) .$$

Consider the situation (5; 5, 1, 1, 1). The exact representation for the first party is 3.125, and for any of the other parties it is 0.625. Apart from permutations among the three small parties, the possible allotments are (4, 1, 0, 0) and (3, 1, 1, 0). If the former is an element of F(5; 5, 1, 1, 1), (9.1) and consistency implies that the latter also is. This, in turn, implies F(5; 5, 1, 1, 1) → (2, 1, 1, 1); by consistency and the case assumption. This violates the lower bound condition.

(B) F(3; 5, 1) → (3, 0). By membership monotonicity,

(9.2) $$F(2; 5, 1) \to (2, 0) .$$

Consider F(3; 5, 1, 1, 1). The exact representation for the first party is 1.875, and for the others it is 0.375. When permutations again are ignored, the only possible allotments are (1, 1, 1, 0) and (2, 1, 0, 0). The former implies the latter, by (9.2) and consistency; while the latter implies F(3; 5, 1, 1, 1) → (3, 0, 0, 0) by consistency and the case assumption. The upper bound condition is violated. This proves (a) and (b).

To prove (c), assume that F_{MF} violates the lower bound condition in a situation with three parties. Assume, without loss of generality, that the violation occurs for party 1, that is, assume $F(n; x_1, x_2, x_3) \to (r_1, r_2, r_3)$ and $r_1 \leq \hat{x}_1 - 1$. Since $\hat{x}_1 + \hat{x}_2 + \hat{x}_3 = n = r_1 + r_2 + r_3$, it follows that $r_2 \geq \hat{x}_2 + \frac{1}{2}$ or $r_3 \geq \hat{x}_3 + \frac{1}{2}$. Therefore, there exist i and j satisfying (8.13) and (8.14) in the proof of Theorem 8(a) in Section 8.3, except that inequality (8.13) need not be strict. As was pointed out in note 69, strict inequality in either (8.13) or (8.14) is sufficient to obtain a contradiction to the definition of F_{MF}. The proof is even simpler for a situation with two parties. The upper bound condition is treated similarly. Since F_{MF} is consistent, the proof is complete.

□

10. MONOTONICITY AND LOWER AND UPPER BOUND

The method of the largest remainder has a number of desirable properties; it satisfies the lower and the upper bound condition, and one can easily prove that it is externally and internally vote monotone. But it is not consistent and not membership monotone. According to Theorem 9, it is impossible to achieve consistency and keep these desirable properties. But if consistency is given up, can membership monotonicity hold together with the other conditions? I do not know the answer to this question, but it may nevertheless be worthwhile to formulate it precisely and make some comments on it.

Problem:

Does there exist any allotment method F satisfying (i) - (v) below?

(i) F satisfies the lower bound condition.

(ii) F satisfies the upper bound condition.

(iii) F is membership monotone.

(iv) F is externally vote monotone.

(v) F is internally vote monotone.

If the number of parties is limited to three, the method of major fractions satisfies (i) - (v), by Theorem 9(c) and general properties of divisor methods. To solve the problem when the number of parties is limited to K for $K \geq 4$ seems to me to be as difficult as to solve the general problem.

In [4], a characterization of methods satisfying (i) - (iii) is given.[87] Any positive solution to the problem must be a submethod of the method $\underset{\sim}{\overline{Q}}$ of that paper.[88]

If one of the conditions (i) - (iv) is removed, the remaining conditions can be met. For (i), (ii) and (iii), this is trivial: (ii) - (v) are satisfied by a submethod of the method of the smallest divisor;[89] (i) and (iii) - (v) are satisfied by the method of the highest average; and (i), (ii), (iv) and (v) are satisfied by the method of the largest remainder.

(i) - (iii) and (v) are satisfied by the "quota method" of [1]. In the framework of this paper, this method, F_Q, can be defined recursively as follows:

$$F_Q(0; x_1, \ldots, x_k) = (0, \ldots, 0).$$

To compute $F_Q(n+1; x_1, \ldots, x_k)$, pick any element $(r_1, \ldots, r_k)$ of $F_Q(n; x_1, \ldots, x_k)$. Find those parties which can receive seat number n+1 without violating the upper bound condition. These parties are said to be eligible in the given situation. Formally, the set E of eligible parties is defined by

$$i \in E \text{ if and only if } r_i < (n+1) \frac{x_i}{\sum_{i=1}^{k} x_i} .$$

Among the eligible parties, choose one according to the criterion of the method of the highest average. That is, find j such that $j \in E$ and $\frac{x_j}{r_j + 1} \geq \frac{x_i}{r_i + 1}$ for all $i \in E$.

Then give seat number n+1 to party j, that is, let $F_Q(n+1; x_1, \ldots x_k) \rightarrow (s_1, \ldots, s_k)$ with $s_j = r_j + 1$ and $s_i = r_i$ for $i \neq j$.

$F_Q(n+1; x_1, \ldots, x_k)$ shall consist of all allotments which can be obtained in this way, by (possibly) different choices of $\bar{r}$ and j.

It is obvious that E is never empty, hence F_Q is a well-defined allotment method. By definition, it is membership monotone and satisfies the upper bound condition. It satisfies the lower bound condition because F_{HA} does: It follows from the definition of F_Q that a party in a given situation gets at least as many seats by F_Q as by F_{HA}, unless the upper bound condition causes problems and makes the party ineligible at some stage; but if the upper bound condition causes problems, the lower bound condition is certainly satisfied.[90]

Internal vote monotonicity obviously holds. But F_Q is not externally vote monotone. For example, $F_Q(16; 74, 71, 7, 6) = (8, 8, 0, 0)$, while $F_Q(16; 74, 73, 7, 6) = (8, 7, 1, 0)$; hence the second party loses a seat because it gains two votes.[91] The reason for this phenomenon can be described as follows: In both situations, the first 14 seats will be divided equally between the two large parties; there is no problem of eligibility. When the 15th seat is to be given out in the first situation, party 1 will be eligible for its 8th seat; its exact representation is slightly above 7. Party 1 then gets the 15th seats; the two small parties are far behind by the highest average criterion. When the 16th seat is to be awarded, party 2 is eligible for its 8th seat and gets it. The increase in the vote of party 2 from the first to the second situation reduces party 1's share of the vote. Its exact representation for 15 seats is reduced below 7, and only the two small parties are eligible for receiving the 15th seat. Hence one of them gets this seat, which is then irretrievably lost for the two large parties. For 16 seats there are no eligibility problems for the large parties, and party 1 wins the seat by the highest average criterion. Party 2, whose vote has increased, must bear the loss.

I know of no method which satisfies (i) - (iv).

11. A CHARACTERIZATION OF DIVISOR METHODS

11.1 The characterization

Theorem 10 below essentially says that the divisor methods are exactly the allotment methods which are consistent, membership monotone, vote monotone and scale independent.

This characterization must, however, be qualified in two ways. First, not only complete divisor methods satisfy the conditions. Second, it is necessary to treat separately the case of "divisor methods with $d_1 = 0$." Definition 13 simply rules out this possibility; hence Theorem 10 must do the same. The condition necessary to achieve this is the following: If one party gets an overwhelming majority of the total vote, it shall get all the seats. At least when allotment methods are used in proportional elections, this seems to be a reasonable condition.[92] It is removed in Section 11.3 below.

Theorem 10

(a) Let F be a divisor method. Then (i) - (v) hold.

- (i) F is consistent.
- (ii) F is membership monotone.
- (iii) F is vote monotone.[93]
- (iv) F is scale independent.
- (v) For any positive integer α, there exist x and y (depending on α), such that $F(\alpha; x, y) \rightarrow (\alpha, 0)$.[94]

(b) Let F be an allotment method which satisfies (i) - (v). Then there exists a divisor method G such that $F \subseteq G$.

Proof

Part (a) follows from earlier results and remarks. In (v), one can choose $y = 1$ and let x be any number greater than $\frac{d_\alpha}{d_1}$, where d_α and d_1 are the divisors given by Definition 13.

To prove (b), assume that F satisfies (i) - (v). For any positive integer α, consider the class of situations $(\alpha; x, y)$ in which F can give one or more seats to the second party. In order for F to be the divisor method with divisors $d_1, d_2, \ldots$, this class must include all situations $(\alpha; x, y)$

with $\frac{x}{y} < \frac{d_\alpha}{d_1}$ and none with $\frac{x}{y} > \frac{d_\alpha}{d_1}$. Hence define

$$D_\alpha = \{a \mid \text{there exist } x, y, r \text{ and } s \text{ with } a = \frac{x}{y}, \; s > 0 \text{ and } F(\alpha; x, y) \to (r, s)\}.$$

Since the votes x and y are always positive, D_α is a well-defined set of positive numbers. It is also bounded from above, as can be shown as follows: Choose x and y, by (v), to satisfy $F(\alpha; x, y) \to (\alpha, 0)$. If $\frac{x}{y}$ is not an upper bound[95] on D_α, there must exist x', y', r and s with $\frac{x'}{y'} > \frac{x}{y}$, $r < \alpha$ and $F(\alpha; x', y') \to (r, s)$. By scale independence, $F(\alpha; x'y, y'y) \to (r, s)$. Scale independence, applied to the definition of x and y, also gives $F(\alpha; xy', yy') \to (\alpha, 0)$. Since $x'y > xy'$, this contradicts (external) vote monotonicity.

It now follows that D_α has a supremum or least upper bound.[96] Define

$$d_\alpha = \sup D_\alpha .$$

By membership monotonicity, $D_\alpha \subseteq D_{\alpha+1}$; hence $d_\alpha \leq d_{\alpha+1}$. Consistency implies $F(1; 1, 1) \to (0, 1)$, which gives $1 \in D_1$. No number greater than 1 can be an element of D_1, since that would imply the existence of x and y with $x > y$ and $F(1; x, y) \to (0, 1)$, contradicting (internal) vote monotonicity. Therefore, $d_1 = 1$. Hence the sequence $d_1, d_2, \ldots,$ satisfies the conditions of Definition 13(a). Let G be the complete divisor method given by these divisors.

It must now be proved that $F \subseteq G$. Since F and G are consistent and membership monotone, Lemma 1(d) implies that this need only be proved for situations with two parties. If F is not a submethod of G, therefore, there must exist numbers x_1, x_2, r_1 and r_2 such that

$$F(r_1+r_2; x_1, x_2) \to (r_1, r_2) \tag{11.1}$$

and

$$\frac{x_1}{d_{r_1}} < \frac{x_2}{d_{r_2+1}} .$$

It is possible to find positive integers a and y such that

$$a\frac{x_1}{d_{r_1}} < y < a\frac{x_2}{d_{r_2+1}} .\text{[97]} \tag{11.2}$$

(11.1) and scale independence give

(11.3) $$F(r_1+r_2;\ ax_1,\ ax_2) \rightarrow (r_1,\ r_2).$$

Choose s_1, s_2 and s_3 such that

(11.4) $$F(r_1+r_2+1;\ ax_1,\ ax_2,\ y) \rightarrow (s_1,\ s_2,\ s_3),$$

and consider two cases:

(A) $s_3 = 0$. Then the distribution of seats between the first and the third party in (11.4), contradicts (11.2) and the conditions (i) - (v). Formally, this is shown as follows: $s_1 + s_2 > r_1 + r_2$; therefore, by (11.3) and Lemma 1(b), there is no loss in generality in assuming $s_1 \geq r_1$ and $s_2 \geq r_2$. From (11.2), $\frac{ax_1}{y} < d_{r_1}$; hence it is possible to find x', y', r_1' and r_3' with $r_3' > 0$, $F(r_1;\ x',\ y') \rightarrow (r_1',\ r_3')$, and

(11.5) $$\frac{ax_1}{y} < \frac{x'}{y'} .$$

Membership monotonicity implies the existence of t_1 and t_3 with $t_3 \geq r_3' > 0$ and $F(s_1;\ x',\ y') \rightarrow (t_1,\ t_3)$. Scale independence gives $F(s_1;\ x'y,\ yy') \rightarrow (t_1,\ t_3)$. (11.4), consistency, scale independence and the case assumption give $F(s_1;\ ax_1y',\ yy') \rightarrow (s_1,\ 0)$. Since $t_1 < s_1$ and $x'y > ax_1y'$, this contradicts (external) vote monotonicity.

(B) $s_3 > 0$. Here the distribution of seats between the second and third party in (11.4) will give a contradiction. Since $s_1 + s_2 \leq r_1 + r_2$, it is possible to assume $s_1 \leq r_1$ and $s_2 \leq r_2$, by (11.3) and Lemma 1(b). This gives $s_2 + s_3 \geq r_2 + 1$. (11.4) and Lemma 1(b) then imply the existence of t_2 and t_3 with $t_2 \leq s_2$, $t_3 \leq s_3$ and $F(r_2+1;\ ax_2,\ y) \rightarrow (t_2,\ t_3)$. Since $s_2 \leq r_2$, this gives $t_3 > 0$ and

$$\frac{ax_2}{y} \in D_{r_2+1} .$$

The definition of the numbers d_α then implies

$$\frac{ax_2}{y} \leq d_{r_2+1} .$$

This contradicts (11.2), and the proof is complete.

□

Nowhere does the proof make any use of votes which are not integers. Neither are situations with more than three parties ever used; hence the result holds if the number of parties is limited to K for $K \geq 3$. It is wrong if the number of parties is limited to two.[98] The characterization of divisor methods provided by Theorem 10 has the same shortcoming as Theorem 5; see remarks in Section 8.2. Partial divisor method must be permitted in the conclusion of part (b),[99] but not all such methods will satisfy the conditions. (Examples of partial divisor methods which violate some of the conditions are given in Section 11.2 below.)

In spite of this, the strength of the result should not be understated. It does say that any method satisfying (i) - (v) will be equal to a divisor method, except possibly on situations in which the divisor method produces a tie. For all divisor methods, ties are extremely rare.

11.2 Necessity and independence of conditions (i) - (v)

All the conditions (i) - (v) are necessary to prove Theorem 10(b), and hence they are also independent of each other. This is shown by examples below; for each of the conditions, an allotment method is constructed which satisfies the other four conditions but which is not a complete or partial divisor method. By the theorem, the method must then violate the fifth condition.

Consistency is the only condition which connects allotments for situations with a different number of parties. Therefore, the other conditions, including both internal and external vote monotonicity, will be satisfied by the method

F_i defined by

$$F_i(n;\, x_1,\, x_2) = F_{HA}(n;\, x_1,\, x_2)$$

$$F_i(n;\, x_1,\, \ldots,\, x_k) = F_{MF}(n;\, x_1,\, \ldots,\, x_k)$$

$$\text{for } k > 2.$$

This method is also neutral. The argument at the end of Section 7 shows that if two complete or partial divisor methods are equal on all situations with two parties, they have essentially the same divisors. If F_i is such a method, it must therefore be a submethod of F_{HA}. But this is clearly impossible.

A method which is not membership monotone but satisfies the other conditions cannot be balanced, not even weakly balanced. (This follows from Theorem 3.) An example of such a method is F_{ii}, defined as follows:

> For a given situation $(n;\, x_1,\, \ldots,\, x_k)$, compute an element of $F_{HA}(\lfloor\frac{n}{2}\rfloor;\, x_1,\, \ldots,\, x_k)$. Double each party's representation. If n is even, this gives an element of $F_{ii}(n;\, x_1,\, \ldots,\, x_k)$. If n is odd, there is one seat left; give it to the party with the largest vote (or to any one of these, if there are two or more parties with an equal vote).
>
> $F_{ii}(n;\, x_1,\, \ldots,\, x_k)$ shall consist of all allotments that can be obtained as described above, for all possible choices of elements in $F_{HA}(\lfloor\frac{n}{2}\rfloor;\, x_1,\, \ldots,\, x_k)$ and all possible choices of the party which gets the last seat when n is odd.

It is not difficult to see that F_{ii} satisfies neutrality, internal and external vote monotonicity, scale independence, and condition (v). Consistency is a little more complicated but can also be proved.[100] $F_{ii}(3;\, 3,\, 2) = (3,\, 0)$ and $F_{ii}(4;\, 3,\, 2) = (2,\, 2)$; hence the method is not membership monotone. To see that it is not a partial divisor method, suppose that $F_{ii} \subseteq G$ where G is a divisor method with divisors $d_1,\, d_2,\, \ldots$. Since $F_{ii}(4;\, 2,\, 1) = \{(4,\, 0),\, (2,\, 2)\}$, this would imply $\frac{2}{d_3} = \frac{2}{d_4} = \frac{1}{d_1} = \frac{1}{d_2}$. This contradicts $F_{ii}(3;\, 3,\, 2) = (3,\, 0)$.

Vote monotonicity will be violated and the rest of the conditions satisfied by the method F_{iii}, which reacts in the "wrong" way to an increase in a party's vote: $F_{iii}(n;\, x_1,\, \ldots,\, x_k) = F_{HA}(n;\, \frac{1}{x_1}\, ,\, \ldots,\, \frac{1}{x_k})$.[101]

Any partial or complete divisor method is vote monotone; therefore, F_{iii} is not such a method.

Scale independence is not satisfied by F_{iv}, defined as follows:

$$F_{iv}(n; x_1, \ldots, x_k) = F_{HA}(n; 2^{x_1}, \ldots, 2^{x_k}).$$

It is easy to show that F_{iv} satisfies conditions (i) - (iii) and (v) of Theorem 10. Suppose $F_{iv} \subseteq G$, where G is a divisor method with divisors d_1, d_2, ... $F_{iv}(3; 2, 1) = (2, 1)$, which implies $\frac{1}{d_1} \geq \frac{2}{d_3}$. On the other hand, $F_{iv}(3; 5, 3) = (3, 0)$, which gives $\frac{5}{d_3} \geq \frac{3}{d_1}$. This is a contradiction.

Let F_v be the method of the smallest divisor, modified as described in note 63, so as to be vote monotone. It is easily seen that F_v satisfies (i) - (iv), but (v) is violated, since $F_v(2; x, y) = (1, 1)$ for all x and y. All complete and partial divisor methods satisfy (v); therefore, F_v is not such a method.

This completes the demonstration that all five conditions are necessary to prove Theorem 10(b).

As mentioned in Section 11.1, it is necessary to include partial divisor methods in the conclusion of Theorem 10(b), but not all such methods satisfy the conditions of the theorem. One can then ask whether the five conditions are independent even within the class of partial divisor methods. The answer is no; it is easy to see that any partial divisor method is internally and externally vote monotone and satisfies condition (v). But (i), (ii) and (iv) are independent within this class; for each of these conditions one can find a partial divisor method which violates that condition but satisfies the others.

In note 49, an example is given of a method which satisfies conditions (ii) - (v) but which is not consistent and not neutral. An example which satisfies neutrality can be constructed by using the fact that consistency is the only condition which links allotments for situations with a different number of parties. Let F_1 be the submethod of F_{HA} obtained by breaking ties in favor of the smaller party when there are only two parties and in favor of the larger one when there are three or more parties.[102] Then $F_1(2; 4, 2) = (1, 1)$ and $F_1(2; 4, 2, 1) = (2, 0, 0)$, contradicting consistency. Neutrality and conditions (ii) - (v) are easily verified.

A method which is membership monotone but not consistent cannot be balanced, and if it is a submethod of a divisor method, this cannot be strict. Consider the method F_2, defined as follows:

Compute an element of $F_2(n; x_1, \ldots, x_k)$ by first distributing $\lfloor \frac{n}{2} \rfloor$ seats according to the method of the highest average, ties being broken in favor of the largest party.[103] Suppose that this gives an allotment $(r_1, \ldots, r_k) \in T_{k,\lfloor \frac{n}{2} \rfloor}$. If n is odd, find the party which is next in line for seat number $\lfloor \frac{n}{2} \rfloor + 1$ according to F_{HA}, but now breaking ties in favor of the smallest party. That is, find a j such that, for all i, $\frac{x_j}{r_j + 1} > \frac{x_i}{r_i + 1}$ or $(\frac{x_j}{r_j + 1} = \frac{x_i}{r_i + 1}$ and $x_j \leq x_i)$. Then let $F_2(n; x_1, \ldots, x_k) \to (s_1, \ldots, s_k)$, where $s_i = 2r_i$ for $i \neq j$, and $s_j = 2r_j$ if n is even and $s_j = 2r_i + 1$ if n is odd.

$F_2(n; x_1, \ldots, x_k)$ shall consist of all allotments that can be obtained in this way.

It is easy to see that F_2 satisfies (iii) - (v), and consistency can also be proved.[104] $F_2(3; 2, 1) = (2, 1)$ and $F_2(4; 2, 1) = (4, 0)$; hence the method is not membership monotone. F_2 is a submethod of the divisor method given by $d_\alpha = \lfloor \frac{\alpha + 1}{2} \rfloor$, that is, by the series of divisors 1, 1, 2, 2, 3, ...

Finally, let F_4 be the submethod of F_{HA} obtained by breaking ties in favor of the party i for which $|x_i - 100|$ is minimized. Conditions (i) - (iii) and (v) are satisfied, but scale independence is contradicted, since $F_4(2; 2, 1) = (2, 0)$ and $F_4(2; 200, 100) = (1, 1)$.

11.3 Generalized divisor methods

The allotment method F_V, used in Section 11.2 as an example of a method which satisfies conditions (i) - (iv) but not condition (v) of Theorem 10, is almost a divisor method. It can be defined as described in Definition 13, except that $d_1 = 0$. This is no coincidence; it will be shown below that conditions (i) - (iv) essentially characterize a class of methods obtained by allowing divisors equal to 0 in Definition 13.

Definition 16

An allotment method F is a generalized divisor method if F can be described as follows:

There is given a finite or infinite sequence $M_0, M_1, \ldots$ of integers, with $M_0 = 0$ and $M_\beta < M_{\beta+1}$ whenever M_β and $M_{\beta+1}$ exist. For each β for which M_β exists, there is given a sequence of real numbers $d_\alpha^{(\beta)}$, where $\alpha = 1, 2, \ldots M_{\beta+1} - M_\beta$ if $M_{\beta+1}$ exists and α varies over all positive integers if $M_{\beta+1}$ does not exist. These numbers satisfy $d_\alpha^{(\beta)} > 0$ and $d_\alpha^{(\beta)} \leq d_{\alpha+1}^{(\beta)}$.

To compute $F(n; x_1, \ldots, x_k)$, find the largest integer β such that $M_\beta \cdot k \leq n$. Distribute $n' = n - M_\beta \cdot k$ seats by the divisor method G_β given by the divisors $d_1^{(\beta)}, d_2^{(\beta)}, \ldots$. If $M_{\beta+1}$ exists, no party shall get more than $M_{\beta+1} - M_\beta$ seats at this stage. Suppose that $G_\beta(n'; x_1, \ldots, x_k) \rightarrow (r_1, \ldots, r_k)$. Then $F(n; x_1, \ldots, x_k) \rightarrow (M_\beta + r_1, \ldots, M_\beta + r_k)$.

$F(n; x_1, \ldots, x_k)$ shall consist of all allotments which can be obtained in this way.

The description is well defined. Since $M_0 = 0$ and the sequence M_β is strictly increasing, there will always exist a largest β for which $M_\beta \cdot k \leq n$. When β is chosen in this way, $(M_{\beta+1} - M_\beta) \cdot k > n - M_\beta \cdot k$, hence the restriction on the number of seats each party can get when G_β is used, cannot cause problems.

The choice of β is unique for any given situation $(n; x_1, \ldots, x_k)$. Therefore, $F(n; x_1, \ldots, x_k)$ will have exactly as many elements as $G_\beta(n'; x_1, \ldots, x_k)$.

The numbers M_β will represent "thresholds"; no party can get more than $M_{\beta+1}$ seats before all parties have reached M_β. Between two thresholds, a generalized divisor method works as an ordinary divisor method. There may be infinitely many thresholds, or there may be a largest threshold beyond which the generalized divisor method works exactly as an ordinary one.

For a given generalized divisor method, the sequence M_0, M_1, ... is uniquely determined. For each β, the numbers $d_1^{(\beta)}$, $d_2^{(\beta)}$, ... are unique up to multiplication by a positive constant.

Every ordinary divisor method is a generalized divisor method; M_β will not exist for $\beta > 0$ and the numbers $d_\alpha^{(0)}$ will be the ordinary divisors.

The method F_V of Section 11.2 (that is, the method F_{SD} modified as described in note 63), is also a generalized divisor method. Here $M_0 = 0$, $M_1 = 1$, M_β does not exist for $\beta > 1$, $d_1^{(0)} = 1$ and $d_\alpha^{(1)} = \alpha$ for $\alpha = 1, 2, \ldots$ In general, if one wants to guarantee each party at least m seats, a generalized divisor method with $M_1 = m$ can be used.[105]

Theorem 11

(a) Let F be a generalized divisor method. Then F satisfies (i) - (iv) of Theorem 10.

(b) Let F be an allotment method which satisfies (i) - (iv) of Theorem 10. Then there exists a generalized divisor method G such that $F \subseteq G$.

Proof

Part (a) is easily proved.

To prove (b), let F be an allotment method which satisfies (i) - (iv). Let $M_0 = 0$, and define $D_\alpha^{(0)}$, for $\alpha = 1, 2, \ldots$, exactly as D_α was defined in the proof of Theorem 10. As in that proof, it can be shown that $\sup D_1^{(0)} = 1$ and $D_\alpha^{(0)} \subseteq D_{\alpha+1}^{(0)}$. If all the sets $D_\alpha^{(0)}$ are bounded from above, one can set $d_\alpha^{(0)} = \sup D_\alpha^{(0)}$ and use the proof of Theorem 10 to prove that F is a submethod of the ordinary divisor method given by $d_1^{(0)}$, $d_2^{(0)}$, ... Otherwise, there will exist a largest number α for which $D_\alpha^{(0)}$ is bounded from above. Let M_1 be equal to this number α. Then $M_1 \geq 1 > M_0$. Define

$$d_\alpha^{(0)} = \sup D_\alpha^{(0)}, \text{ for } \alpha = 1, \ldots, M_1.$$

Assume, by induction, that M_β has been defined for the positive integer β. For $\alpha = 1, 2, \ldots$, define

$$D_\alpha^{(\beta)} = \{a \mid \text{there exist } x, y, r \text{ and } s \text{ with } a = \frac{x}{y}, \; s > M_\beta$$
$$\text{and } F(2M_\beta+\alpha;\; x, y) \to (r, s)\}.$$

By consistency and vote monotonicity, $\sup D_1^{(\beta)} = 1$; and by membership monotonicity, $D_\alpha^{(\beta)} \subseteq D_{\alpha+1}^{(\beta)}$. If all the sets $D_\alpha^{(\beta)}$ are bounded from above, the numbers $M_{\beta+1}$, $M_{\beta+2}$ etc. shall not be defined. Otherwise, find the largest α such that $D_\alpha^{(\beta)}$ is bounded from above, and define $M_{\beta+1} = M_\beta + \alpha$. Since $\alpha \geq 1$, $M_{\beta+1} > M_\beta$. In any case, define

$$d_\alpha^{(\beta)} = \sup D_\alpha^{(\beta)},$$

for all positive integers α if $M_{\beta+1}$ is not defined, and for $\alpha = 1, \ldots, M_{\beta+1}-M_\beta$ otherwise.

If $M_{\beta+1}$ is defined, one then goes on to define $D_1^{(\beta+1)}$, etc. The numbers M_β and $d_\alpha^{(\beta)}$ defined in this way will satisfy the requirements of Definition 16. Let G be the generalized divisor method determined by these numbers.

In order to prove $F \subseteq G$, situations with two parties will first be considered. Suppose that there exists such a situation which contradicts $F \subseteq G$. That is, assume that there exist numbers x_1, x_2, r_1 and r_2 such that

$$F(r_1+r_2;\; x_1, x_2) \to (r_1, r_2) \tag{11.6}$$

and

$$\text{not } G(r_1+r_2;\; x_1, x_2) \to (r_1, r_2). \tag{11.7}$$

Let these numbers be chosen so that $\min(r_1, r_2)$ is as small as possible, subject to (11.6) and (11.7). Let β be the largest number such that $M_\beta \leq \min(r_1, r_2)$, which is well defined by the properties of the sequence $M_0, M_1, \ldots$. Define

$$m_i = r_i - M_\beta, \text{ for } i = 1, 2.$$

There must exist numbers r_1^* and r_2^* with $G(r_1+r_2;\ x_1,\ x_2) \to (r_1^*,\ r_2^*)$, and there is no loss of generality in assuming $r_1 > r_1^*$. $r_1 + r_2 \geq 2M_\beta$, which implies $r_1^* \geq M_\beta$ and $r_1 > M_\beta$. If $M_{\beta+1}$ exists and $r_2 \geq M_{\beta+1}$, then $r_2^* > M_{\beta+1}$. This implies $r_1^* \geq M_{\beta+1}$ by the definition of G. Hence $r_1 > M_{\beta+1}$, which contradicts the definition of β. This contradiction shows that $r_2 < M_{\beta+1}$, provided that the latter number exists.

There are two ways in which (11.7) can hold. For one thing, r_1 can exceed the next "threshold" $M_{\beta+1}$. Formally,

(11.8) $$M_{\beta+1} \text{ exists and } m_1 > M_{\beta+1} - M_\beta.$$

If (11.8) does not hold, $M_\beta \leq r_i \leq M_{\beta+1}$ for i = 1, 2 (or $M_\beta \leq r_i$ and $M_{\beta+1}$ does not exist). Then (11.7) implies that the parties' representation above the "threshold" M_β is inconsistent with the divisor method G_β of Definition 16. Since $r_1 > r_1^*$, this implies

$$\frac{x_1}{d_{m_1}^{(\beta)}} < \frac{x_2}{d_{m_2+1}^{(\beta)}} .$$

Then positive integers a and y can be found such that

(11.9) $$a\frac{x_1}{d_{m_1}^{(\beta)}} < y < a\frac{x_2}{d_{m_2+1}^{(\beta)}} .$$

$d_{m_2+1}^{(\beta)}$ is also defined when (11.8) holds, in which case it is clearly possible to choose a and y such that the right-hand inequality in (11.9) is satisfied.

(11.6) and scale independence give $F(r_1+r_2;\ ax_1,\ ax_2) \to (r_1,\ r_2)$.

Now s_1, s_2 and s_3 can be chosen such that

(11.10) $$F(3M_\beta+m_1+m_2+1;\ ax_1,\ ax_2,\ y) \to (s_1,\ s_2,\ s_3).$$

If $s_3 < M_\beta$, then $s_i > M_\beta$ for i = 1 or 2. By consistency, there exists a situation with two parties in which F can give the allotment $(s_i,\ s_3)$. This allotment can never be produced by G, and since $s_3 < M_\beta \leq \min(r_1,\ r_2)$, the choice of r_1 and r_2 is contradicted. Two cases remain:

(A) $s_3 = M_\beta$. As in case (A) of the proof of Theorem 10, the allotment of seats to the first and the third party in (11.10) will contradict the assumptions. If (11.8) holds, $D_{m_1}^{(\beta)}$ is not bounded from above, and it is clearly possible to find numbers x', y', r_1' and r_3' with $r_3' > M_\beta$, $F(2M_\beta + m_1;\ x',\ y') \rightarrow (r_1',\ r_3')$, and

$$\frac{ax_1}{y} < \frac{x'}{y'} .$$

If (11.9) holds, such numbers can be found by the left-hand inequality of (11.9) and the definition of $d_{m_1}^{(\beta)}$. As in (A) of the proof of Theorem 10, one can assume $s_1 \geq r_1$ and $s_2 \geq r_2$. $2M_\beta + m_1 = M_\beta + r_1 \leq M_\beta + s_1$; therefore, membership monotonicity implies the existence of t_1 and t_3 with $t_3 \geq r_3' > M_\beta$ and $F(M_\beta + s_1;\ x',\ y') \rightarrow (t_1,\ t_3)$. From this and (11.10), a contradiction to external vote monotonicity can be derived.

(B) $s_3 > M_\beta$. This is similar to (B) in the proof of Theorem 10. $s_1 \leq r_1$ and $s_2 \leq r_2$ can be assumed, and $s_2 + s_3 \geq r_2 + M_\beta + 1$ follows. Hence there exist t_2 and t_3 with $t_2 \leq s_2$, $t_3 \leq s_3$ and $F(M_\beta + r_2 + 1;\ ax_2,\ y) \rightarrow (t_2,\ t_3)$. $t_2 \leq s_2 \leq r_2$ gives $t_3 > M_\beta$; therefore, $\frac{ax_2}{y} \in D_{m_2+1}^{(\beta)}$. The right-hand inequality of (11.9) is contradicted.

F and G are both consistent and membership monotone. $F \subseteq G$ now follows by Lemma 1(d), and the proof of Theorem 11 is complete.

□

12. MAXIMIZING UTILITY

Let a sequence $u_0, u_1, \ldots,$ of real numbers be given, where u_α is a measure of the utility[106] enjoyed by a voter if the party for which that person has voted is awarded α seats in an election. Furthermore, assume that total social utility can be found by adding the utility of the voters. If there are k parties which get $x_1, \ldots, x_k$ votes and $r_1, \ldots, r_k$ seats, the social utility is

$$U(\overline{x}, \overline{r}) = \sum_{i=1}^{k} x_i \cdot u_{r_i} \tag{12.1}$$

Now it is possible to define an allotment method by requiring that social utility be maximized. That is, an allotment method can be defined by

$$F(n; \overline{x}) = \{\overline{r} \in T_{k,n} | U(\overline{x}, \overline{r}) \geq U(\overline{x}, \overline{s}) \text{ for all } \overline{s} \in T_{k,n}\}. \tag{12.2}$$

Presumably, the persons who vote for a party will prefer that the party get as many seats as possible. Therefore, it is reasonable to assume

$$u_\alpha < u_{\alpha+1}, \quad \text{for } \alpha = 0, 1, \ldots . \tag{12.3}$$

Another reasonable assumption is that a seat is more precious to a party and its voters when the party has few seats than when it has many. Formally, this means

$$u_\alpha - u_{\alpha-1} > u_{\alpha+1} - u_\alpha, \quad \text{for } \alpha = 1, 2, \ldots .$$

A slightly weaker condition is obtained by requiring that the seats at least not become more valuable as a party gets more of them; that is

$$u_\alpha - u_{\alpha-1} \geq u_{\alpha+1} - u_\alpha, \quad \text{for } \alpha = 1, 2, \ldots . \tag{12.4}$$

[107]

When conditions (12.3) and (12.4) hold, the method defined by (12.2) will be a divisor method, as shown by the following theorem.

Theorem 12

Let a sequence u_0, u_1, ..., of real numbers be given, assume that (12.3) and (12.4) are satisfied, and let the allotment method F be defined by (12.2). Let G be the divisor method given by the divisors $d_\alpha = \frac{1}{u_\alpha - u_{\alpha-1}}$, for $\alpha = 1, 2, \ldots$ Then F = G.

Proof

It follows from (12.3) and (12.4) that the sequence $d_1, d_2, \ldots$, is well-defined, and that $0 < d_\alpha \leq d_{\alpha+1}$ for all α. The sequence therefore defines a divisor method by Definition 13.

When party i gets its seat number r_i, the social utility is increased by $x_i \cdot (u_{r_i} - u_{r_i - 1})$. By (12.4), this expression is, for each i, a non-decreasing function of r_i. Therefore, it is possible to find the elements of $F(n; \overline{x})$ by distributing one seat at a time, each time giving the seat to the party for which the number $x_i \cdot (u_{r_i+1} - u_{r_i})$ is largest, where r_i is the number of seats party i has previously been awarded. Since $x_i \cdot (u_{r_i+1} - u_{r_i}) = \frac{x_i}{d_{r_i+1}}$, this is equivalent to the computation of allotments for the divisor method G, as explained in Section 7.[108] The theorem follows.

□

Conversely, any divisor method can be obtained by maximizing a utility criterion; if the divisors are $d_1, d_2, \ldots$, the utility u_α of having voted for a party which gets α seats can be defined by

$$u_0 = 0,$$

$$u_\alpha = \sum_{\beta=1}^{\alpha} \frac{1}{d_\beta}, \quad \text{for } \alpha = 1, 2, \ldots .$$

13. MERGER AND DIVISION

13.1 Definitions

The question to be asked in this section is: What happens when two parties merge? It will be assumed that the merger does not cause any real change; that is, it is assumed that the merged party's vote is equal to the sum of its parts' votes, while everything else remains unchanged. In formal terms, the question is: What is the connection between $(r_1, r_2, \ldots, r_k)$ and $(s_1, s_3, \ldots, s_k)$, when $F(n; x_1, x_2, \ldots, x_k) \rightarrow (r_1, r_2, \ldots, r_k)$ and $F(n; x_1 + x_2, x_3, \ldots, x_k) \rightarrow (s_1, s_3, \ldots, s_k)$?

In general, it is not reasonable to require that the new party get exactly as many seats as its parts got. For one thing, such a requirement will immediately contradict internal vote monotonicity.[109] It is possible, however, to require that the merging parties never lose because of the merger. Instead, it can be required that they never gain. Weaker conditions are obtained by requiring that the merging parties shall never lose (or win) more than one seat. Still weaker versions are of course possible, but will not be considered here. When the possiblity of ties is taken into account, this leads to the following definition.

Definition 17[110]

An allotment method F is said to

(a) encourage merger if $F(n; x_1, x_2, \ldots, x_k) \rightarrow (r_1, r_2, \ldots r_k)$ implies the existence of $(s_1, s_3, \ldots, s_k)$ with $F(n; x_1 + x_2, x_3, \ldots, x_k) \rightarrow (s_1, s_3, \ldots, s_k)$ and $s_1 \geq r_1 + r_2$;

(b) restrict loss by merger if $F(n; x_1, x_2, \ldots, x_k) \rightarrow (r_1, r_2, \ldots, r_k)$ implies the existence of $(s_1, s_3, \ldots, s_k)$ with $F(n; x_1 + x_2, x_3, \ldots, x_k) \rightarrow (s_1, s_3, \ldots, s_k)$ and $s_1 \geq r_1 + r_2 - 1$;

(c) discourage merger if $F(n; x_1, x_2, \ldots, x_k) \rightarrow (r_1, r_2, \ldots, r_k)$ implies the existence of $(s_1, s_3, \ldots, s_k)$ with $F(n; x_1 + x_2, x_3, \ldots, x_k) \rightarrow (s_1, s_3, \ldots, s_k)$ and $s_1 \leq r_1 + r_2$;

(d) <u>restrict gain by merger</u> if $F(n; x_1, x_2, \ldots, x_k) \rightarrow (r_1, r_2, \ldots, r_k)$ implies the existence of $(s_1, s_3, \ldots, s_k)$ with $F(n; x_1 + x_2, x_3, \ldots, x_k) \rightarrow (s_1, s_3, \ldots, s_k)$ and $s_1 \leq r_1 + r_2 + 1$.[111]

In a case where no ties are involved, condition (a) simply says that the merging parties do not lose. When there are ties, the condition says that no matter how the ties are broken prior to the merger, it is possible to break the tie after the merger in such a way that the merging parties have not lost. This does not imply that a loss can never occur if the ties before and after the merger are broken independently.[112] The latter condition is stronger than (a); for example, it is not satisfied by the method of the highest average,[113] while that method satisfies (a), as will be seen below. Similar comments apply to parts (b) - (d).

The trivial method F_T satisfies all the conditions of Definition 17. This means that the conditions by themselves are not very strong. Together with other conditions they will, however, have considerable strength; see, for example, Theorem 15 below.

Because of the phenomenon just described, Definition 17 does not treat merger and division symmetrically. The definition considers the possible effects of merger, and another set of concepts is obtained by looking at division instead. For example, the following definition would correspond to (a):[114]

F is said to discourage division if $F(n; x_1 + x_2, x_3, \ldots, x_k) \rightarrow (s_1, s_3, \ldots, s_k)$ implies the existence of $(r_1, r_2, \ldots, r_k)$ with $F(n; x_1, x_2, \ldots, x_k) \rightarrow (r_1, r_2, \ldots, r_k)$ and $r_1 + r_2 \leq s_1$.

The difference between these two sets of definitions is not very important; only when ties occur can there by any difference at all.[115]

13.2 <u>The method of the largest remainder</u>

The method of the largest remainder neither encourages nor discourages merger,[116] but it satisfies the weaker conditions of Definition 17.

Theorem 13

The method of the largest remainder restricts loss by merger and restricts gain by merger.[117]

Proof

Suppose that $F_{LR}(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k)$ and $F_{LR}(n; x_1 + x_2, x_3, \ldots, x_k) \to (s_1, s_3, \ldots, s_k)$. Since the total vote and the total number of seats remain unchanged, so does the exact representation of parties 3, ..., k. The exact representation of the merged party is the sum of these numbers for its parts. Moreover, it is clear that $\lfloor \hat{x}_1 \rfloor + \lfloor \hat{x}_2 \rfloor \leq \lfloor \hat{x}_1 + \hat{x}_2 \rfloor \leq \lfloor \hat{x}_1 \rfloor + \lfloor \hat{x}_2 \rfloor + 1$.

Assume that $s_1 < r_1 + r_2 - 1$. Then parties 1 and 2 must both have been awarded seats for their remainders in the first situation, while the merged party did not get such a seat in the second situation, and $\lfloor \hat{x}_1 \rfloor + \lfloor \hat{x}_2 \rfloor = \lfloor \hat{x}_1 + \hat{x}_2 \rfloor$. This implies $\hat{x}_2 - \lfloor \hat{x}_2 \rfloor > 0$, and hence $\hat{x}_1 + \hat{x}_2 - \lfloor \hat{x}_1 + \hat{x}_2 \rfloor > \hat{x}_1 - \lfloor \hat{x}_1 \rfloor$. Since the same number of seats are distributed on the basis of the remainders in the two situations and the remainders for parties 3, ..., k are unchanged, this is a contradiction.

Assume that $s_1 > r_1 + r_2 + 1$. It follows that $\lfloor \hat{x}_1 + \hat{x}_2 \rfloor = \lfloor \hat{x}_1 \rfloor + \lfloor \hat{x}_2 \rfloor + 1$. Moreover, the merged party must have been awarded a seat for its remainder, while neither party 1 nor party 2 got such a seat in the first situation. Since $\hat{x}_2 - \lfloor \hat{x}_2 \rfloor < 1$, it also follows that $\hat{x}_1 + \hat{x}_2 - \lfloor \hat{x}_1 + \hat{x}_2 \rfloor < \hat{x}_1 - \lfloor \hat{x}_1 \rfloor$. This is a contradiction, because fewer seats are distributed on the basis of the remainders in the second situation than in the first.

□

The theorem is true for any quota method (see Definition 5 in Section 4.3). The proof can be applied essentially unchanged; $V = V(n, \sum_{i=1}^{k} x_i)$ will be the same in the two situations under consideration, and $\frac{x_i}{V}$ is substituted for $\hat{x}_i$ in the proof.[118]

13.3 Divisor methods

For complete divisor methods, there exists a straightforward connection between the divisors defining the method and its satisfying the various parts of Definition 17. The characterization also permits "divisor methods" with $d_1 = 0$, as described in Definition 14(c).

Theorem 14[119]

Let F be a complete divisor method, or a "divisor method with $d_1 = 0$" as described in Definition 14(c). Let $d_1, d_2, \ldots,$ be the divisors which define F. Then:

(a) F encourages merger if and only if $d_{\alpha+\beta} \leq d_\alpha + d_\beta$ for all α and β.[120]

(b) F restricts loss by merger if and only if $d_{\alpha+\beta-1} \leq d_\alpha + d_\beta$ for all α and β.

(c) F discourages merger if and only if $d_{\alpha+\beta-1} \geq d_\alpha + d_\beta$ for all α and β.

(d) F restricts gain by merger if and only if $d_{\alpha+\beta} \geq d_\alpha + d_\beta$ for all α and β.

Proof

(a) and (d) will be proved; the proofs of (b) and (c) are similar.

For (a), let $d_{\alpha+\beta} \leq d_\alpha + d_\beta$ for all α and β, let $F(n; x_1, x_2, \ldots, x_k) \rightarrow (r_1, r_2, \ldots, r_k)$, and assume that the allotment has been found by computing quotients $\frac{x_i}{d_\alpha}$ and ordering them as explained after Definition 13 in Section 7. Let a be the value of the largest quotient which belongs to any of the parties 3, ..., k and for which no seat was awarded. Then $\frac{x_1}{d_{r_1}} \geq a$ and $\frac{x_2}{d_{r_2}} \geq a$.[121] The assumption $d_{r_1+r_2} \leq d_{r_1} + d_{r_2}$ then implies $\frac{x_1 + x_2}{d_{r_1+r_2}} \geq a$. By the definition of a, parties 3, ..., k can at most have $n - (r_1 + r_2)$ quotients greater than a. In the situation $(n; x_1 + x_2, x_3, \ldots, x_k)$, it is therefore possible to give the first party at least $r_1 + r_2$ seats. Hence F encourages merger.

To prove the converse implication, assume that $d_{\alpha+\beta} > d_\alpha + d_\beta$ for some α and β. Then it is possible to find positive integers x and y with

$\frac{x}{d_\alpha} > \frac{x+y}{d_{\alpha+\beta}}$ and $\frac{y}{d_\beta} > \frac{x+y}{d_{\alpha+\beta}}$. Assume that $\frac{x}{d_\alpha} \leq \frac{y}{d_\beta}$; the opposite case can be treated similarly. By breaking ties in favor of parties 1 and 2, one can find r_1, r_2 and r_3 such that $F(2\alpha+\beta-1;\ x,\ y,\ x) \to (r_1,\ r_2,\ r_3)$ and $r_3 < \alpha$, which gives $r_1 + r_2 \geq \alpha + \beta$. But $F(2\alpha+\beta-1;\ x + y,\ x) \to (s_1,\ s_3)$ implies $s_3 \geq \alpha$ and $s_1 < \alpha + \beta$; therefore, F does not encourage merger.

To prove (d), let $d_{\alpha+\beta} \geq d_\alpha + d_\beta$ for all α and β, and assume that $F(n;\ x_1,\ x_2,\ \ldots,\ x_k) \to (r_1,\ r_2,\ \ldots,\ r_3)$. Let b be the value of the smallest quotient for which any of the parties 3, ..., k got a seat. Then $b \geq \frac{x_1}{d_{r_1+1}}$ and $b \geq \frac{x_2}{d_{r_2+1}}$. Since $d_{r_1+r_2+2} \geq d_{r_1+1} + d_{r_2+1}$, this implies $b \geq \frac{x_1 + x_2}{d_{r_1+r_2+2}}$. Parties 3, ..., k have at least $n - (r_1 + r_2)$ quotients equal to or greater than b; in the situation $(n;\ x_1 + x_2,\ x_3,\ \ldots,\ x_k)$ one can therefore avoid giving the first party its seat number $r_1 + r_2 + 2$. This shows that F restricts gain by merger.

Conversely, if $d_{\alpha+\beta} < d_\alpha + d_\beta$ for some α and β, there exist integers x and y with $\frac{x}{d_\alpha} < \frac{x+y}{d_{\alpha+\beta}}$ and $\frac{y}{d_\beta} < \frac{x+y}{d_{\alpha+\beta}}$. There is no loss of generality in assuming $\frac{x}{d_\alpha} \geq \frac{y}{d_\beta}$. By breaking ties in favor of parties 3 and 4, one can find r_1, r_2, r_3 and r_4 such that $F(3\alpha+\beta-2;\ x,\ y,\ x,\ x) \to (r_1,\ r_2,\ r_3,\ r_4)$, $r_3 \geq \alpha$ and $r_4 \geq \alpha$, which gives $r_1 + r_2 \leq \alpha + \beta - 2$. $F(3\alpha+\beta-2;\ x + y,\ x,\ x) \to (s_1,\ s_3,\ s_4)$ implies $s_1 \geq \alpha + \beta$; hence F does not restrict gain by merger.[122]

□

Parts (a) and (d) of Definition 17 make up a natural pair of conditions; when ties are ignored, they say that a merger never leads to a loss of seats, and if there is a gain it amounts to at most one seat. Theorem 14 implies that a divisor method will satisfy these two conditions if and only if $d_{\alpha+\beta} = d_\alpha + d_\beta$ for all α and β. If $d_1 > 0$, there is no loss of generality in assuming $d_1 = 1$, and it follows that $d_\alpha = \alpha$ for all α.[123] Hence the method of the highest average is the only divisor method which encourages merger and restricts gain by merger. Similarly, parts (b) and (c) of Definition 17 will hold for a method covered by Theorem 14 if and only if $d_{\alpha+\beta-1} = d_\alpha + d_\beta$. This gives $d_1 = 0$, and if $d_2 > 0$, the method of the smallest divisor emerges.

13.4 A result from 1907

The conditions of Definition 17(a) and (d) will characterize the method of the highest average not only within the class of divisor methods, but within a considerably larger class. A result to this effect was announced as early as 1907 by the Danish mathematician A. K. Erlang. In [9], Erlang made the following claim:[124]

> In order that none of the voters, by using their votes in especially sophisticated ways (instead of voting for their party list), shall be able to alter the outcome of the election to their advantage, it is necessary and sufficient that one observes the following ground rules:
>
> I. A party must not be able to gain by a part of it not voting.
>
> II. A party must not be able to gain by a part of it voting for another list.
>
> III. A. By merger of two smaller parties into a larger one, none of the parts can lose (by division of the larger party none of the parts can gain).
>
> III. B. By merger of two parties into one, the total gain (by division of one party in two, the total loss) cannot exceed one seat.
>
> It shall now be shown that there is only one usable method, the one proposed by the Belgian d'Hondt, ...[125]

As quoted, Erlang's claim is not correct. His proof uses conditions which are not stated.[126] The idea, however, is correct. The proofs of Theorems 15 and 16 below are based on Erlang's work, and Theorem 16 corresponds closely to his result.

13.5 Characterizations of F_{HA} and F_{SD}

The characterizations of F_{HA} and F_{SD} given at the end of Section 13.3 are generalized by the following theorem.

Theorem 15[127]

(a) An allotment method F is consistent and strongly balanced, encourages merger and restricts gain by merger if and only if $F = F_{HA}$.

(b) An allotment method F is consistent and strongly balanced, discourages merger and restricts loss by merger if and only if $F = F_{SD}$.

Note that the conclusion of the theorem is equality, not only inclusion as in several other theorems. It is necessary to require that F be strongly balanced, since F_T is weakly balanced and satisfies the rest of the conditions.[128]

It follows from the proof below that the theorem is correct if the number of parties is limited to K for $K \geq 4$.

In the proof, it is necessary to make iterated use of the conditions of Definition 17. But it is not immediately obvious how this can be done. If F encourages merger and $F(n; x_1, x_2, x_3, x_4) \rightarrow (r_1, r_2, r_3, r_4)$, then there exist s_1, s_3 and s_4 with $F(n; x_1 + x_2, x_3, x_4) \rightarrow (s_1, s_3, s_4)$ and $s_1 \geq r_1 + r_2$. Moreover, there exist t_1 and t_4 with $F(n; x_1 + x_2 + x_3, x_4) \rightarrow (t_1, t_4)$ and $t_1 \geq s_1 + s_3$. But this does not necessarily give $t_1 \geq r_1 + r_2 + r_3$, since nothing is known about the relationship between r_3 and s_3. This problem could have been taken care of by a slight strengthening of Definition 17, but this is unnecessary since consistency and membership monotonicity imply that it is possible to make iterated use of the conditions, that is, apply them to merger of more than two parties. A general result to this effect can be proved from Lemma 1(b), but only a special case is needed for the proof of Theorem 15.

Lemma 2

Let F be a consistent and membership monotone allotment method, assume $F(r_1 + r_2; x_1, x_2) \rightarrow (r_1, r_2)$, and let n be a positive integer.

(a) If F encourages merger, then there exist s_1 and s_2 with $F(nr_1 + r_2; nx_1, x_2) \rightarrow (s_1, s_2)$ and $s_1 \geq nr_1$.

(b) If F restricts loss by merger and $r_1 \geq 1$, then there exist s_1 and s_2 with $F(nr_1-n+1+r_2;\ nx_1,\ x_2) \rightarrow (s_1,\ s_2)$ and $s_1 \geq nr_1 - n + 1$.

(c) If F discourages merger, then there exist s_1 and s_2 with $F(nr_1+r_2;\ nx_1,\ x_2) \rightarrow (s_1,\ s_2)$ and $s_1 \leq nr_1$.

(d) If F restricts gain by merger, then there exist s_1 and s_2 with $F(nr_1+n-1+r_2;\ nx_1,\ x_2) \rightarrow (s_1,\ s_2)$ and $s_1 \leq nr_1 + n - 1$.

Proof

Parts (a) and (d) will be proved; (b) and (c) are similar. The proof is by induction on n.

For n = 1, all the parts are trivial.

Assume that (a) holds for a given n, that is, suppose that there exist s_1 and s_2 such that

$$F(nr_1+r_2;\ nx_1,\ x_2) \rightarrow (s_1,\ s_2) \tag{13.1}$$

and $s_1 \geq nr_1$. This implies $s_2 \leq r_2$. Choose r_1', s_1' and s_2' such that $F((n+1)r_1+r_2;\ nx_1,\ x_1,\ x_2) \rightarrow (s_1',\ r_1',\ s_2')$, and such that the sum

$$|s_1-s_1'| + |r_1-r_1'| + |s_2-s_2'| \tag{13.2}$$

is minimized. Assume $s_2' > r_2$. (This will lead to a contradiction.) Then $s_2' > s_2$. Since $s_1' + r_1' + s_2' = (n+1)r_1 + r_2 = s_1 + r_1 + s_2$, this implies that either $s_1' < s_1$ or $r_1' < r_1$. If the former holds, (13.1) and Lemma 1(b) can be used to change s_1' and s_2' so as to reduce the sum in (13.2), contradicting the choice of s_1' and s_2'. If the latter holds, let $m = r_1' + s_2'$. Since $F(r_1+r_2;\ x_1,\ x_2) \rightarrow (r_1,\ r_2)$, Lemma 1(a) implies the existence of $r_1^{(m)}$ and $r_2^{(m)}$ with $F(m;\ x_1,\ x_2) \rightarrow (r_1^{(m)},\ r_2^{(m)})$ and either $r_i^{(m)} \leq r_i$ for i = 1 and 2, or $r_i^{(m)} \geq r_i$ for i = 1 and 2. Substituting $r_1^{(m)}$ and $r_2^{(m)}$ for r_1' and s_2', which is possible by consistency, reduces the sum in (13.2) and gives a contradiction. Hence $s_2' \leq r_2$. Since F encourages merger, there exist t_1 and t_2 with $F((n+1)r_1+r_2;\ (n+1)x_1,\ x_2) \rightarrow (t_1,\ t_2)$ and $t_1 \geq s_1' + r_1'$.

Since $s_1' + r_1' = (n+1)r_1 + r_2 - s_2' \geq (n+1)r_1$, this proves part (a) for $n + 1$ and completes the induction step.

Then suppose (d) holds for n. That is, there exist s_1 and s_2 with

$$F(nr_1+n-1+r_2;\ nx_1,\ x_2) \to (s_1,\ s_2)$$

and $s_1 \leq nr_1 + n - 1$. Hence $s_2 \geq r_2$. Choose r_1', s_1', s_2' and s_2'' such that

$$F((n+1)r_1+n-1+2r_2;\ nx_1,\ x_1,\ x_2,\ x_2) \to (s_1',\ r_1',\ s_2',\ s_2'').$$

These numbers shall be chosen so as to minimize a sum similar to (13.2). From the assumption $s_2' + s_2'' < 2r_2$, a contradiction can be derived. ($s_2' + s_2'' < 2r_2 \leq r_2 + s_2$ gives $s_1' + r_1' > s_1 + r_1$. Hence either $s_1' > s_1$ or $r_1' > r_1$. Also, either $s_2' < r_2 \leq s_2$ or $s_2'' < r_2 \leq s_2$. Then one can proceed as above.) Therefore, $s_2' + s_2'' \geq 2r_2$. Since F restricts gain by merger, there exist t_1, t_2 and t_3 with

$$F((n+1)r_1+n-1+2r_2;\ (n+1)x_1,\ x_2,\ x_2) \to (t_1,\ t_2,\ t_3)$$

and $t_1 \leq s_1' + r_1' + 1 = (n+1)r_1 + n - 1 + 2r_2 - (s_2' + s_2'') + 1 \leq (n+1)r_1 + n$. This also gives $t_2 + t_3 \geq 2r_2 - 1$. Hence $t_2 \geq r_2$ or $t_3 \geq r_2$; there is no loss of generality in assuming the former. By Lemma 1(b), there exist t_1' and t_2' with

$$F((n+1)r_1+n+r_2;\ (n+1)x_1,\ x_2) \to (t_1',\ t_2'),$$

and either $t_i \leq t_i'$ for i = 1 and 2 or $t_i \geq t_i'$ for i = 1 and 2. In both cases, $t_1' \leq (n+1)r_1 + n$. This proves (d) for n + 1, and the induction step is completed.

□

<u>Proof of Theorem 15</u>

F_{HA} and F_{SD} clearly satisfy the conditions of parts (a) and (b), respectively.

To prove the other half of (a), assume that the allotment method F is consistent and strongly balanced, encourages merger and restricts gain by merger. F is membership monotone by Theorem 3.

First it will be proved that F is internally vote monotone. Suppose this is not the case. By consistency, a counterexample can be reduced to a situation with two parties, and there exist x_1, x_2, r_1 and r_2 with

$$(13.3) \qquad F(r_1+r_2;\ x_1,\ x_2) \rightarrow (r_1,\ r_2),$$

$$x_1 < x_2$$

and

$$r_1 > r_2.$$

If $r_1 = r_2 + 1$, one can find an integer n such that $(n+1)x_1 < nx_2$. By Lemma 2(a) there exist s_1 and s_2 with

$$F((n+1)r_1+r_2;\ (n+1)x_1,\ x_2) \rightarrow (s_1,\ s_2)$$

and $s_1 \geq (n+1)r_1$, which gives $s_2 \leq r_2$. Lemma 2(d) and consistency can now be applied to prove the existence of t_1 and t_2 with

$$F(s_1+ns_2+n-1;\ (n+1)x_1,\ nx_2) \rightarrow (t_1,\ t_2)$$

and $t_2 \leq ns_2 + n - 1$, implying $t_1 \geq s_1$. By assumption, $(n+1)x_1 < nx_2$. Moreover, $t_2 \leq ns_2 + n - 1 \leq nr_2 + n - 1 = nr_1 - 1 \leq s_1 - r_1 - 1 \leq s_1 - 2 \leq t_1 - 2$.

Therefore, if $r_1 = r_2 + 1$ it is possible to find another counterexample to internal vote monotonocity in which the smaller party gets at least two seats more than the larger one. Hence there is no loss of generality in assuming $r_1 \geq r_2 + 2$ in the first place.

Then choose r_1', r_1'' and r_2' such that

$$F(r_1+r_2;\ x_1,\ x_2 - x_1,\ x_2) \rightarrow (r_1',\ r_1'',\ r_2').$$

Clearly $r_1' + r_2' \leq r_1 + r_2$, and by (13.3) and Lemma 1(b), $r_2' \leq r_2$ can be assumed. Hence $r_1' + r_1'' \geq r_1$. Since F encourages merger, there exist s_1 and s_2 such that

$$F(r_1+r_2;\ x_2,\ x_2) \rightarrow (s_1,\ s_2)$$

and $s_1 \geq r_1' + r_1''$. Hence $s_1 \geq r_1$, and $s_2 \leq r_2 \leq r_1 - 2$. This contradicts the assumption that F is strongly balanced, and internal vote monotonicity is proved.

$F \subseteq F_{HA}$ and $F_{HA} \subseteq F$ will be proved separately. In each case, it is enough to prove the inclusion for situations with two parties, by Lemma 1(d).

If $F \subseteq F_{HA}$ does not hold, there will therefore exist x_1, x_2, r_1 and r_2 such that

$$F(r_1+r_2;\ x_1,\ x_2) \to (r_1,\ r_2) \tag{13.4}$$

and

$$\frac{x_1}{r_1} < \frac{x_2}{r_2+1}\ . \tag{13.5}$$

Lemma 2(a), applied to (13.4) with $n = r_2 + 1$, implies the existence of s_1 and s_2 such that

$$F((r_2+1)r_1+r_2;\ (r_2+1)x_1,\ x_2) \to (s_1,\ s_2)$$

and $s_1 \geq (r_2+1)r_1$, which gives $s_2 \leq r_2$. Lemma 2(d), with $n = r_1$, implies the existence of t_1 and t_2 such that

$$F(s_1+r_1s_2+r_1-1;\ (r_2+1)x_1,\ r_1x_2) \to (t_1,\ t_2)$$

and $t_2 \leq r_1s_2 + r_1 - 1$. Hence $t_1 \geq s_1 > (r_2+1)r_1 - 1 \geq r_1s_2 + r_1 - 1 \geq t_2$. But (13.5) gives $r_1x_2 > (r_2+1)x_1$. Therefore, internal vote monotonicity is contradicted, and $F \subseteq F_{HA}$ is proved.

If $F_{HA} \subseteq F$ does not hold, a counterexample with two parties exists. Since $F \subseteq F_{HA}$, this must be a situation in which F_{HA} leads to a tie between two allocations of the seats, while F produces a unique allotment. That is, there exist x_1, x_2, r_1 and r_2 such that $r_1 \geq 1$,

$$F(r_1+r_2;\ x_1,\ x_2) = (r_1,\ r_2) \tag{13.6}$$

and

$$\frac{x_1}{r_1} = \frac{x_2}{r_2+1}\ . \tag{13.7}$$

Among the situations for which (13.6) and (13.7) hold, choose one for which $r_1 + r_2$ is minimized.

$r_1 = r_2 + 1$ is impossible; it would imply $x_1 = x_2$ and contradict consistency of F. Hence two cases must be considered:

(A) $r_1 < r_2 + 1$. Then $x_1 < x_2$ and

$$\frac{x_1}{r_1} = \frac{x_2-x_1}{r_2-r_1+1} .$$

Since $r_1 \geq 1$, the choice of r_1 and r_2 implies:

$$F(r_2;\ x_1,\ x_2-x_1) \rightarrow (r_1-1,\ r_2-r_1+1).$$

$F \subseteq F_{HA}$ and consistency then give

$$F(r_1+r_2;\ x_1,\ x_1,\ x_2-x_1) \rightarrow (r_1-1,\ r_1,\ r_2-r_1+1).$$

F is neutral by Theorem 2, and the fact that F encourages merger can be applied to the second and third party to give the existence of s_1 and s_2 such that

$$F(r_1+r_2;\ x_1,\ x_2) \rightarrow (s_1,\ s_2)$$

and $s_2 \geq r_2 + 1$. This contradicts (13.6).

(B) $r_1 > r_2 + 1$. Then $x_1 > x_2$ and

$$\frac{x_1-x_2}{r_1-r_2-1} = \frac{x_2}{r_2+1} .$$

Here the choice of r_1 and r_2 implies:

$$F(r_1-1;\ x_1-x_2,\ x_2) \rightarrow (r_1-r_2-2,\ r_2+1).$$

Since $F \subseteq F_{HA}$, consistency now gives:

$$F(r_1+2r_2;\ x_1-x_2,\ x_2,\ x_2,\ x_2) \rightarrow (r_1-r_2-2,\ r_2,\ r_2+1,\ r_2+1).$$

F restricts gain by merger, hence there exist s_1, s_2 and s_2' such that

$$F(r_1+2r_2;\ x_1,\ x_2,\ x_2) \rightarrow (s_1,\ s_2,\ s_2')$$

and $s_1 \leq r_1 - 1$. It follows that $s_2 + s_2' \geq 2r_2 + 1$, which gives $s_2 \geq r_2 + 1$ or $s_2' \geq r_2 + 1$; assume the former. Since $F \subseteq F_{HA}$, $s_1 = r_1 - 1$ and $s_2 = r_2 + 1$. Consistency now contradicts (13.6).

This completes the proof of part (a) of the theorem.[129]

For (b), assume that the allotment method F is consistent and strongly balanced, discourages merger and restricts loss by merger. F is membership monotone. F will not be internally vote monotone, but it will satisfy the following slightly weaker condition:

(13.8) There do not exist numbers x_1, x_2, r_1 and r_2 such that

$$F(r_1+r_2;\ x_1,\ x_2) \rightarrow (r_1,\ r_2),$$

$x_1 < x_2$, $r_1 > r_2$ and $r_1 \geq 2$.

This means that internal vote monotonicity perhaps can be violated by a smaller party getting one seat while a larger party gets none, but this is the only possible type of violation.

If (13.8) does not hold, one can find numbers x_1, x_2, r_1 and r_2 which contradict it. There exists a positive integer n such that $(n+2)x_1 < nx_2$. By means of Lemma 2(b) and (c), a situation can now be constructed in which a smaller party gets at least three seats more than a larger one; hence $r_1 \geq r_2 + 3$ can be assumed in the first place. Then find r_1', r_1'', r_2' and r_2'' such that

$$F(r_1+2r_2;\ x_1,\ x_2-x_1,\ x_2,\ x_2) \rightarrow (r_1',\ r_1'',\ r_2',\ r_2'').$$

Choose these numbers such that $r_2' + r_2''$ is minimized. If $r_2' + r_2'' > 2r_2$, then $r_1' < r_1$, and there is no loss of generality in assuming $r_2' > r_2$. Lemma 1(b) can be used to increase r_1' and decrease r_2', contradicting the choice of r_2' and r_2''. Hence $r_2' + r_2'' \leq 2r_2$ and $r_1' + r_1'' \geq r_1$. Since F restricts loss by merger, there exist s_1, s_2 and s_2' such that

$$F(r_1+2r_2;\ x_2,\ x_2,\ x_2) \rightarrow (s_1,\ s_2,\ s_2'),$$

and $s_1 \geq r_1' + r_2' - 1 \geq r_1 - 1$. Then $s_2 + s_2' \leq 2r_2 + 1$, and $s_2 \leq r_2$ or $s_2' \leq r_2$. Hence F is not strongly balanced. This contradiction proves (13.8).

Next it will be proved that F never gives two seats to a party unless every party has been given at least one seat. If this is not true, there exist x_1, x_2 and r with $r \geq 2$ and

$$F(r;\ x_1,\ x_2) \to (r,\ 0).$$

Lemma 2(c) then gives

$$F(r;\ x_1,\ nx_2) \to (r,\ 0)$$

for all n. But it is always possible to choose n so large that $nx_2 > x_1$; hence this contradicts (13.8).

This shows that $F(n;\ x_1,\ \ldots,\ x_k) \subseteq F_{SD}(n;\ x_1,\ \ldots,\ x_k)$ whenever $n \leq k$. It also shows that when there are more seats than parties, F gives every party at least one seat. If F is not a submethod of F_{SD}, a counterexample can be reduced to a situation with two parties, and there must exist numbers x_1, x_2, r_1 and r_2 with $r_1 \geq 2$, $r_2 \geq 1$,

$$F(r_1+r_2;\ x_1,\ x_2) \to (r_1,\ r_2)$$

and

$$\frac{x_1}{r_1-1} < \frac{x_2}{r_2}\ .$$

Lemma 2(b) and (c) can now be used to find n, t_1 and t_2 with

$$F(n;\ r_2x_1,\ (r_1-1)x_2) \to (t_1,\ t_2)$$

and $t_1 \geq r_2r_1 - r_2 + 1$ and $t_2 \leq (r_1-1)r_2$. Then $t_1 \geq 2$ and $t_1 > t_2$, and the weakened version (13.8) of internal vote monotonocity is contradicted. $F \subseteq F_{SD}$ follows.

If F and F_{SD} are not equal, there must exist a two-party situation in which they are unequal. This must be a situation in which F_{SD} produces a tie while F has a unique allotment. Specifically, there must exist x_1, x_2, r_1 and

r_2 with

(13.9) $$F(r_1+r_2;\ x_1,\ x_2) = (r_1,\ r_2)$$

and either

$$r_1 = 1 \text{ and } r_2 = 0$$

or

$$r_1 \geq 2,\ r_2 \geq 1 \text{ and } \frac{x_1}{r_1-1} = \frac{x_2}{r_2}.$$

There must exist a rational number a such that $x_1 = n_1 a$ and $x_2 = n_2 a$, for positive integers n_1 and n_2. Fix such an a. If necessary, change the numbers x_1, x_2, r_1 and r_2 such that, among the two-party situations in which the votes are integer multiples of a and for which F and F_{SD} are not equal, the sum $x_1 + x_2$ is minimized. This is well defined since it is equivalent to minimizing $n_1 + n_2$.[130]

$x_1 = x_2$ will contradict consistency. If $x_1 < x_2$,

$$F(2r_1+r_2;\ x_1,\ x_1,\ x_1,\ x_2-x_1) \to (r_1-1,\ r_1-1,\ r_1,\ r_2-r_1+2),$$

by the choice of x_1 and x_2, consistency, and the fact that $F \leq F_{SD}$. If $x_1 > x_2$,

$$F(r_1+r_2;\ x_1-x_2,\ x_2,\ x_2) \to (r_1-r_2-1,\ r_2,\ r_2+1),$$

for the same reason. In both cases, it can be concluded that there exist t_1 and t_2 such that

$$F(r_1+r_2;\ x_1,\ x_2) \to (t_1,\ t_2)$$

and $t_1 \leq r_1 - 1$, since F restricts loss by merger and discourages merger. (Lemma 1(b) is used in the first of the two cases.) This contradicts (13.9).

The proof of Theorem 15 is complete.

□

It is possible to obtain a result which is a bit closer to Erlang's claim than is Theorem 15(a), by removing the condition that F be strongly balanced, and instead require that it be internally vote monotone. (Erlang's conditions I and II presumably correspond to external vote monotonicity, which, by Theorem 4, is at least as strong a condition as internal vote monotonicity for consistent methods.)

Theorem 16

An allotment method F is consistent and internally vote monotone, encourages merger and restricts gain by merger if and only if $F = F_{HA}$.

Sketch of Proof

F_{HA} satisfies the conditions. Conversely, it is sufficient to prove that the conditions imply that F is weakly balanced. Then the proof of Theorem 15(a) can be applied, since that proof uses the condition that F is strongly balanced (as opposed to just weakly balanced) only to derive vote monotonicity.

First it can be proved that for any positive integers m and n, there exist r and s such that

$$F(n; mx, x) \rightarrow (r, s)$$

and $r \geq \min(m, n)$. The proof is by induction on m.

If there exist x, r_1 and r_2 with

$$F(r_1+r_2; x, x) \rightarrow (r_1, r_2)$$

and $r_1 \geq r_2+3$, one can conclude that

$$F(r_1+r_2; x, \frac{x}{r_1}, x) \rightarrow (r_2, 0, r_1)^{131}$$

Merging the first two parties will give a counterexample to vote monotonicity, since F restricts gain by merger.

If

$$F(2r+2; x, x) \rightarrow (r+2, r),$$

one can find s_1, s_2 and s_3 such that

$$F(3r+4; \frac{x}{r+2}, x, x, x) \rightarrow (0, s_1, s_2, s_3).$$

By consistency, $s_3 \geq s_2 \geq s_1$ can be assumed. $s_3 \geq r + 3$ will imply $s_3 \geq s_1 + 3$, which is impossible by the above and consistency. Two possibilities remain: (A) $s_3 = s_2 = r + 2$, $s_1 = r$. Then the first two parties can be merged, giving a contradiction to internal vote monotonicity. (B) $s_1 = s_2 = r + 1$, which implies that F is weakly balanced in the situation (2r+2; x, x).

Consistency now implies that F is weakly balanced, and Theorem 16 follows.

□

A corresponding result for F_{SD} would be:

An allotment method F is consistent and satisfies condition (13.8), discourages merger and restricts loss by merger if and only if $F = F_{SD}$.

I have, however, not been able to prove this without introducing membership monotonicity in the premise, in which case it becomes an immediate consequence of the proof of Theorem 15(b). (The fact that F is balanced is only used to prove (13.8) and membership monotonicity.)

14. LARGE AND SMALL PARTIES

14.1 Definition

In political discussion of methods for proportional representation, an important issue is whether a certain method favors small parties, or favors large parties, or neither. An unqualified statement that "the method F favors large parties" is hardly meaningful. But one can ask whether a method favors large or small parties compared to something else, be it a criterion[132] or another method. The latter possibility suggests the following definition.

Definition 18

Let F and G be two allotment methods. F favors small parties compared to G if

$$F(n; x_1, \ldots, x_k) \to (r_1, \ldots, r_k),$$

$$G(n; x_1, \ldots, x_k) \to (s_1, \ldots, s_k),$$

$$x_i < x_j, \text{ and}$$

$$r_i + r_j = s_i + s_j$$

imply

$$r_i \geq s_i.$$ [133]

The condition says that if party i is smaller than party j and F and G give the two parties the same total number of seats, then F shall give the smallest party at least as many seats as G gives it. The definition treats F and G symmetrically; hence there is no need for a separate condition "favors large parties."

Because of the possibility of ties, F will generally not favor small parties compared to F.[134] If F favors small parties compared to G, $F' \subseteq F$ and $G' \subseteq G$, then F' favors small parties compared to G'.

Definition 18 does not give rise to a complete relation among methods. That is, there exist two methods such that neither favors small parties compared to the other. For consistent methods, the relation is transitive.

That is, if F favors small parties compared to F' which favors small parties compared to F", then F favors small parties compared to F".[135]

In Definition 18, F and G are applied to the same situation. An alternative formulation would have allowed two different situations and would have compared the representation for two parties with equal votes in the two situations and equal total representation.[136] For consistent methods, the two formulations are equivalent; in general, the latter version is stronger than Definition 18.

14.2 Divisor methods

For divisor methods, it can be easily determined from the divisors whether one method favors small parties compared to another.

Theorem 17

Let F and F' be complete or partial divisor methods, defined by the divisor sequences d_1, d_2, ..., and d_1', d_2', ..., respectively.

(a) If $\frac{d_\alpha}{d_\beta} > \frac{d_\alpha'}{d_\beta'}$, for all α and β with $\alpha > \beta \geq 1$, then F favors small parties compared to F'.[137]

(b) If F favors small parties compared to F', then $\frac{d_\alpha}{d_\beta} \geq \frac{d_\alpha'}{d_\beta'}$ for all α and β with $\alpha > \beta \geq 1$.

Proof

(a) Assume that $\frac{d_\alpha}{d_\beta} > \frac{d_\alpha'}{d_\beta'}$ for all $\alpha > \beta$, and suppose that the premise of Definition 18 holds. Then $\frac{x_j}{d_{r_j}} \geq \frac{x_i}{d_{r_i+1}}$, by the definition of F. If $r_j > r_i + 1$, the assumption then gives $\frac{x_j}{d'_{r_j}} > \frac{x_i}{d'_{r_i+1}}$. If $r_j = r_i + 1$, the same formula follows because $x_j > x_i$. Hence, when the method F' is used, party j must get its seat number r_j before party i gets number $r_i + 1$. If $r_j \leq r_i$, the same follows by internal vote monotonicity of the method F'. $r_i + r_j = s_i + s_j$ now implies $s_i \leq r_i$, and the proof is complete.

(b) Suppose that $\frac{d_\alpha}{d_\beta} < \frac{d_\alpha'}{d_\beta'}$ for some α and β with $\alpha > \beta$. Then there exist positive integers x_1 and x_2 such that $\frac{d_\alpha}{d_\beta} < \frac{x_1}{x_2} < \frac{d_\alpha'}{d_\beta'}$. $\alpha > \beta$ implies $d_\alpha \geq d_\beta$, and hence $x_1 > x_2$. Moreover,

$$F(\alpha+\beta-1;\ x_1,\ x_2) \rightarrow (r_1,\ r_2)$$

implies $r_2 \leq \beta - 1$, while

$$F'(\alpha+\beta-1;\ x_1,\ x_2) \rightarrow (s_1,\ s_2)$$

implies $s_2 \geq \beta$. Hence $r_2 < s_2$. A counterexample to Definition 18 has been constructed, and the proof is complete.

□

The theorem holds and the proof applies even if the strengthened version of Definition 18 is used.[138] If F and F' are strict, complete divisor methods and all their divisors, when appropriately normalized, are rational numbers, strict inequality is obtained in (b).[139] The proof works even if $d_1 = 0$ or $d_1' = 0$ is allowed (see Definition 14(c) in Section 8.1).

The theorem implies, for example, that F_{SD} favors small parties compared to F_{MF}, which favors small parties compared to F_{HA}. More generally, let a and a' be real numbers between -1 and 1, and let F and F' be defined from these numbers by equation (8.19) of Section 8.5. Then F favors small parties compared to F' if and only if a < a'.

14.3 Preserving the majority

If a party gets a majority of the total vote, it is certainly a "large" party. It might seem reasonable to require that such a party get at least half the seats. This motivates the following definition.

Definition 19

An allotment method F preserves the majority if

$$F(n;\ x_1,\ \ldots,\ x_k) \rightarrow (r_1,\ \ldots,\ r_k)$$

and

$$x_i > \sum_{j \neq i} x_j$$

imply

$$r_i \geq \frac{n}{2}.$$

The idea behind this concept is related to the one underlying the lower bound condition; see Definition 6. Both conditions establish a "minimum requirement" for a party's representation. In general, preservation of the majority is the weaker condition; it only applies in certain situations and to the largest party, while the lower bound condition is non-vacuous for all parties whose "exact representation" is at least one. But when the number of seats is odd, none of the two conditions implies the other.[140]

The converse of Definition 19, that is, a requirement that a party can get a majority of the seats only if it has a majority of the vote, does not seem like a reasonable condition. (I am here thinking of proportional elections, not other uses of allotment methods.) It would imply that in a situation where one party has almost half the vote and the other parties are small, seats would have to be given to parties with almost no votes.

The method of the highest average preserves the majority. Within the class of consistent methods, F_{HA} will essentially be the boundary between those methods which do and those which do not preserve the majority, when methods are ordered by the relation of Definition 18.

<u>Theorem 18</u>

(a) If F is a submethod of F_{HA}, then F preserves the majority.

(b) Let F be a consistent allotment method and assume that F_{HA} favors small parties compared to F. Then F preserves the majority.

(c) Let F be a consistent allotment method, and assume that F favors small parties compared to F_{HA} and is not a submethod of F_{HA}. Then F does not preserve the majority.

Proof

(a) Let $F \subseteq F_{HA}$, and assume

$$F(n;\ x_1,\ \ldots,\ x_k) \rightarrow (r_1,\ \ldots,\ r_k)$$

and

$$x_i > \sum_{j \neq i} x_j \ .$$

For all j, one has

$$\frac{x_i}{r_i+1} \leq \frac{x_j}{r_j} \ . \qquad (14.1)$$

Summing this for all $j \neq i$ gives

$$\frac{x_i}{r_i+1} \leq \frac{\sum_{j \neq i} x_j}{\sum_{j \neq i} r_j} \ .$$

Together with the assumptions, this implies

$$r_i + 1 > \sum_{j \neq i} r_j \ ,$$

which gives

$$r_i \geq \frac{n}{2} \ ,$$

and (a) is proved. (Cases where $r_j = 0$ for some or all $j \neq i$ must formally be treated separately but cause no problems.)

(b) Let F be consistent, and suppose that F_{HA} favors small parties compared to F. Assume $F(n;\ x_1,\ \ldots,\ x_k) \rightarrow (r_1,\ \ldots,\ r_k)$ and $x_i > \sum_{j \neq i} x_j$. Then, clearly, $x_i > x_j$ for all $j \neq i$. The assumptions now imply that if s_i and s_j are chosen such that $F_{HA}(r_i + r_j;\ x_i,\ x_j) \rightarrow (s_i,\ s_j)$, then $s_i \leq r_i$ and $s_j \geq r_j$. (14.1) follows, and the proof can proceed as in (a).

(c) Suppose that F is a consistent method which favors small parties compared to F_{HA} and preserves the majority. Moreover, assume that F is not a submethod of F_{HA}. From this a contradiction will be derived.

F_{HA} is consistent and membership monotone. Lemma 1(d) applies, and there exists a two-party counterexample to $F \subsetneq F_{HA}$. That is, there exist x_1, x_2, r_1 and r_2 such that

(14.2) $$F(r_1+r_2;\ x_1,\ x_2) \rightarrow (r_1,\ r_2)$$

and

(14.3) $$\frac{x_1}{r_1} < \frac{x_2}{r_2+1}\ .$$

(14.3) implies the existence of s_1 and s_2 with $s_2 > r_2$ and $F_{HA}(r_1+r_2;\ x_1,\ x_2) \rightarrow (s_1,\ s_2)$. If $x_1 > x_2$, this contradicts the assumption that F favors small parties compared to F_{HA}.

Now it will be proved that F is strongly balanced. If not, consistency implies the existence of x, r and s such that $r \leq s - 2$ and $F(r+s;\ x,\ x) \rightarrow (r,\ s)$. It is possible to find a rational number[141] y such that $y > 2x$ but $\frac{y}{2x} < \frac{r+s+2}{r+s+1}$ and hence $\frac{y}{2x} < \frac{r+s+1}{r+s}$. Then choose t_1, t_2 and t_3 such that

$$F(2r+2s+1;\ x,\ x,\ y) \rightarrow (t_1,\ t_2,\ t_3).$$

$t_3 \geq r + s + 1$, since F preserves the majority. If $t_3 > r + s + 1$, then $t_1 + t_2 \leq r + s - 1$, and there is no loss of generality in assuming $t_1 \leq \frac{r+s-1}{2}$. The choice of y implies $\frac{y}{t_3} \leq \frac{y}{r+s+2} < \frac{x}{\frac{1}{2}(r+s+1)} \leq \frac{x}{t_1+1}$. By consistency, the first and third party can now be used to construct an example satisfying (14.2) and (14.3), with $x_1 > x_2$. (Use y, x, t_3 and t_1 for x_1, x_2, r_1 and r_2, respectively.) This is impossible by an earlier argument. Hence $t_3 = r + s + 1$. Then $t_1 + t_2 = r + s$, and $t_1 = r$ can be assumed by consistency. Then

$$\frac{y}{t_3} < \frac{x}{\frac{1}{2}(r+s)} \leq \frac{x}{t_1+1}\ ,$$

and again an example can be constructed in which (14.2) and (14.3) hold and $x_1 > x_2$. This contradiction proves that F is strongly balanced.

Theorem 3 now implies that F is membership monotone.

In (14.2) and (14.3), the possibility $x_1 > x_2$ has already been ruled out. If $x_1 = x_2$, then $r_1 > r_2 + 1$, contradicting the fact that F is strongly balanced. Hence $x_1 < x_2$. Since F preserves the majority, $r_1 \leq r_2$. It is possible to find a positive rational number y such that $y < x_2 - x_1$ and

$$\frac{x_2}{r_2+1} < \frac{y}{r_2-r_1+1} . \tag{14.4}$$

Choose s_1, s_2 and s_3 such that

$$F(2r_2+1;\ x_1,\ x_2,\ y) \rightarrow (s_1,\ s_2,\ s_3),$$

and such that s_2 is minimized. $x_2 > x_1 + y$, and since F preserves the majority, $s_2 \geq r_2 + 1$. If $s_1 < r_1$, (14.2) and Lemma 1(b) can be used to transfer one or more seats from the second to the first party, contradicting the choice of s_2. Hence $s_1 \geq r_1$, which implies $s_3 \leq r_2 - r_1$. But then the representation of the second and third party, together with (14.4), contradicts the assumption that F favors small parties compared to F_{HA}. The proof is complete.

□

Part (b) is true if the number of parties is limited to K for any $K \geq 2$. If $K \geq 4$, the same is true for part (c). (Four parties are needed in order to apply Theorem 3. If one is willing to assume membership monotonicity, the case K = 3 can be included.) The assumption that F is not a submethod of F_{HA}, is essential in part (c).[142]

If F is a complete or partial divisor method, the assumption of consistency can be dropped, and parts (b) and (c) will still hold.[143] Formally, this statement is neither more nor less general than the theorem, since partial divisor methods need not be consistent.

14.4 Quota methods

For quota methods, defined in Section 4.3, there exists a result similar to Theorem 17.

Theorem 19

Let F and F' be quota methods defined by the quota criteria V and V' respectively.

(a) If $V(n, y) > V'(n, y)$ for all y and all $n \geq 2$, then F favors small parties compared to F'.

(b) If F favors small parties compared to F', then $V(n, y) \geq V'(n, y)$ for all y and all $n \geq 2$.

According to Definition 5, $V(n, y)$ and $V'(n, y)$ are only defined when $n > 1$. Hence the condition of Theorem 19 only applies to such values of n.

Proof

(a) Assume that F and F' satisfy the premise but not the conclusion. Then there exist n, $\overline{x}$, $\overline{r}$, $\overline{s}$, i and j such that

$$F(n; \overline{x}) \to \overline{r},$$

$$F'(n; \overline{x}) \to \overline{s},$$

$$x_i < x_j,$$

$$r_i + r_j = s_i + s_j$$

and

$$r_i < s_i.$$

$n \leq 1$ implies $s_i = 0$, hence $n > 1$ can be assumed. Let y be the sum of the votes in the vector $\overline{x}$, and define

$$\hat{x}_i = \frac{x_i}{V(n, y)},$$

$$\tilde{x}_i = \frac{x_i}{V'(n, y)},$$

and similarly for party j. Then $\tilde{x}_j > \hat{x}_j$. r_j is either $\lfloor \hat{x}_j \rfloor$ or $\lfloor \hat{x}_j \rfloor + 1$, while s_j is either $\lfloor \tilde{x}_j \rfloor$ or $\lfloor \tilde{x}_j \rfloor + 1$. Since $r_j > s_j$, this is only possible if

$$r_j = \lfloor \hat{x}_j \rfloor + 1$$

and

(14.5) $$s_j = \lfloor \tilde{x}_j \rfloor = \lfloor \hat{x}_j \rfloor.$$

The assumptions and (14.5) imply

(14.6) $$0 < \tilde{x}_i - \hat{x}_i < \tilde{x}_j - \hat{x}_j < 1.$$

If $\lfloor \tilde{x}_i \rfloor = \lfloor \hat{x}_i \rfloor$, $r_i < s_i$ implies $r_i = \lfloor \hat{x}_i \rfloor$ and $s_i = \lfloor \tilde{x}_i \rfloor + 1$. Since F gives party j a seat on the basis of its remainder while party i gets no such seat, $\hat{x}_j - \lfloor \hat{x}_j \rfloor \geq \hat{x}_i - \lfloor \hat{x}_i \rfloor$. Similarly, $\tilde{x}_i - \lfloor \tilde{x}_i \rfloor \geq \tilde{x}_j - \lfloor \tilde{x}_j \rfloor$. Together with the assumption $\lfloor \tilde{x}_i \rfloor = \lfloor \hat{x}_i \rfloor$, this contradicts (14.5) and (14.6). The only other possibility permitted by (14.6) is $\lfloor \tilde{x}_i \rfloor = \lfloor \hat{x}_i \rfloor + 1$. Then (14.6) gives $\hat{x}_i - \lfloor \hat{x}_i \rfloor > \hat{x}_j - \lfloor \hat{x}_j \rfloor$ and $\tilde{x}_j - \lfloor \tilde{x}_j \rfloor > \tilde{x}_i - \lfloor \tilde{x}_i \rfloor$. Therefore, $r_i = \lfloor \hat{x}_i \rfloor + 1$ and $s_i = \lfloor \tilde{x}_i \rfloor$, contradicting $r_i < s_i$.

(b) If the conclusion fails, there exist y and n such that $n > 1$ and $V(n, y) < V'(n, y)$. A rational number x can now be chosen such that[144]

$$\frac{y - (n-1)V'(n, y)}{2} < x < \frac{y - (n-1)V(n, y)}{2}.$$

x is positive since V' satisfies (4.3) of Definition 5. It is also clear that $x < \frac{y}{2}$. Now $F(n; y-x, x) = (n, 0)$, while $F'(n; y-x, x) \rightarrow (r_1, r_2)$ implies $r_2 > 0$. Hence F does not favor small parties compared to F', and the proof is complete.

□

The conditions (4.2) and (4.3) of Definition 5 are never explicitly used in the proof of (a). But it is assumed that the number of seats a party gets is either the whole number of times the quota divides its vote, or one more than this. If F and F' are methods given by quota criteria that do not satisfy (4.2) and (4.3), Theorem 19(a) will still apply to any situation in which the latter statement holds for both methods. Part (b) uses (4.3) but not (4.2) of Definition 5.

Part (a) does not hold if Definition 18 is strengthened as indicated in the last paragraph of Section 14.1. If V(n, y) and V'(n, y) are always rational numbers (or irrational votes are allowed), strict inequality can be obtained in the conclusion of (b).

If F and F' are defined by equation (8.18) of Section 8.5, from real numbers a and a', respectively, then F will favor small parties compared to F' if and only if $a < a'$. This also holds in the limiting cases where a or a' is 1 or -1. In particular, the method of the largest remainder favors small parties compared to the method obtained by setting $a = 1$, that is, the method obtained by going to the limit of (4.2) in Definition 5.[145]

The limiting method just mentioned will preserve the majority. This holds even if ties are broken against the largest party. But no quota method which satisfies (4.2) preserves the majority; a situation with three parties is sufficient to provide a counterexample.

15. PRIORITY METHODS[146]

15.1 Definitions

The class of divisor methods, studied in Sections 7, 8 and 11, is a fairly wide and general class of allotment methods. In particular, Theorem 10 shows that the class contains every method which satisfies a set of reasonable conditions given there.

As mentioned in the comments to Theorem 10(b), the characterization of that theorem is not, however, entirely satisfactory; see also the comments to Theorem 5(b). The problem is that a divisor method will allow relatively many ties, that is, there are many situations in which more than one allotment is possible. Some ties will occur in any consistent method.[147] But as pointed out in connection with Theorem 10, it is possible to find proper submethods of divisor methods which satisfy all the requirements of that theorem.[148] That is, at least some ties can be eliminated without destroying consistency and other desirable properties.

A complete divisor method distributes the seats according to a certain criterion of priority, based on quotients of the form $\frac{x}{d_r}$. When quotients are equal, there is a tie. A way of breaking the ties could be to refine the priority criterion so that one party sometimes is placed ahead of another even if the quotients are equal. This is the idea behind the approach in this section.

Definition 20

(a) A priority relation is a binary relation R defined on all pairs of the form (x, r), where x is a positive rational number (the vote of a party) and r is a positive integer (a number of seats), such that

(i) R is transitive, that is, whenever (x, r)R(y, s) and (y, s)R(z, t), then (x, r)R(z, t),

and

(ii) R is complete, that is, for all x, y, r and s, at least one of (x, r)R(y, s) and (y, s)R(x, r) is true.[149]

(b) A priority relation R is <u>semi-strict</u> if (x, r)R(x, r+1) for all x and r.

(c) A priority relation R is <u>strict</u> if (x, r+1)R(x, r) is false for all x and r.

A pair (x, r) is supposed to represent the priority for a party with x votes of getting its seat number r. The relation R shall mean "has at least as high a priority as." Therefore, if (x, r)R(y, s) and a situation is given in which one party has x votes and another has y votes, the former party has at least as strong a claim on its seat number r as the latter has on its seat number s.[150]

When R is a priority relation, P will denote the relation defined by

(15.1) (x, r)P(y, s) if and only if (x, r)R(y, s) and not (y, s)R(x, r).[151]

It can be proved that P is transitive. P is also asymmetric, that is, (x, r)P(y, s) and (y, s)P(x, r) cannot both be true. In particular, P is irreflexive, that is, (x, r)P(x, r) is wrong for all x and r. Clearly, (x, r)P(y, s) implies (x, r)R(y, s).[152] The intuitive meaning of P is "has a strictly higher priority than."

If the priority relation R is semi-strict, a party will never have a higher priority of getting its next seat than it had of getting the last one. More generally, an earlier seat always has at least as high a priority as a later one; Definition 20(b) implies that (x, r)R(x, s) whenever $r \leq s$. For a strict relation R, $r < s$ implies (x, r)P(x, s). Hence the priority for a later seat will be strictly lower than the priority for an earlier one. Any strict relation is semi-strict.

It should now be clear how an allotment method is defined from a given priority relation.

<u>Definition 21</u>

An allotment method F is a <u>priority method</u>[153] if there exists a priority relation R such that F can be defined recursively in the following way:

(i) $F(0; x_1, \ldots, x_k) = (0, \ldots, 0)$.

(ii) For $n \geq 1$, $F(n; x_1, \ldots, x_k)$ shall consist of all allotments $(s_1, \ldots, s_k)$ which can be constructed as follows:

Take any $(r_1, \ldots, r_k)$ in $F(n-1; x_1, \ldots, x_k)$, and find an i such that

$$(x_i, r_i+1)R(x_j, r_j+1) \text{ for all } j = 1, \ldots, k.$$

Then let

$$s_i = r_i + 1$$
$$s_j = r_j \text{ for } j \neq i .$$

F is a <u>semi-strict</u> priority method (a <u>strict</u> priority method) if R can be chosen semi-strict (strict).

It is easily seen that for any priority relation R and any $x_1, \ldots, x_k$ and $r_1, \ldots, r_k$, there exists an i such that $(x_i, r_i+1)R(x_j, r_j+1)$ for all $j = 1, \ldots, k$. (This i need not be unique.) Hence any priority relation R will define a priority method F. Except for certain uninteresting cases involving relations that are not semi-strict, there exists only one R which corresponds to a given priority method.[154]

Ties can and will occur in any priority method; their frequency depends on how "fine" the priority relation is. In particular, $F(n; x_1, \ldots, x_k)$ can get more than one element because $F(n-1; x_1, \ldots, x_k)$ has more than one element, or because there is more than one i such that

$$(x_i, r_i+1)R(x_j, r_j+1) \text{ for all } j = 1, \ldots, k.$$

Any divisor method is a priority method. Specifically, the divisor method given by the divisors $d_1, d_2, \ldots,$ can be derived from the priority relation given by

$$(x, r)R_d(y, s) \text{ if and only if } \frac{x}{d_r} \geq \frac{y}{d_s} .$$

All generalized divisor methods (see Definition 16) are also priority methods; the corresponding priority relation can easily be constructed from the numbers M_β and $d^{(\beta)}_\alpha$.

The partial divisor method described in note 50, which is a proper submethod of the divisor method F_{HA}, is also a priority method. Its priority relation is given by

$$(x, r)R(y, s) \text{ if and only if } \frac{x}{r} > \frac{y}{s} \text{ or}$$

$$(\frac{x}{r} = \frac{y}{s} \text{ and } x \geq y).$$

Any priority relation can be represented by a real-valued function. That is, for any priority relation R there exists a function f_R such that

$$(x, r)R(y, s) \text{ if and only if } f_R(x, r) \geq f_R(y, s).$$ [155]

Conversely, if a function F_R is given and R is defined as above, R will be a priority relation. Sometimes, it may be convenient to represent a priority method by f_R rather than by R. For example, for the divisor method with divisors d_1, d_2, ..., a representing function is given by

$$f_{R_d}(x, r) = \frac{x}{d_r} .$$

In other cases, the representing function offers no advantage. For the method of note 50, such a function does exist as a mathematical object. But it is not easy to describe such a function, and if the function was given by an algorithm or a formula, it would not necessarily be obvious which method it represented. Theoretically, it makes no difference whether priority methods are represented by relations or functions.[156]

Let F be defined by the priority relation R. It can be proved that F is (externally and internally) vote monotone if

$$x > y \text{ implies } (x, r)P(y, r), \text{ for all } x, y \text{ and } r.$$

Moreover, F is scale independent if

$(x, r)R(y, s)$ if and only if $(ax, r)R(ay, s)$, for all x, y, r and s, and all positive rational numbers a.

If R is semi-strict, these conditions are also necessary for the corresponding properties of F.[157]

15.2 Consistent and membership monotone methods

It is immediate from Definition 21 that any priority method is membership monotone. Perhaps one would expect all priority methods to be consistent as well. For methods that are not semi-strict this is, however, not necessarily the case.[158] But the converse implication holds.

Theorem 20

Let F be a consistent and membership monotone allotment method. Then F is a priority method.[159]

Proof

Let F be consistent and membership monotone, and define the relation R^* by

$$(x, r)R^*(y, s) \text{ if and only if}$$
$$F(r+s-2;\ x,\ y) \rightarrow (r-1,\ s-1)$$
$$\text{and } F(r+s-1;\ x,\ y) \rightarrow (r,\ s-1),$$

for x and y positive rational numbers and r and s positive integers. R^* is, in a sense, the priority relation induced by F; $(x, r)R^*(y, s)$ means that there is a situation in which a party with x votes is in line for its seat number r while a party with y votes is in line for its seat number s, and the next seat is given to the former party.

In general, R^* is not complete and transitive, hence it is not a priority relation. But it does have a property related to transitivity, given by (15.2) below. Let P^* be defined from R^* as P was defined from R by (15.1). If $(x, r)R^*(y, s)$ and $(y, s)R^*(z, t)$, transitivity would require $(x, r)R^*(z, t)$. This need not hold, but one can at least conclude that $(z, t)P^*(x, r)$ does not hold. Similar conditions hold for longer chains of pairs connected by R^*. This turns out to be sufficient to continue the proof.

Formally, the following statement holds:

(15.2) There does not exist any sequence $(x_1, r_1), \ldots, (x_k, r_k)$ of pairs, where $k \geq 2$, $x_1, \ldots, x_k$ are positive rational numbers and $r_1, \ldots, r_k$ are positive integers, such that

$$(x_i, r_i)R^*(x_{i+1}, r_{i+1}) \text{ for } i = 1, \ldots, k-1$$

and

$$(x_k, r_k)P^*(x_1, r_1).$$

To prove (15.2), assume the opposite, that is, suppose that such a sequence exists. The definitions of R^* and P^* give

(15.3) $F(r_i+r_{i+1}-2; x_i, x_{i+1}) \rightarrow (r_i-1, r_{i+1}-1)$ for $i = 1, \ldots, k-1$,

(15.4) $F(r_i+r_{i+1}-1; x_i, x_{i+1}) \rightarrow (r_i, r_{i+1}-1)$ for $i = 1, \ldots, k-1$,

(15.5) $F(r_1+r_k-2; x_1, x_k) \rightarrow (r_1-1, r_k-1)$,

(15.6) $F(r_1+r_k-1; x_1, x_k) \rightarrow (r_1-1, r_k)$,

and

(15.7) not $F(r_1+r_k-1; x_1, x_k) \rightarrow (r_1, r_k-1)$.

For a given n, one can ask whether it is possible to find $s_1, \ldots, s_k$ such that

(15.8) $s_i < r_i$ for $i = 1, \ldots, k$, and

(15.9) $F(n; x_1, \ldots, x_k) \rightarrow (s_1, \ldots, s_k)$.

Since $r_1, \ldots, r_k$ are positive integers, this clearly is possible for $n = 0$, but for sufficiently large n some s_i would have to be greater than or equal to r_i. Let n be the largest number for which (15.8) and (15.9) can be satisfied, and choose $s_1, \ldots, s_k$ accordingly.

By membership monotonicity and the choice of n, there will now exist $t_1, \ldots, t_k$ and i such that

(15.10) $F(n+1; x_1, \ldots, x_k) \to (t_1, \ldots, t_k)$

(15.11) $t_j = s_j < r_j$ for $j \neq i$

(15.12) $t_i = s_i + 1 = r_i$.

Choose i as small as possible for (15.10) - (15.12) to hold.

If $i > 1$, consider the two parties $i - 1$ and i. If $t_{i-1} = s_{i-1} = r_{i-1} - 1$, (15.4) and consistency can be used to move one seat from party i to party $i - 1$ in the allotment given by (15.10). This contradicts the choice of i. Therefore, $t_{i-1} \leq r_{i-1} - 2$, which implies $t_{i-1} + t_i \leq r_{i-1} + r_i - 2$. Then (15.3) and Lemma 1(b) can be used to transfer one or more seats from party i to party $i - 1$, after which $t_j \leq r_j - 1$ for $j = i - 1, i$. This contradicts the choice of n.

Hence $i = 1$. If $t_k \leq r_k - 2$, Lemma 1(b) can be applied to (15.5), in the same way as above, to get a contradiction to the choice of n. The only remaining possibility is $t_k = r_k - 1$. Since $t_1 = r_1$, consistency now contradicts (15.7). The proof of (15.2) is complete.[160]

Now it is possible to extend R* to a transitive and complete relation. That is, there exists a priority relation R such that, for all x, y, r and s:

(15.13) $(x, r)R^*(y, s)$ implies $(x, r)R(y, s)$,

and

(15.14) $(x, r)P^*(y, s)$ implies $(x, r)P(y, s)$.

(P is defined from R as before.) The meaning of (15.13) is that if (x, r) has at least as high a priority as (y, s) according to the incomplete relation R*, then the same is true in R. And (15.14) says that if this relation is strict in R*, it is also strict in R. Note that if (x, r) and (y, s) are incomparable in R*, nothing is said about their relation in R; they can have equal priority, or one can have higher priority than the other. R is complete, hence any two pairs must be comparable.[161]

Let F_0 be constructed from R by Definition 21. The proof will be complete if it can be shown that F and F_0 are equal. Let k and $x_1, \ldots, x_k$ be given. $F(n; x_1, \ldots, x_k) = F_0(n; x_1, \ldots, x_k)$ will be proved by induction on n.

All methods are equal when there are no seats to distribute. Let a positive integer n be given, and assume that F and F_0 are equal on $(n-1;\ x_1,\ \ldots,\ x_k)$. This is the induction hypothesis.

Suppose

$$F(n;\ x_1,\ \ldots,\ x_k) \rightarrow (r_1,\ \ldots,\ r_k). \tag{15.15}$$

Since F is membership monotone, there exist $s_1, \ldots, s_k$ and i such that

$$F(n-1;\ x_1,\ \ldots,\ x_k) \rightarrow (s_1,\ \ldots,\ s_k) \tag{15.16}$$

$$s_j = r_j \text{ for } j \neq i \tag{15.17}$$

$$s_i = r_i - 1\ . \tag{15.18}$$

Since F is consistent, (15.16) and (15.15) give

$$F(s_i+s_j;\ x_i,\ x_j) \rightarrow (s_i,\ s_j)$$

and

$$F(s_i+s_j+1;\ x_i,\ x_j) \rightarrow (s_i+1,\ s_j)$$

for all $j \neq i$. This gives

$$(x_i,\ s_i+1)R^*(x_j,\ s_j+1) \text{ for all } j \neq i.$$

(15.13) and the completeness of R then give

$$(x_i,\ s_i+1)R(x_j,\ s_j+1) \text{ for all } j. \tag{15.19}$$

The induction hypothesis is

$$F_0(n-1;\ x_1,\ \ldots,\ x_k) \rightarrow (s_1,\ \ldots,\ s_k). \tag{15.20}$$

The definition of F_0 gives the desired conclusion, namely

$$F_0(n;\ x_1,\ \ldots,\ x_k) \rightarrow (r_1,\ \ldots,\ r_k). \tag{15.21}$$

Conversely, assume that $r_1, \ldots, r_k$ satisfy (15.21). By the definition of F_0, there must exist $s_1, \ldots, s_k$ and i satisfying (15.17) - (15.20). The induction hypothesis gives (15.16). By membership monotonicity of F, there must now exist $t_1, \ldots, t_k$ and i' such that

(15.22) $$F(n; x_1, \ldots, x_k) \rightarrow (t_1, \ldots, t_k),$$
$$t_j = s_j \text{ for } j \neq i', \text{ and } t_{i'} = s_{i'} + 1.$$

If i' = i, (15.22) is the same as (15.15), which completes the induction step. Otherwise,

$$(x_{i'}, s_{i'}+1)R^*(x_i, s_i+1)$$

can be derived from (15.16) and (15.22), by consistency of F. If $(x_{i'}, s_{i'}+1)P^*(x_i, s_i+1)$, (15.14) implies $(x_{i'}, s_{i'}+1)P(x_i, s_i+1)$, which contradicts (15.19). Hence

$$(x_i, s_i+1)R^*(x_{i'}, s_{i'}+1).$$

The definition of R* then requires

$$F(s_i+s_{i'}+1; x_i, x_{i'}) \rightarrow (s_i+1, s_{i'}).$$

Consistency can then be applied to (15.22), to give (15.15).

Hence it follows that F and F_0 are equal for situations in which n seats shall be distributed. By induction, $F = F_0$. Theorem 20 is proved.

□

It should be noted that the proof of (15.2) uses the assumption that the number of parties is unlimited. The rest of the proof of Theorem 20 makes no use of situations which have more parties than the situations about which statements are made. In fact, (15.2) and Theorem 20 hold even if the number of parties is limited to K for $K \geq 3$.[162] The proof of this statement is a little complicated; therefore, it was not incorporated in the proof above but is given separately below.

Now let F be consistent and membership monotone when the number of parties is limited to K for $K \geq 3$. According to what has just been said, (15.2) holds, and R and F_0 can be constructed as in the proof above. F_0 is defined for any number of parties, and the proof shows that F and F_0 are equal when there are K or fewer parties. F_0 is membership monotone, and it can also be proved that F_0 is consistent.[163] Hence F can be consistently extended to situations with any number of parties. The extension is unique, by Lemma 1(d).

Proof of (15.2) when the number of parties is limited

Assume that (15.2) is false, and find the shortest possible sequence $(x_1, r_1), \ldots, (x_k, r_k)$ which provides a counterexample. $k = 2$ will immediately contradict the definition of P*. If $k = 3$, the proof above will not use more than three parties and can be applied. The case $k \geq 4$ remains.

If $(x_1, r_1)R^*(x_{k-1}, r_{k-1})$, then the pairs $(x_2, r_2), \ldots, (x_{k-2}, r_{k-2})$ can be deleted from the sequence, and it will still provide a counterexample to (15.2). Since $k \geq 4$, at least one pair is deleted and the sequence is made shorter. This contradicts the original choice of sequence. If $(x_{k-1}, r_{k-1})P^*(x_1, r_1)$, the pair (x_k, r_k) can be deleted while the violation of (15.2) is maintained. Again the original choice of sequence is contradicted. The conclusion is that (x_1, r_1) and (x_{k-1}, r_{k-1}) must be incomparable according to R*.

The next step is to prove the following:

(15.23) Let $m \geq 2$ and let $(x_1, r_1), \ldots, (x_m, r_m)$ be given. If $(x_i, r_i)R^*(x_{i+1}, r_{i+1})$ for all $i = 1, \ldots, m - 1$, then there exists $s_m < r_m$ such that

$$F(r_1+s_m; x_1, x_m) \rightarrow (r_1, s_m).$$

For $m = 2$, this is immediate from the definition of R*; s_2 can be chosen equal to $r_2 - 1$. To prove (15.23) by induction, assume that it holds for all sequences of length m, and consider a sequence of length $m + 1$. The induction hypothesis guarantees the existence of $s_m < r_m$ satisfying

$$F(r_1+s_m; x_1, x_m) \rightarrow (r_1, s_m). \qquad (15.24)$$

This, together with $(x_m, r_m)R^*(x_{m+1}, r_{m+1})$, must be used to find $s_{m+1} < r_{m+1}$ such that

(15.25) $$F(r_1+s_{m+1};\ x_1,\ x_{m+1}) \rightarrow (r_1,\ s_{m+1})$$

Find the largest n for which there exist t_1, t_m, t_{m+1} satisfying $t_i \leq r_i - 1$ for i = 1, m, m + 1 and

$$F(n;\ x_1,\ x_m,\ x_{m+1}) \rightarrow (t_1,\ t_m,\ t_{m+1}).$$

Add a seat number n+1. By membership monotonicity, this seat can be given to one of the parties without affecting the others. Let the resulting allotment be $(t_1',\ t_m',\ t_{m+1}')$. Then $t_m' + t_{m+1}' \leq r_m + r_{m+1} - 1$; hence the second clause in the definition of $(x_m,\ r_m)R^*(x_{m+1},\ r_{m+1})$ can be used, together with Lemma 1(b), to redistribute seats between the two last parties to obtain $t_{m+1}' \leq r_{m+1} - 1$. By (15.24), the same lemma can then be used on the first two parties, resulting in either $t_1' \leq r_1$ and $t_m' \leq s_m \leq r_m - 1$, or $t_1' \geq r_1$. $t_1' < r_1$ will contradict the choice of n. Hence $t_1' \geq r_1$ and, by consistency,

(15.26) $$F(t_1' + t_{m+1}';\ x_1,\ x_{n+1}) \rightarrow (t_1',\ t_{m+1}').$$

If $t_1' = r_1$, this immediately gives (15.25), since $t_{m+1}' < r_{m+1}$. If $t_1' > r_1$, one can take one seat at a time away from the allotment of (15.26), each time using membership monotonicity to take the seat from one party and leave the representation of the other party unchanged. Eventually, the first party is brought down to r_1 seats, and we have (15.25) with some $s_{m+1} \leq t_{m+1}' < r_{m+1}$. (Formally, Lemma 1(a) can be invoked here.) This completes the proof of (15.23).

Returning to the assumed counterexample to (15.2), one can apply (15.23) for m = k - 1 and find the corresponding s_{k-1}. Choose t_1, t_{k-1} and t_k such that

(15.27) $$F(r_1+r_{k-1}+r_k-3;\ x_1,\ x_{k-1},\ x_k) \rightarrow (t_1,\ t_{k-1},\ t_k).$$

If the situation of (15.27) allows more than one allotment, t_{k-1} shall be chosen in the following way:

(i) If possible, choose $t_{k-1} \leq r_{k-1} - 1$.

(ii) Among the possible choices of t_{k-1} for which $t_{k-1} \leq r_{k-1} - 1$, or among those for which $t_{k-1} \geq r_{k-1}$ if $t_{k-1} \leq r_{k-1} - 1$ is impossible, choose t_{k-1} to minimize $|t_{k-1} - (r_{k-1} - 1)|$.

In other words, t_{k-1} shall be chosen equal to the earliest possible number in the sequence $r_{k-1} - 1$, $r_{k-1} - 2$, ..., 1, 0, r_{k-1}, $r_{k-1} + 1$, ...

If $t_{k-1} \geq r_{k-1}$, then either $t_1 < r_1 - 1$ or $t_k < r_k - 1$. In the former case, (15.23) and Lemma 1(b) can be used to transfer seats between the first two parties in such a way that either $t_1 \geq r_1$ or $t_{k-1} \leq s_{k-1} \leq r_{k-1} - 1$. This must reduce t_{k-1}, hence the original choice of t_{k-1} contradicted (i) or (ii). In the latter case, that is, if $t_k < r_k - 1$, reduction in t_{k-1} can similarly be obtained by applying Lemma 1(b) to the last two parties in (15.27), using the fact that

$$(15.28) \qquad F(r_{k-1}+r_k-2;\ x_{k-1},\ x_k) \rightarrow (r_{k-1} - 1,\ r_k-1),$$

which follows from the assumption $(x_{k-1}, r_{k-1})R^*(x_k, r_k)$. Hence $t_{k-1} \geq r_{k-1}$ is impossible.

If $t_{k-1} < r_{k-1} - 1$, then $t_1 + t_k \geq r_1 + r_k - 1$. $F(r_1+r_k-1;\ x_1,\ x_k) \rightarrow (r_1 - 1,\ r_k)$ follows from the assumption $(x_k, r_k)P^*(x_1, r_1)$; therefore, one can assume $t_1 \geq r_1 - 1$ and $t_k \geq r_k$. Then $t_{k-1} + t_k \leq r_{k-1} + r_k - 2$, and seats can be transferred from the third to the second party, by (15.28), such that $t_{k-1} \leq r_{k-1} - 1$ and $t_k \leq r_k - 1$. This transfer increases t_{k-1} without bringing it above $r_{k-1} - 1$, hence the original choice cannot have been made according to (ii).

Therefore, t_{k-1} must be equal to $r_{k-1} - 1$. The assumption (x_k, r_k) $P^*(x_1, r_1)$ implies $F(r_1+r_k-2;\ x_1,\ x_k) \rightarrow (r_1-1,\ r_k-1)$, and by consistency, it is possible to choose $t_1 = r_1 - 1$ and $t_k = r_k - 1$. Then consistency can be applied to the first two parties in (15.27) to get

$$F(r_1+r_{k-1}-2;\ x_1,\ x_{k-1}) \rightarrow (r_1-1,\ r_{k-1}-1).$$

By membership monotonicity, this implies that $F(r_1+r_{k-1}-1;\ x_1,\ x_{k-1})$ contains at least one of $(r_1,\ r_{k-1}-1)$ and $(r_1-1,\ r_{k-1})$. Hence either $(x_1, r_1)R^*(x_{k-1}, r_{k-1})$ or $(x_{k-1}, r_{k-1})R^*(x_1, r_1)$. But it was shown above that (x_1, r_1) and (x_{k-1}, r_{k-1}) are incomparable according to R^*. This contradiction completes the proof of (15.2).

□

15.3 A characterization of consistent and balanced methods

Presumably, one is mainly interested in methods which are balanced, in the strong or weak sense of Definition 11. One should expect strict and semi-strict priority relations to correspond to strongly balanced and weakly balanced methods, respectively. Indeed, this is the case. When these conditions are imposed, all priority methods are consistent. Hence a complete characterization of these methods is obtained.

Theorem 21

(a) An allotment method is consistent and weakly balanced if and only if it is a semi-strict priority method.

(b) An allotment method is consistent and strongly balanced if and only if it is a strict priority method.

The three classes of methods characterized by Theorems 20, 21(a) and 21(b) are really different. The method which gives all seats to the largest party (or to one of the largest parties in case of a tie), is consistent and membership monotone, but not weakly balanced. Hence it is a priority method, but it is not semi-strict.[164] F_T is consistent, membership monotone and weakly balanced; therefore, it is a semi-strict but not strict priority method. (The corresponding relation gives all pairs equal priority.) Finally, any strict divisor method, such as F_{HA}, is consistent, membership monotone and strongly balanced, and is therefore a strict priority method.

Proof

Assume that F is consistent and weakly balanced. By Theorem 3, F is membership monotone, and by Theorem 20, it is a priority method. Let R be the priority relation constructed in the proof of Theorem 20. Since F is weakly balanced,

$$F(2r; x, x) \to (r, r),$$

for any x and r. This allotment can be constructed by the procedure of Definition 21, and it is easy to see that this is possible only if (x, r)R(x, r+1). Hence R is semi-strict, and the "only if" part of (a) is proved.

If F is consistent and strongly balanced, everything of the above applies. In particular, R is semi-strict. If R is not strict, there exist x and r such that (x, r) and (x, r + 1) have equal priority according to R. Then

$$F(2r; x, x) \rightarrow (r + 1, r - 1),$$

which contradicts the assumption that F is strongly balanced. This proves the "only if" part of (b).

Conversely, assume that F is a semi-strict priority method, defined from the semi-strict priority relation R. In order to prove that F is consistent, assume that the premise of Definition 8 holds. Suppose that $r_{i_1} > s_{i_2}$ and $r_{j_1} < s_{j_2}$; if the inequalities go the other way, a similar proof applies; and if $r_{i_1} = s_{i_2}$, there is nothing to prove.

$F(n_1; \overline{x}) \rightarrow \overline{r}$, and it must be possible to construct this allotment by the procedure of Definition 21. In particular, party i_1 must have received its seat number r_{i_1} at a time when party j_1 had r_{j_1} or fewer seats. Therefore, there exists r' with $r' \leq r_{j_1}$ and

$$(x_{i_1}, r_{i_1})R(x_{j_1}, r' + 1).$$

A similar argument can be applied to the allotment $F(n_2; \overline{y}) \rightarrow \overline{s}$, to prove the existence of s' with $s' \leq s_{i_2}$ and

$$(y_{j_2}, s_{j_2})R(y_{i_2}, s' + 1).$$

Since $x_{i_1} = y_{i_2}$, $x_{j_1} = y_{j_2}$, $s' + 1 \leq s_{i_2} + 1 \leq r_{i_1}$ and $r' + 1 \leq r_{j_1} + 1 \leq s_{j_2}$, semi-strictness of R gives

$$(y_{i_2}, s' + 1)R(x_{i_1}, r_{i_1})$$

and

$$(x_{j_1}, r' + 1)R(y_{j_2}, s_{j_2}).$$

These four relations constitute a cycle of pairs connected by R. By transitivity of R, all the pairs involved must have equal priority. Semi-strictness implies that all the following pairs have equal priority:

$$(y_{i_2}, s' + 1), \ldots, (y_{i_2}, r_{i_1}),$$

$$(y_{j_2}, r' + 1), \ldots, (y_{j_2}, r_{j_1} + 1),$$

$$\ldots \quad (y_{j_2}, s_{j_2}).$$

Now consider a construction of the allotment $F(n_2; \overline{y}) \to \overline{s}$, as described in Definition 21. A new allotment for the same situation can be constructed, by distributing the seats exactly as in the old allotment until party j_2 is about to get its seat number r_{j_1}+1. Party i_2 now has at most s' seats, and by the above argument it must have at least as high a priority as party j_2 for getting the next seat. The priorities must, in fact, be equal, since party j_2 got the seat originally. Hence the next seat can be given to party i_2. This change from the original allotment does not change the highest of the priorities of the seats which are next in line for parties i_2 and j_2. Hence it is possible to go on awarding seats to parties different from i_2 and j_2, exactly as this was done in the original allotment. Every time the original allotment gives a seat to party j_2, an argument similar to the one above can be used to show that this seat can be given to party i_2 instead, as long as the number of seats for this party does not exceed r_{i_1}. But this number will never exceed r_{i_1}; when all the n_2 seats are distributed, party j_2 will have r_{j_1} seats, each party j different from i_2 and j_2 will have s_j seats, and hence party i_2 will have exactly r_{i_1} seats. This is the allotment $\overline{t}$, the existence of which should be proved. F is consistent.

If F is not weakly balanced, there must exist n and x such that

(15.29) $$F(n; x, x) \to (r, s)$$

implies $|r - s| > 1$. For given n and x, choose r and s to satisfy (15.29) such that $|r - s|$ is minimized. Assume $r > s$; the case $r < s$ is similar.

Since R is semi-strict and $r > s + 1$, $(x, s + 1)R(x, r)$. Hence one can prove that

$$F(n; x, x) \to (r - 1, s + 1),$$

by simulating the construction of allotment (15.29) but giving the first party's seat number r to the second party. $r - 1 \geq s + 1$ and $|(r-1) - (s+1)| < |r - s|$; therefore, this contradicts the choice of r and s. This contradiction shows that F is weakly balanced and completes the proof of the "if" part of (a).

The "if" part of (b) is now trivial. If F is a strict priority method defined by the strict relation R, F is semi-strict, and therefore consistent and weakly balanced by (a). If F is not strongly balanced, there exist x, r and s such that $r > s + 1$ and

$$F(r+s; x, x) \to (r, s).$$

Since the first party got its seat number r at some time, there must exist $s' \leq s$ with

$$(x, r)R(x, s' + 1).$$

But $s' + 1 < r$; hence this contradicts the assumption that R is strict.

□

The theorem holds and the proof can be applied without change if the number of parties is limited to K for $K \geq 4$. Four parties are necessary because of the reference to Theorem 3. If membership monotonicity is made a part of the premise, the case $K = 3$ can be included.

The construction of the priority relation R in the proof of Theorem 20 is implicit and indirect. Only the incomplete relation R* is constructed directly from F, and then R* is extended by methods which have nothing to do with F. When F is weakly balanced, it is possible to construct R directly. This provides an alternative proof of Theorem 21, and this proof is given below.

Alternative proof of Theorem 21

The alternative proof applies to the "only if" part of Theorem 21(a), which is the part of the proof which depends on Theorem 20. The purpose is to eliminate this reference.

Assume, therefore, that F is consistent and weakly balanced. Then F is membership monotone. The point is to prove that F is a priority method; otherwise, the earlier proof can be applied.

Define R by

$(x,r)R(y, s)$ if and only if there exist r' and s' with

$r' \geq r$, $s' < s$ and $F(r'+s'; x, y) \to (r', s')$.

For any x, y, r and s, there exist r' and s' with $F(r+s-1; x, y) \to (r', s')$. $r' \geq r$ implies $(x, r)R(y, s)$, while $r' < r$ implies $(y, s)R(x, r)$. This proves that R is complete.

R is also transitive, which can be proved as follows: Assume $(x, r)R(y, s)$ and $(y, s)R(z, t)$. By definition there exist r', s', s'' and t'' such that

$$F(r'+s'; x, y) \to (r', s'), \tag{15.30}$$

$$F(s''+t''; y, z) \to (s'', t''), \tag{15.31}$$

$$r' \geq r,$$

$$s'' \geq s > s', \text{ and}$$

$$t > t''.$$

Find r_1, s_1 and t_1 such that

$$F(r+s+t-1; x, y, z) \to (r_1, s_1, t_1).$$

Choose r_1 as large as possible for this to hold. If $r_1 < r$, then $s_1 + t_1 \geq s + t$. (15.31) and Lemma 1(b) can be used to change s_1 and t_1 such that either $s_1 \geq s''$ or $t_1 \leq t'' \leq t$. In both cases, $s_1 \geq s > s'$. But then (15.30) can be used to transfer at least one seat from the second to the first party, contradicting the choice of r_1. Hence $r_1 \geq r$. Then (15.31) can again be used to guarantee either $s_1 \geq s'' \geq s$ or $t_1 \leq t''$. Since $s_1 + t_1 \leq s + t - 1$, both these possibilities give $t_1 < t$. Consistency gives

$$F(r_1 + t_1;\ x,\ z) \rightarrow (r_1,\ t_1),$$

from which $(x, r)R(z, t)$ follows.

Hence R is a priority relation. Let F_0 be the priority method defined from R. Then F_0 is membership monotone. $F \subseteq F_0$ can be proved by an argument similar to the one used in the proof of Theorem 20; (15.19) follows from (15.15) and the definition of R.[165]

$F_0 \subseteq F$ remains. Since F is weakly balanced, $F(2r;\ x,\ x) \rightarrow (r,\ r)$ for all x and r. This gives $(x, r)R(x, r+1)$, and R is semi-strict. F_0 is consistent by the "if" part of Theorem 21(a), Lemma 1(d) applies, and $F_0 \subseteq F$ need only be proved for two-party situations.

The following implication will be proved below: Assume that x, y, r_1, r_2, s_1 and s_2 satisfy $r_1 > r_2$, $s_1 < s_2$,

(15.32) $$F(r_1+s_1;\ x,\ y) \rightarrow (r_1,\ s_1)$$

and

(15.33) $$F(r_2+s_2;\ x,\ y) \rightarrow (r_2,\ s_2).$$

Then

(15.34) $$F(r_1+s_1;\ x,\ y) \rightarrow (r_1-1,\ s_1+1).$$

For any x and y, $F_0(n;\ x,\ y) \subseteq F(n;\ x,\ y)$ will now be proved by induction on n. This is obvious for n = 0. Assume that it holds for situations with n - 1 seats, and assume

(15.35) $$F_0(n;\ x,\ y) \rightarrow (r,\ s).$$

There is no loss of generality in assuming that when this allotment is constructed, as described in Definition 21, the last seat is given to the second party. Hence

(15.36) $$F_0(n-1;\ x,\ y) \rightarrow (r,\ s - 1)$$

and $$(y,\ s)R(x,\ r+1).$$

By the definition of R, there exist r' and s' such that $r' \leq r$, $s' \geq s$ and

(15.37) $$F(r'+s';\ x,\ y) \rightarrow (r',\ s').$$

The induction hypothesis implies that (15.36) holds for F. Since F is membership monotone, either

(15.38) $$F(n;\ x,\ y) \to (r + 1,\ s - 1)$$

or

(15.39) $$F(n;\ x,\ y) \to (r,\ s).$$

If (15.38) is true, one can define $r_1 = r + 1$, $s_1 = s - 1$, $r_2 = r'$ and $s_2 = s'$. Then $r_1 > r_2$ and $s_1 < s_2$. (15.32) holds by (15.38), and (15.33) is true by (15.37). (15.34) is now the same as (15.39). The latter formula must therefore be true in any case. This completes the induction step, and $F_0 \subseteq F$ must hold.

The proof that (15.32) and (15.33) imply (15.34) remains. Assume that the two first formulas hold. By Lemma 1(b), there exist r_3 and s_3 such that

$$F(r_1+s_1;\ x,\ y) \to (r_3,\ s_3)$$

and either $r_3 \geq r_2$ and $s_3 \geq s_2$, or $r_3 \leq r_2$ and $s_3 \leq s_2$. In both cases, $r_3 < r_1$, and $s_3 > s_1$. Hence there exists a positive number m such that

(15.40) $$F(r_1+s_1;\ x,\ y) \to (r_1 - m,\ s_1 + m).$$

Among the positive integers m for which (15.40) holds, choose the smallest. Assume, in order to derive a contradiction, that $m \geq 2$.

Find r_4, s_4 and s_4' such that

(15.41) $$F(r_1+2s_1;\ x,\ y,\ y) \to (r_4,\ s_4,\ s_4').$$

If possible, choose r_4 to satisfy $r_1 - m \leq r_4 < r_1$. If this is not possible, choose r_4 to minimize $|r_1 - r_4|$. Let $|s_4 - s_4'| \leq 1$; this is possible since F is weakly balanced. Assume $r_4 < r_1 - m$. Then both s_4 and s_4' are greater than s_1, which gives $r_4 + s_4 < r_1 + s_1$. (15.32) and Lemma 1(b) can be applied to the first two parties. This leads to an increase in r_4 without bringing it above r_1, contradicting the original choice of r_4. Next assume $r_4 \geq r_1$. Then $s_4 \leq s_1$ and $s_4' \leq s_1$, and $r_4 + s_4 \geq r_1 + s_1$. (15.40) can be used to reduce r_4 without bringing it below $r_1 - m$; again contradicting the original choice. Hence $r_1 - m \leq r_4 < r_1$.

If $r_1 - m < r_4 < r_1$, at least one of s_4 and s_4' must lie strictly between s_1 and $s_1 + m$; there is no loss of generality in assuming $s_1 < s_4 < s_1 + m$. Now use Lemma 1(b) on (15.41) and (15.32), letting $I = \{1, 2\}$ and $\sigma(i) = i$. The conclusion can be written

$$F(r_1+s_1;\ x,\ y) \rightarrow (r_1 - m',\ s_1 + m'),$$

where either $r_1 - m' \leq r_4$ and $s_1 + m' \leq s_4$, or else these two inequalities are reversed. In both cases, $0 < m' < m$. This contradicts the original choice of m.

The only remaining possibility is $r_4 = r_1 - m$. Then $s_4 + s_4' = 2s_1 + m$, and $s_4 = s_1 + \lfloor \frac{m}{2} \rfloor$ can be assumed. In analogy with (15.41), choose r_5, r_5' and s_5 such that

$$F(2r_1+s_1-m;\ x,\ x,\ y) \rightarrow (r_5,\ r_5',\ s_5).$$

An argument entirely similar to the one applied to (15.41) can now be pursued. $s_1 + m$ and $r_1 - m$ play the roles of r_1 and s_1, respectively. It is possible to choose s_5 such that $s_1 \leq s_5 < s_1 + m$. If $s_1 < s_5 < s_1 + m$, a contradiction to the original choice of m follows. The possibility $s_5 = s_1$ remains; then $r_5 = r_1 - \lceil \frac{m}{2} \rceil$ can be assumed. Now $r_4 + s_4 = r_5 + s_5$. By consistency, r_5 and s_5 can be substituted for r_4 and s_4 in (15.41). Since $m \geq 2$, $r_1 - m < r_4 < r_1$ will then hold. But then an earlier argument applies, and again the choice of m is contradicted.

The assumption $m \geq 2$ has led to a contradiction. Hence $m = 1$. But then (15.40) is (15.34), and the proof is complete.

□

FOOTNOTES

1. The author is a doctoral candidate at John F. Kennedy School of Government, Harvard University, Cambridge, Massachusetts, USA. His work is supported by a fellowship from the Norwegian Research Council for Science and the Humanities.

2. Some of the results in this paper were presented at a seminar arranged by the Kennedy School and the Department of Economics, Harvard University, in March 1977.

3. This distribution is not the only aspect of an electoral system. In particular, individual representatives must be chosen from among each party's candidates, and the system must contain rules which determine this choice. All issues but the distribution of seats among the parties are ignored in this paper. Other aspects of electoral systems are discussed, to some extent, in [16].

4. This loose statement can be made precise in different ways; see, for example, Theorem 9 in Section 9 below.

5. In particular, I will mention the recent contributions of M. L. Balinski and H. P. Young. Before starting working on this paper, I was aware of their article [1]. Later, I have had the opportunity to study other works by Balinski and Young, in which many of the issues raised in this paper are discussed. Specific references are made, and their approach and results are compared to mine, in subsequent notes.

6. See, for example, notes 96 and 161.

7. If the election is conducted and the result determined separately in several districts, the entire discussion here applies to any one district. Some problems related to aggregation of results over many districts are discussed in [16].

8. The concept is used in social choice theory. A procedure is said to be anonymous if interchanging two persons' ballots never changes the outcome. It then follows that no change in outcome will result if the ballots are permuted in an arbitrary way among the individuals. See, for example, [20] page 72 or [11] page 183.

9. To be precise, for any k and n, the set of vectors $\overline{x}$ for which $F(n; \overline{x})$ has more than one element, is of dimension k-1.

10. In particular, this example shows that all consistent methods must contain ties, see Definition 8 in Section 5 below.

11. In connection with the treatment of ties, there is a difference between the approach of this paper and the one of Balinski and Young; see, for example, [2]. While an allotment method in this paper is a set-valued function, a method in the sense of [2] is a set of single-valued functions, defined on the same domain. The difference has the following consequence: Let a "method," in some intuitive sense, be given. Consider two situations in

which the allotment of seats to parties is not unique; assume, for example, that there are two possible allocations in each situation. Then there are four "combined outcomes." An allotment method will give these four possibilities equal status. A method in the sense of [2] will distinguish between combined outcomes that can be obtained by the use of one function from the method, and outcomes that can only occur if different functions are used in the two situations. There will exist a method which contains functions corresponding to all four possibilities, and methods which are proper subsets of this one. The former and any of the latter will formally be different methods. But the "full" method and some of its proper subsets will have the property that they, for every situation, allow exactly the same allocations of seats.

The "standard" methods studied by Balinski and Young will be uncountable sets of functions, although the set of situations is countable and only a finite number of allocations is possible in each situation. Each standard method has an infinity of subsets which allow the same allotments everywhere but are formally different methods. This richness of structure does not seem to add anything.

On the whole, the difference between the two approaches is not very important, but it does have an impact on some of the definitions given below.

If M is a method in the sense of [2], there is a natural way of defining a corresponding allotment method, namely by $F(n; x_1, \ldots, x_k) = \{f(n; x_1, \ldots, x_k) | f \in M\}$. It is not in general obvious how one shall go the other way. One can set $M = \{f | f(n; x_1, \ldots, x_k) \in F(n; x_1, \ldots, x_k)$ for all n and $x_1, \ldots, x_k\}$, but this may destroy desirable properties, such as membership monotonicity (see Definition 9 in Section 5). This issue is discussed further in note 36.

12. It is hard to imagine how a party can exist and participate in an election without getting any votes. It is equally difficult to believe that the restriction $x_i > 0$ will cause any problems in other contexts where allotment methods can be used, such as the apportionment of representatives among geographical districts.

13. For an example, see Theorem 5(b) in Section 8.

14. This may be difficult in connection with proportional representation, but possible in other applications.

15. Note that this restriction normally makes a condition weaker, since it need only hold on a more limited set of situations.

16. In fact, almost all the results can be proved without assuming the existence of anything but integer votes. The proofs are formulated so as to make this clear. Exception can be found in Sections 13 and 14; see Theorems 16, 18(c) and 19(b), and notes 131, 141 and 144.

17. Usually, it would not make any difference if this generalization were made and irrational votes allowed; most results would still hold. (Exceptions are pointed out in notes 130 and 155, see also note 120.) Even when the votes are assumed to be rational (or integers), certain other quantities which occur in description of methods and in proofs must be allowed to take on irrational values; see notes 52 and 96.

18. It is sufficient to assume that (4.1) holds for integer values of a. For if $a = \frac{p}{q}$, where p and q are positive integers, then the weaker form of (4.1) gives $F(n; x_1, \ldots, x_k) = F(n; px_1, \ldots, px_k)$ and $F(n; ax_1, \ldots, ax_k) = F(n; qax_1, \ldots, qax_k)$. Since $qax_i = px_i$, this implies (4.1) for the given a, which was an arbitrary positive rational number.

    If one allows irrational votes, the a of Definition 2 must be allowed to take on any positive real value.

19. The measures are closely related to what is known as the ℓ^p-norm (and the ℓ^∞-norm). F_{ℓ^∞} can, in effect, be obtained as the limit of F_{ℓ^p} as p tends to infinity. (To be precise, it is the "lexicographic version" of F_{ℓ^∞} which is equal to this limit, see note 23 and accompanying text.)

20. Equivalently, one can say that F_{ℓ^2} minimizes the length of the "error vector" $(\hat{x}_1 - r_1, \ldots, \hat{x}_k - r_k)$.

21. In [2], this method is referred to as the Hamilton method. The name used in this paper is also used elsewhere; see, for example, [18] and [17]. The reason for choosing this name is that the last seats are given to the parties which have the largest "remainders," measured by $\hat{x}_i - \lfloor \hat{x}_i \rfloor$, when each party has been awarded one seat for each time $V_{\overline{x}}$ divides its vote (which is $\lfloor \hat{x}_i \rfloor$ times for party i).

    The word "complete" is used, whenever necessary, to distinguish F_{LR} from any of its submethods.

22. This result is by no means new; in whole or in part it can be found in [12], [1] and [8].

23. To be precise, let $\overline{r}$ and $\overline{s}$ be two possible allocations for a situation $(n; \overline{x})$, let $a_1, \ldots, a_k$ be the numbers $|\hat{x}_1 - r_1|, \ldots, |\hat{x}_k - r_k|$ in decreasing order, and define $b_1, \ldots, b_k$ similarly from $|\hat{x}_1 - s_1|$, $\ldots, |\hat{x}_k - s_k|$. $\overline{r}$ is "better than" $\overline{s}$ if there is an i such that $a_i < b_i$ and $a_j = b_j$ for all $j = 1, \ldots, i - 1$; $\overline{r}$ and $\overline{s}$ are "equally good" if $a_i = b_i$ for $i = 1, \ldots, k$. The lexicographic F_{ℓ^∞} chooses the allocations in $T_{k,n}$ which are "best" in this sense.

24. The equivalence of this description and Definition 4 is easy to prove. The deductions may bring a party's vote below 0, but no seat will be awarded on the basis of a non-positive vote.

25. Several quota methods are discussed in the literature, but I have not seen the general concept formulated elsewhere. (I have used it previously in [14].) The concept is not related to the "quota method" of [1].

26. It would not make much difference if the quota also were allowed to depend on the number of parties.

27. For example, the methods F_{HA}, F_{MF} and F_{SD} (see Definition 14) can be described as "quota methods" in this broader sense; see Section 8.4.

28. It is known as Droop's method or Hagenbach-Bischof's method. If the number of votes is large compared to the number of seats, one can define $V(n, y) = \left\lfloor \frac{y}{n+1} \right\rfloor + 1$, which corresponds to the standard way of describing the method. If the number of votes is small, however, this will not satisfy (4.3) and can cause problems. (This was pointed out in [12].) Hence "slightly greater" must be defined in some other way.

29. In [1] and [5], these conditions are called "lower quota" and "upper quota"; a method is said to "satisfy quota" if both are fulfilled.

30. Formally, the weakened lower bound condition is $r_i \geq \hat{x}_i - 1$, and the weakened upper bound condition is $r_i \leq \hat{x}_i + 1$. There are three theorems below where any of these conditions appear in the premise, namely Theorems 5(b), 6(b) and 9. In each case, the weakened conditions are sufficient, see notes 60, 65 and 86. When the conditions occur in the conclusion of a theorem, the strong version is always established.

31. When there are two parties, the "extended" condition is no stronger than Definition 6(a). For three parties, the extended lower bound condition is equivalent to the conjunction of the lower and the upper bound condition.

32. The situation (2; 1, 1, 1, 1), with four parties, will also show the inconsistency of the condition. As above (see text accompanying note 30), the condition can be weakened slightly by permitting $\sum_{i \in I} r_i = \sum_{i \in I} \hat{x}_i - 1$. This condition is satisfied by F_{LR} when the number of parties is limited to 4, but it is inconsistent for larger number of parties by the example in the text.

33. This is not the same as the condition "binary fairness" of [8]. The latter condition is equivalent to the definition of $F_{\ell 1}$; see Definition 3.

Hence, by Theorem 1, binary fairness characterizes the method of the largest remainder. (This is also pointed out in [3].) The condition "binary consistency" of [8] is, however, related to pairwise fairness; see note 66 and Theorems 7 and 8.

34. The definition is inspired by [1]. The consistency condition of [1] is, however, considerably weaker than the condition of this paper. In [2] and [5], consistency is defined in a way which is close to the definition here. The definition of [2] and [5] applies only to membership monotone methods, as is made clear in [5]. It is also a little weaker than consistency in this paper. The premise of Definition 8 says that two parties have the same votes and the same total representation in two situations; no reference is made to the order in which the seats were awarded (nor would such a reference make sense). In [5], the corresponding premise includes the condition that the last seat, in each of the two situations, were given to one of the two designated parties. A stronger premise, of course, makes a weaker condition. See further discussion in notes 36 and 158.

35. In [2] and [5], the corresponding condition is called monotonicity or house monotonicity.

36. When F is membership monotone, there will in some cases be a natural link between allotments for different situations. (This is not the case for an arbitrary allotment method, see the discussion in note 11.) In particular, if $F(n; \overline{x}) \to \overline{r}$, then $F(n+1; \overline{x})$ will contain some allotments $\overline{s}$ for which $s_i \geq r_i$ for all i, and perhaps some for which this is not true.

More can now be said about the connection between allotment methods and methods as defined by Balinski and Young. If F is a membership monotone allotment method, a method M in the sense of [2] is given by $M = \{f \mid f$ is house monotone and $f(n; \overline{x}) \in F(n; \overline{x})$ for all n and $\overline{x}\}$. Then M will be non-empty and house monotone, and if an allotment method F' is constructed from M as described in note 11, F' and F will be equal. If F is consistent, so is M (consistency taken in the appropriate sense each time). But M can be consistent without F being (an example of this is given in note 158). Therefore, consistency as defined in this paper is, in a sense, a stronger condition than consistency as defined in [2] and [5]; see also note 34. In this paper, membership monotonicity will follow from consistency and another reasonable condition (see Theorem 3 below), and in most results where consistency is a part of the premise, it is not necessary to make membership monotonicity a separate assumption. In the works of Balinski and Young, both consistency and house monotonicity are generally assumed. This difference is a consequence of the different definitions and is not very important.

If the order in which F and M are constructed from each other is reversed, the correspondence is a little more complicated. Let M be a house monotone method in the sense of [2], and let F be obtained from M as in note 11. Then F is a membership monotone allotment method. But if M' is constructed from F as above, it can only be concluded that $M \subseteq M'$. (Consistency of M is not enough to guarantee M' = M.) Any of M and F can be consistent without the other being. This lack of symmetry is caused by the "richness" of the system of possible methods M; many different methods M collapse into the same F by the construction of note 11, and only certain "complete" methods can be obtained from an F by the construction described in this note.

37. To be pedantic, one first uses the fact that $F(r_i+r_j; x_i, x_j)$ is non-empty to find a pair (r_i', r_j') in this set. Since $r_i' + r_j' = r_i + r_j$, consistency can be applied to the situations $(n_1; \overline{x})$ and $(r_i+r_j; x_i, x_j)$, to get $F(r_i+r_j; x_i, x_j) \to (r_i, r_j)$.

38. In [2], this condition is called symmetry. The word neutrality is often used in social choice theory about a related condition; see, for example, [20] page 72 or [11] page 183. "Neutrality" means neutrality in relation to parties. Neutrality in relation to voters is called anonymity and is embodied in the definition of allotment methods; see note 8 and accompanying text.

39. The problem is that if $F(n; x_1, x_2, x_3) \to (r_1, r_2, r_3)$ and $F(n; x_2, x_1, x_3) \to (s_2, s_1, s_3)$, it can conceivably be the case that $r_i + r_j \neq s_i + s_j$ for any two parties i and j. Then consistency cannot immediately be applied.

When the number of parties is limited to two, consistency and neutrality are equivalent conditions. A consistent method which is not neutral cannot be weakly balanced, see Definition 11 and Theorem 3 below. The generalization of consistency to more than two parties, see the end of Section 5.1, immediately implies neutrality.

40. The term "balanced" is taken from [2].

41. Example: If $x > y$ and $n \leq 3$, $F(n; x, y) = (n, 0)$ and $F(n; y, x) = (0, n)$. Otherwise, $F(n; x, y) = F_{\overline{LR}}(n; x, y)$. The theorem breaks down because consistency is a rather weak condition when the number of parties is limited to two (see note 39 above). I do not know whether the theorem is correct when the number of parties is limited to three, but I would guess that the answer is yes.

42. See note 9 and accompanying text.

43. The "quota method" of [1] is internally but not externally vote monotone. (It is also membership monotone.) This method is described in Section 10 below. The method $F(n; \overline{x}) = (n, 0, \ldots, 0)$, which gives all seats to party 1 regardless of the votes, is externally but not internally vote monotone. It is also possible to find neutral methods which have these properties, but my examples are fairly complicated (and not very interesting).

44. This again has to do with the weakness of the consistency condition when the number of parties is limited to two. The following method, defined only for situations with two parties, is consistent, membership monotone, and internally but not externally vote monotone: If $x > y \geq x - 1$, then $F(n; x, y) = (n, 0)$ and $F(n; y, x) = (0, n)$. Otherwise, $F(n; x, y) = F_{LR}(n; x, y)$. (F is strange, but it is not entirely inconceivable that it captures the spirit of actually existing political phenomena: If the vote is a close call, the winner will "overcompensate" and take all the seats; if the winning party has a more comfortable lead in the votes, it can afford to share power and seats with the loser.)

45. The idea of defining this class of methods emerges from the studying of special divisor methods such as F_{HA} and F_{MF} (defined and discussed in Section 8 below). The general concept has earlier been defined in [14], [1] and [2]. $d(\alpha)$ in [1] and [2] corresponds to $d_{\alpha+1}$ in this paper. The definition used in [1] is somewhat different from the one used in the other papers, but the equivalence of the two formulations can easily be proved.

In [13], divisor methods are studied, but the general concept is not formulated explicitly.

The word "complete" is used whenever necessary to distinguish divisor methods from partial divisor methods, defined in (b) below.

46. Note that a partial divisor method is not a partial function in the sense that it may be undefined for some situations; a partial divisor method F is an allotment method, and $F(n; \bar{x})$ must be defined and non-empty for all situations $(n; \bar{x})$.

47. In practice, it will often be unnecessary to compute all these numbers. If it is clear from the outset that party i will get at least r_i^- and at most r_i^+ seats, only $\frac{x_i}{d_{r_i^-}}, \ldots, \frac{x_i}{d_{r_i^+}}$ need actually be computed.

48. A convenient way of organizing the computation is to arrange the quotients in an $n \times k$ matrix, the entry in column i and row α being $\frac{x_i}{d_\alpha}$. Then one checks off the n largest quotients, and counts the number of checks in each column.

49. For example, take any divisor method and break all ties in favor of the first party or parties; "first" referring to the ordering of the parties' votes in the vector $\bar{x} = (x_1, \ldots, x_k)$. That is, if the quotients $\frac{x_i}{d_\alpha}$, on the basis of which the parties compete for their next seat, are equal for two or more parties, then the seat shall be given to the first of the parties which have this largest quotient. This is a membership

monotone and scale independent partial divisor method which is not consistent. Neither is it neutral. For examples of neutral methods which satisfy all but one of the three conditions mentioned above, see Section 11.2.

50. For later reference, an example is given: Consider the divisor method given by $d_\alpha = \alpha$ for all α. (This is the method of the highest average (F_{HA}), discussed in Section 8 below.) Whenever there is a tie, break it in favor of the largest party. (If the parties have an equal number of votes, break the tie arbitrarily.) Hence the unique allotment in the situation (2; 2, 1) is (2, 0), while $F_{HA}(2; 2, 1) = \{(2, 0), (1, 1)\}$. This method is actually in use, in Norwegian municipal elections; see "lov om kommunestyrevalg og fylkestingsvalg" (local elections law) from 10 July 1925 no. 6, section 34. (The law gives an algorithm which is somewhat different from the one described above, but the equivalence is easily proved.) For electorates as large as in actual elections, ties are highly unlikely. The difference between this method and F_{HA} is therefore mainly theoretical.

51. Again, the statements need not be true for partial divisor methods. One part is true, however: If F is a submethod of a strict divisor method, then F is strongly balanced.

52. The example involves only two parties, the votes are integers, and the divisors d_α and d_α' are allowed to be irrational. In particular, it follows that the full strength of the class of divisor methods can only be obtained by allowing irrational divisors, although the votes are supposed to be rational; the same would be true if only integer votes were allowed. See notes 82 and 84 for a method with irrational divisors which is in actual use.

53. The method has many names. It often called the d'Hondt method, after Victor d'Hondt, a Belgian who proposed the method in articles published approximately 100 years ago. In [2], it is called the Jefferson method. The phrase highest (or largest) average is used in [18] and [17] (in the former, however, this name refers to a broader class of methods). The reason for the name is the following: If party i has already been awarded r_i seats, it will compete for its next on the basis of the number $\frac{x_i}{d_{r_i+1}} = \frac{x_i}{r_i+1}$. (See the discussion of computational aspects, after Definition 13 in Section 7.) This number is the average number of votes behind each of the party's seats if it gets the one for which it presently is competing. Thus the seat is given to the party for which this average is highest. Yet another name is "the method of the greatest divisor" (used in [13]); this name is explained below, see note 74.

54. This method also has several names: the odd number method, because the sequence of divisors equivalently could have been given as 1, 3, 5, ...; the Sainte-Laguë method, after a Belgian who proposed the method early in this century (this name is used in [18] and [17]); and the Webster method, used in [2]. The name "the method of major fractions," also used in [13], is explained below; see note 71 and accompanying text.

55. Again an explanation of the name is given below; see text accompanying note 75.

56. The method is intended to guarantee each party at least one seat, and it does not make much sense when $n < k$. Therefore, some ad hoc rule must be applied in this case. The one chosen in Definition 14(c) is designed to make possible the characterization in Theorem 6 below. As defined here, F_{SD} is not vote monotone. Another possibility is to give the n seats to the largest parties when $n < k$; see note 63.

57. Essentially the same result is proved in [6]. In that paper, membership monotonicity is included in the premise; see the discussion in note 36.

58. The non-trivial part of the proof, namely the proof of part (b), is based on a technique used in [1].

59. Indeed, the sequence can be chosen in uncountably many different ways.

60. The proof will work even if one only assumes the weakened versions of the lower bound condition mentioned in comments after Definition 6; see text accompanying note 30. When this weakened condition holds, it can be proved that F is weakly balanced, which is enough for the proof in the text to proceed. (Later in the proof, only the weakened condition is needed.) The proof that F is weakly balanced goes like this: By consistency, it is sufficient to prove that Definition 11(b) holds for situations with two parties. Assume that $F(2r; x, x) \to (r+1, r-1)$ for some x and r, which is the most "unbalanced" allotment permitted by the weakened lower bound condition. Consider the situation $(3r+1; x, x, x)$. Here the exact representation of each party is $r + \frac{1}{3}$. Apart from permutations, the only possible allotment is $(r+1, r, r)$. Consistency gives $F(2r; x, x) \to (r, r)$, which is exactly what is needed to conclude that F is weakly balanced.

61. See note 9 and accompanying text.

62. A similar result can be found in [6].

63. If F_{SD} is redefined for situations $(n; x_1, \ldots, x_k)$ with $n < k$ so as to require that the n seats be given to the n largest parties, the method and all its submethods become internally and externally vote monotone. (See note 56.) The theorem holds for the redefined method if either internal or external vote monotonicity is added to the premise of (b). In (a) and (c), both these conditions can be added to the conclusion.

64. Here it is essential that all votes are strictly positive.

65. As in Theorem 5(b), it is sufficient to assume the weakened form of the upper bound condition; see note 30. By an argument similar to the one used in note 60, it can be proved that F is weakly balanced. (If $F(2r; x, x) \to (r+1, r-1)$, the situation $(3r-1; x, x, x)$ is considered.)

66. The concept originates in [8], where it is called "binary consistency." (To be precise, binary consistency, as formulated in [8], is not contradicted by certain strongly biased allotments which do contradict relative well-roundedness. But the concept is presumably intended to be equivalent to Definition 15.) The name "relative well-roundedness," as well as the formulation of Definition 15, is taken from [3].

67. This description is strictly correct only when $\hat{x}_i + 1 > r_i > \hat{x}_i + \frac{1}{2}$. But if $r_i \geq \hat{x}_i + 1$, the over-rounding is even more striking.

68. In [3], membership monotonicity is a part of the premise, see discussion in note 36 above.

69. For later reference, note that the contradiction is obtained even if only one of (8.13) and (8.14) is a strict inequality.

70. Fractions which are equal to one half have been ignored so far. To obtain an allotment with the correct total number of seats, it may be necessary to round some such fractions upwards and others downwards.

71. This explains the name "the method of major fractions."

72. Ostensibly, the range permitted by (8.15) depends not only on n and $\overline{x}$, but also on $\overline{r}$. It can be shown, however, that the range does not depend on the particular $\overline{r}$ chosen from $F_{MF}(n; \overline{x})$, if this set has more than one element.

73. The reasoning may perhaps seem circular, since V_{MF} originally is given by (8.15), which in turn depends on F_{MF}. But in fact the right-hand side of (8.17) makes no reference to all of this and does not depend on the earlier definition of F_{MF}; hence it provides an independent definition of the method. The proof that (8.17) defines the same methods as the one given by Definition 14(b) is given informally above and can easily be formalized.

74. The method is often defined this way; see, for example, [9] and [1]. It also explains the name "the method of the greatest divisor."

75. This explains the name. In [1] and [8], the method is defined in this way; in the latter paper it is called "σ-quota."

76. This is strictly true only if ties are ignored. To be precise, the "limit method" is a submethod of F_{HA}. It breaks ties in favor of small parties, in contrast to the submethod described in note 50.

77. Again, ties must be ignored.

78. The limiting cases a = 1 and a = -1 correspond to F_{HA} and F_{SD}, respectively; see the previous paragraph.

79. All methods considered here are scale independent; hence the votes can be normalized to sum to 1. For any k and n, the set of situations then becomes a bounded set. The statement in the text can be made precise in terms of a measure on that set.

80. One gets the "divisor method" defined by $d_\alpha = \frac{2\alpha(\alpha-1)}{2\alpha-1}$. The problems caused by the fact that $d_1 = 0$ are solved as in Definition 14(c).

81. This criterion gives the method of major fractions.

82. One gets the "divisor method" given by $d_\alpha = \sqrt{\alpha(\alpha-1)}$; $d_1 = 0$ is again taken care of as in Definition 14(c). The divisors are irrational numbers and cannot all be made rational by rescaling; see note 52. Also, the method is in actual use; see note 84.

83. Some of the criteria do not uniquely define a method, while others lead to methods which are not membership monotone. From the remaining criteria five different methods emerge, namly the ones described in Definition 14 and notes 80 and 82. A systematic presentation of the five methods and criteria defining them, can be found in [2].

84. The topic discussed in [13] is not distribution of seats to parties in proportional elections, but apportionment of the House of Representatives of the USA. (The U.S. Constitution prescribes that the representatives be apportioned among the states "according to their respective numbers," that is, in proportion to their population.) In my opinion, the remarks in the text apply in this case as well. Professor Huntington, however, argues that certain criteria are clearly superior. These determine a unique "method of equal proportions," namely the one described in note 82. This method is presently used for apportionment of the House of Representatives.

85. This result is also given in [2], where its proof depends on a theorem in [1], the proof of which uses an unlimited number of parties.

86. See the proof of Theorem 5(b). The lower and the upper bound condition can be weakened as described in note 30 without affecting the result; the argument of note 60 can be used to show that F is weakly balanced, and the rest of the proof applies without change.

87. The difference in basic framework, see notes 11 and 36, is unimportant here. Methods satisfying (i) - (iii) are called "quotatone" in [4].

88. To be precise, any solution must be a submethod of the allotment method $F_{\bar{Q}}$ defined from $\underset{\sim}{\bar{Q}}$ as described in note 11.

89. F_{SD} itself, as given by Definition 14(c), is not vote monotone, but the submethod described in note 63 satisfies (ii) - (v).

90. This is not really a proof, but it gives the idea. A formal proof is given in [1], where it is also proved that F_Q essentially is the unique method satisfying (i) - (iii) and a weakened form of consistency; uniqueness is here interpreted as in Theorem 5(b).

91. The use of four parties is essential; the method is externally vote monotone when the number of parties is limited to three.

92. The size of the majority required to take all the seats, will of course depend on the total number of seats. This condition is not necessarily a reasonable one in other connections where allotment methods can be used.

93. By (i), (ii) and Theorem 4, it is unnecessary to distinguish between external and internal vote monotonicity.

94. In the presence of consistency, vote monotonicity and scale independence, this says exactly what was announced above: No matter how many seats there are, if a party has a strong enough majority, it takes them all.

95. A number u is an upper bound for a set U if $u \geq v$ for all $v \in U$.

96. The number u is the supremum of U if u is an upper bound for U and no number $v < u$ is an upper bound for U. If a supremum exists, it clearly is unique. It may seem obvious that any set of real numbers which is bounded from above, has a supremum. But in fact this is a nontrivial property of the real number system. For example, the rational number system does not have this property. It follows that although the set D_α contains rational numbers only, d_α need not be rational; see note 52 above. The existence of suprema is discussed in any basic textbook in real analysis; see, for example, [19], Chapter 1.

97. The number a is included, and a and y are required to be integers, in order for the proof to work even if only integer votes are allowed.

98. This again has to do with the lack of strength of the consistency condition when there are only two parties. An example of a method which is not a complete or partial divisor method but satisfies (i) - (v) when the number of parties is limited to two, can be constructed as

follows: Let $d_\alpha^{(n)} = \alpha - \frac{1}{n+1}$, and let F_n be the divisor method with divisors $d_1^{(n)}$, $d_2^{(n)}$, ... Then define $F(n; x, y) = F_n(n; x, y)$. F satisfies (i) - (v). (The only condition which could possibly cause problems is membership monotonicity, but the proof that this condition holds is not difficult.) It is also easily seen that F cannot be a submethod of any divisor method, even for the restricted case where there are only two parties. F cannot consistently be extended to situations with three or more parties.

99. Again the method of note 50 can serve as an example; it satisfies (i) - (v) but is not a complete divisor method.

100. As noted in the comments to Definition 8, it is sufficient to verify consistency when one of the two situations involved has only two parties. The definition of F_{ii} implies that at most one party gets an odd number of seats. Now one can consider the four cases given by $k_1 = 2$ or $k_2 = 2$, $r_{i_1} + r_{j_1}$ is even or odd, where the symbols are used as in Definition 8. The cases in which $k_2 = 2$ are fairly straightforward. Then consider the cases with $k_1 = 2$ and $k_2 > 2$. If $r_{i_1} + r_{j_1} = s_{i_2} + s_{j_2}$ is odd, y_{i_2} or y_{j_2} must be the largest (or one of the largest) in the vector $(y_1, \ldots, y_{k_2})$. If this number is even and n_2 is odd, there must exist $i \notin \{i_2, j_2\}$ with $y_i \geq y_{i_2}$ and $y_i \geq y_{j_2}$. (If n_2 is even, there is no problem.) In all the cases, the conclusion of Definition 8 can be derived.

101. Even if one only accepts integer votes, $F_{HA}(n; \bar{y})$, as a formal function, is defined for all positive numbers $y_1, \ldots, y_k$. Therefore, F_{iii} is well defined in this case as well.

102. This description is not completely precise, but it should be clear what "breaking ties" means; see also note 49.

103. That is, the $\lfloor \frac{n}{2} \rfloor$ seats are allotted according to the method of note 50. If there is a tie between parties with equal votes, it can be broken any way.

104. F_2 never gives an odd number of seats to more than one party. The proof of consistency can be divided into cases in the same way as was done for the method F_{ii} in note 100. In each case it can be shown, in the terminology of Definition 8, that $r_{i_1} \neq s_{i_2}$ implies $x_{i_1} = x_{j_1}$. Consistency then follows.

105. Such a minimum requirement is probably not a reasonable condition in proportional elections, but it may be imposed in other connections when allotment methods can be used. (Compare note 92 and accompanying text.) Minimum requirements are discussed more thoroughly in [1] and [4], where the possibility of the requirement being different for different parties is also considered.

106. No attempt will be made to explain what "utility" is or how it can be measured. One can think of it as some measure of pleasure or contentment, but more tangible interpretations are also conceivable. In general, the realism of the approach of this section is open to criticism in a number of ways. This issue is not discussed here, but some possible objections are mentioned in [16].

107. This condition corresponds to "diminishing marginal utility" or "concave utility functions," concepts often used in welfare economics. If (12.4) does not hold, the method defined by (12.2) need not be membership monotone.

108. See the second description in the comments to Definition 13.

109. If F is internally vote monotone, $F(1; 2, 2, 3) = (0, 0, 1)$ and $F(1; 4, 3) = (1, 0)$, contrary to the proposed condition.

110. In a sense, this definition assumes that only neutral allotment methods are considered. If F is not neutral, part (a) will, for example, not really say that F encourages merger, but only that F encourages a merger of parties 1 and 2. In the theorems below, neutrality will follow from other conditions and need not be assumed.

111. In [9], concepts similar to parts (a) and (d) were introduced; see quote and discussion below. In [2], the term "stability" refers to a property which for all practical purposes is equivalent to the conjunction of (b) and (d), the phrase "encourage coalitions" is used of (a), and "encourage schisms" denotes (c).

112. Formally, this condition would read:

$$F(n; x_1, x_2, \ldots, x_k) \to (r_1, r_2, \ldots, r_k) \text{ and}$$
$$F(n; x_1 + x_2, x_3, \ldots, x_k) \to (s_1, s_3, \ldots, s_k)$$
$$\text{imply } s_1 \geq r_1 + r_2.$$

113. Example: $F_{HA}(2; 1, 1, 1) \to (1, 1, 0)$ and $F_{HA}(2; 2, 1) \to (1, 1)$, contradicting the condition of note 112.

114. In the strengthened version of note 112, merger and division are treated symmetrically.

The use of the phrase "encourage schisms" in [2] does not mean that division is considered; the concept is equivalent to Definition 17(c).

115. If the "division conditions" are used instead of the conditions of Definition 17, Theorems 13 and 14 still hold. (The proof of Theorem 13 applies unchanged, and the proof of Theorem 14 can easily be modified.) In Theorems 15 and 16, certain technical problems arise; they are not discussed in detail here.

116. Examples: $F_{LR}(2; 7, 7, 6) = (1, 1, 0)$ while $F_{LR}(2; 14, 6) = (1, 1)$; and $F_{LR}(1; 3, 3, 4) = (0, 0, 1)$ while $F_{LR}(1; 6, 4) = (1, 0)$.

117. This result is also proved in [2].

118. Moreover, the proof works even if the conditions of Definition 17(b) and (d) are strengthened as indicated in note 112.

119. Parts of the theorem can be found in [14] and [2].

120. Strict inequality will hold between $d_{\alpha+\beta}$ and $d_\alpha + d_\beta$ if and only if F encourages merger in the strengthened sense of note 112. (If some of the numbers d_α are irrational, irrational votes are required to prove that strict inequality is a necessary condition for the strengthened version of Definition 17(a).) Similar remarks apply to parts (b) - (d).

121. In order to make the proof formally correct, the possibility of $r_i = 0$ or $d_{r_i} = 0$ for $i = 1$ or 2 should be treated separately. It is easily seen that the entire proof can be carried through in these cases.

122. It is essential that this proof makes use of four parties, while only three parties were involved in the corresponding proof of part (a).

123. $d_1 = 0$ will imply $d_\alpha = 0$ for all α. The "divisor method" defined from these divisors is F_T^α.

124. Translated from Danish by this author.

125. Here follows a description of the method of the highest average, similar to the one given in Section 8.4; see note 74 and corresponding text.

126. The unstated conditions are slightly stronger than consistency, but follow from consistency and membership monotonicity. (I interpret Erlang's first two conditions as correspoding to external vote monotonicity.) In [15], it is proved that Erlang's claim, and not only the proof, is erroneous. The allotment method used to show this can be described as follows: Let a situation $(n; x_1, \ldots, x_k)$ be given. If $n \geq 2$ and there is a party whose vote exceeds a fraction $\frac{n}{n+2}$ of the total vote, that party shall get all the seats. Otherwise, the

method of the highest average shall be used. This method obviously satisfies I and II, and III is not difficult to prove. In the situation (2; 3, 2) the method gives both seats to the first party, while $F_{HA}(2; 3, 2) = (1, 1)$. The method is not consistent, but it is neutral and satisfies several other of the conditions previously discussed. In particular, conditions (ii) - (v) of Theorem 10 hold.

127. Similar results are given in [2]. The results of that paper may seem stronger than Theorem 15, since they do not contain any condition corresponding to restriction of gain (or loss) by merger. But such a condition (called stability, see note 111) is used in the proof and is necessary for the result. (The divisor method given by $d_\alpha = \alpha + 1$ is different from F_{HA} but satisfies the rest of the conditions.)

128. This also means that Theorem 15 is not, strictly speaking, a generalization of the characterizations at the end of Section 13.3. The class of divisor method and the class of consistent and strongly balanced method are incomparable; neither is a subset of the other. Theorem 16, however, is a generalization of the result from Section 13.3.

129. A shorter proof can be given by using Theorem 21 in Section 15 below. (This corresponds to the proof given in [2].) Since F is consistent and strongly balanced, it is a strict priority method. Let R be the corresponding priority relation. Lemma 2(a) and (d), and the fact that R is strict, can be used to prove (x, r)R(nx, nr) and (nx, nr)R(x, r) for all x and r and all positive integers n. Then it can be proved that R is exactly the priority relation which defines F_{HA}.

Similar remarks apply to the proof of Theorem 15(b). Here Lemma 2(b) and (c) imply that (x, r) and (nx, nr - n + 1) have equal priority.

130. The method used at the corresponding point in the proof of part (a), which essentially is induction on $r_1 + r_2$, cannot be applied here. The reason is the possibility that $r_1 = 1$, $r_2 = 0$ and $x_1 \neq x_2$. Note that the proof in no way depends on non-integer votes being allowed; if both x_1 and x_2 are integer, one can choose a = 1. The proof does not work, however, if irrational votes are allowed.

131. Here it seems essential that non-integer votes are allowed. If it is assumed that F is scale independent, only integer votes need be used.

132. For example, a standard of comparison can be the "exact representation," $\hat{x}_i$, defined in Section 4.1. It is meaningful to say that a method favors large (or small) parties compared to this criterion.

133. A similar condition is formulated in [1].

134. For example, Definition 18 is not satisfied if F = G is a complete divisor method. But a method which only leads to ties when two parties have an equal number of votes, such as the method described in note 50, will favor small parties compared to itself. These phenomena are consequences of the specifics of the definition and should not cause any problems.

135. It is not clear that transitivity holds for methods which are not consistent. If $F(n; \overline{x}) \rightarrow \overline{r}$, $F''(n; \overline{x}) \rightarrow \overline{t}$ and $r_i + r_j = t_i + t_j$, there will not necessarily exist any $\overline{s}$ with $F'(n; \overline{x}) \rightarrow \overline{s}$ and $s_i + s_j = r_i + r_j$; hence the assumptions about F' cannot be used. If the methods are consistent, everything can be reduced to two-party situations, in which case such an $\overline{s}$ will always exist.

136. Formally, the alternative condition is the following: $F(n_1; x_1, \ldots, x_{k_1}) \rightarrow (r_1, \ldots, r_{k_1})$, $G(n_2; y_1, \ldots, y_{k_2}) \rightarrow (s_1, \ldots, s_{k_2})$, $x_{i_1} = y_{i_2} < x_{j_1} = y_{j_2}$ and $r_{i_1} + r_{j_1} = s_{i_2} + s_{j_2}$ imply $r_{i_1} \geq s_{i_2}$. The definition in [1] is formulated in this way (at least it allows $n_1 \neq n_2$).

137. This part is also proved in [1], using a different but equivalent definition of divisor methods.

138. See note 136 and accompanying text. F and F' of Theorem 17 may be partial divisor methods which are not consistent, hence the two versions of the definition need not be equivalent.

139. If $\frac{d_\alpha}{d_\beta} = \frac{d_\alpha'}{d_\beta'}$ for some $\alpha > \beta$, one can find x_1 and x_2 with $\frac{x_1}{x_2} = \frac{d_\alpha}{d_\beta}$. (Here rationality of the divisors is essential, unless irrational votes are allowed.) Since F is strict, $d_\alpha > d_\beta$ and $x_1 > x_2$. In the situation $(\alpha+\beta-1; x_1, x_2)$, both F and F' produce a tie between $(\alpha-1, \beta)$ and $(\alpha, \beta-1)$. This contradicts Definition 18.

140. One can ask whether it is possible to strengthen the condition of Definition 19 in the same way as was attempted by the "extended lower bound condition" in Section 4.4. That is, one can ask whether there exist methods which guarantee that any coalition of parties which have more than half the vote get at least half the seats. The answer is no, as is shown by the situation (1; 1, 1, 1). For any allotment, there will exist a coalition of two parties with no seats.

141. Here, and again in the discussion of the case $x_1 < x_2$ below, it may be necessary to use a non-integer vote y. This can be avoided if F is scale independent. But then the assumptions, together with membership monotonicity which is proved below in the text, imply that F is a complete or partial divisor method. Hence the proof could have

proceeded as in note 143. The proof that F is a divisor method goes as follows: Consistency and preservation of the majority imply internal vote monotonicity. Hence, by Theorem 11, there exists a generalized divisor method G such that $F \subseteq G$. If, in the notation of Definition 16, M_1 exists, then $F(3M_1; 3, 1, 1) = (M_1, M_1, M_1)$, and F does not preserve the majority. It follows that G is an ordinary divisor method.

142. The submethod of F_{HA} which breaks all ties in favor of the smallest party is consistent and favors small parties compared to F_{HA}. By (a), this method preserves the majority.

143. Let F be a complete or partial divisor method, given by the divisors $d_1, d_2, \ldots$ ($d_1 = 0$ can be permitted.) Suppose that F_{HA} favors small parties compared to F. Theorem 17(b) implies that $\frac{\alpha}{\beta} \geq \frac{d_\alpha}{d_\beta}$ for all $\alpha > \beta \geq 1$. If $F(n; \bar{x}) \to \bar{r}$ and party i has more than half the total vote, $r_i \geq r_j$ for all j. (Being a divisor method, F is internally vote monotone.) Equation (14.1) of the proof of Theorem 18 follows, and one can proceed from there. Conversely, assume that F favors small parties compared to F_{HA}. If $d_1 = 0$, $F(3; 3, 1, 1) = (1, 1, 1)$, and F does not preserve the majority. If $d_1 > 0$, the divisors can be normalized so that $d_1 = 1$. If $d_\alpha = \alpha$ for all α, $F \subseteq F_{HA}$. Otherwise, let α be the smallest integer for which $d_\alpha \neq \alpha$; then $\alpha \geq 2$. $d_\alpha < \alpha$ contradicts Theorem 17(b), hence $d_\alpha > \alpha$. Find integers x and y such that $\alpha x < y < d_\alpha x$. Then $F(2\alpha-1; x, (\alpha-1)x, y) = (1, \alpha-1, \alpha-1)$, and F does not preserve the majority.

144. Again, it may be necessary to choose an x which is not an integer. If F and F' are scale independent, this can be avoided. F is scale independent if $V(n, ay) = aV(n, y)$ for all n, y and positive rational numbers a; and similarly for F'. (In fact, it is sufficient to assume that this holds for positive integers a; see note 18.)

145. See note 28 and accompanying text.

146. The ideas of this section are inspired by [5].

147. For example, consistency requires $F(1; 1, 1) = \{(1, 0), (0, 1)\}$.

148. An example, which has been referred to several times before and will be used again later, is given in note 50.

149. Note that the possibility $x = y$, $r = s$ is included here. Hence $(x, r)R(x, r)$ for all (x, r), that is, R is reflexive.

150. The formulation in [5] is essentially equal. But there the pair (x, r), where $r \geq 0$, represents the party's priority of getting its seat number r + 1, that is, of getting another one when it has r seats. See also note 156.

151. Since R is complete, P could equivalently have been defined by: (x, r)P(y, s) if and only if not (y, s)R(x, r). The version in the text is chosen for later reference.

152. These statements can be proved from transitivity and completeness of R; see discussion in [20], Chapter 1* or [11], Sections 7.1 and 7.2. Unfortunately, the terminology is not uniform. When R and P are as in the text, R will be an ordering in the sense of [20], while P will be a weak order in the sense of [11].

153. The name "Huntington method" is used in [5].

154. Proof: Let R and R' be semi-strict priority relations, defining the methods F and F', respectively. If $R \neq R'$, there exist x, y, r, s such that (x, r)R(y, s) is true and (x, r)R'(y, s) is false (or vice versa). Consider the situation (r+s-1; x, y). It is possible to find r_1, s_1 with $r_1 \geq r$ and F(r+s-1; x, y) $\rightarrow$ (r_1, s_1). But if F'(r+s-1; x, y) $\rightarrow$ (r_2, s_2), the definition implies that (x, r_2)R' (y, s_3+1), where s_3 is the number of seats the second party had when the first party got its seat number r_2. Hence $s_3 \leq s_2$; and if $r_2 \geq r_1$, semi-strictness contradicts the assumptions. Therefore, $F \neq F'$. It is also easy to see that F and F' are different if R is semi-strict while R' is not. (This also follows from the proof of Theorem 21 below.) Assume, however, that R and R' are defined by

(x, r)R(y, s) if and only if $r \geq s$; and

(x, r)R'(y, s) if and only if $r > 1$ or r = s = 1.

Then F and F' are equal; they both consist of all allotments in which one party gets all the seats. But R and R' are formally different.

155. The set of pairs (x, r) is countable, and the existence of f_R can be proved by induction on an enumeration of this set. Proofs can be found in [7], page 200 (Theorem 22) and [10], page 14 (Theorem 2.2). Countability of the set of pairs (x, r) is essential; if irrational votes are allowed, the method of note 50 is an example of a priority method which cannot be represented by a real-valued function. (See discussion in [10], page 27.)

156. In [5], the concepts are defined in terms of representing functions. Such a function is called a rank index.

157. If R is not semi-strict, there may be pairs whose relative priority does not matter, hence the condition on R can be violated while F still has the corresponding property. The method which gives all seats to the largest party (or one of the largest parties, in case of a tie), is vote monotone and scale independent, but it can be represented by relations which do not satisfy the conditions.

158. Examples must be quite strange, but here is one: Let $x \neq y$, let the pairs $(x, 1)$ and $(y, 1)$ have equal priority, and let $(x, 2)$ have higher priority than $(x, 1)$ while $(x, 3)$ and $(y, 2)$ have lower priority than $(x, 1)$. It is clearly possible to find a priority relation for which this holds; the unspecified parts of the relation do not matter. Let F be derived from this relation. Then

$$F(3; x, y, x) \rightarrow (2, 0, 1);$$

the two first seats are given to the first party and the third to the third party. Moreover,

$$F(2; x, y) \rightarrow (1, 1);$$

the first seat is given to the second party. Consistency, applied to the first two parties, now gives

$$F(3; x, y, x) \rightarrow (1, 1, 1).$$

But this is impossible. Before the third seat is distributed, the first or the third party (or both) must have one seat. $(x, 2)$ has higher priority than any other relevant pair, hence that party (one of these parties) must get the third seat and end up with two seats.

In the notation of Definition 8, this counterexample has $k_1 = 2$ and $k_2 > 2$. It is not difficult to see that any priority method will satisfy Definition 8 for $k_2 = 2$.

If a method M in the sense of [2] and [5] is constructed from this F as described in note 36, M will be consistent. In fact, in the framework of [5], every priority method (Huntington method) will be consistent; see the discussion in notes 34 and 36 of the relationship between consistency as defined in this paper and in [5].

The purpose of introducing priority methods was to give a characterization of a broader class of consistent and membership monotone methods than the class of divisor methods (or generalized divisor methods). Therefore, the fact that not all priority methods are consistent can be taken as an indication that the definitions are not entirely satisfactory. (The flaw may lie in the definition of consistency or in that of priority methods.) But if one is interested mainly in balanced methods, Theorem 21 shows that the problem is not a serious one.

159. A similar statement makes up one half of the main result in [5]. (The other half is the opposite implication, which holds in the framework of that paper; see the previous note.) But what is proved in [5] is a little weaker than the stated result, namely the following: Let M_1 be consistent and (membership) monotone. Then there exists a Huntington method (priority method) M_2 such that $M_1 \subseteq M_2$. Examples can be constructed in which $M_2 \neq M_1$. This problem has to do with the "richness of structure" of the system of methods as defined by Balinski and Young

(see discussion in notes 11 and 36), and perhaps also with the relative weakness of their consistency condition. It is a "problem," however, only on a formal level; any consistent and (membership) monotone method is "almost" equal to a Huntington method, and that is what matters for practical applications.

160. The proof of (15.2) is different from, and simpler than, the one in [5]. The difference has to do with differences in basic framework and in the definition of consistency; see notes 11, 34 and 36.

161. The existence of R follows from well-known results in the theory of relations; in particular, Szpilrajn's extension theorem [21]. A discussion can be found in [10], Section 2.3. Only an outline is given here.

Let w_1, w_2, etc. denote pairs of the form (x, r). Define R_0 as the transitive closure of R*. That is, wR_0w' if and only if there exists a sequence $w_1, \ldots, w_k$ such that $w = w_1$; $w_iR^*w_{i+1}$ for $i = 1, \ldots, k - 1$, and $w' = w_k$. (This includes the possibility k = 1, hence wR_0w for all w.) Moreover, define P_0 from R_0 by (15.1). (15.2) can be used to prove that wP*w' implies wP_0w'. Then identify any two pairs w and w' for which wR_0w' and $w'R_0w$. (Formally, this amounts to dividing by an equivalence relation.) No pairs w and w' for which wP*w' will be identified. R_0 and P_0 are well-defined on this reduced domain, and P_0 will be a strict partial order in the sense of [10]. P is now constructed by extending P_0 by Szpilrajn's theorem, and R is defined by wRw' if and only if not w'Pw. (Equivalently, wRw' if and only if w = w' or wPw'.) This gives R on the reduced domain; one gets back to the original domain by letting any pairs which were identified have equal priority in R. It can then be demonstrated that R is a priority relation satisfying (15.13) and (15.14).

The proofs of Szpilrajn's theorem, as given in [21] and [10], use the axiom of choice, or some equivalent mathematical principle. The proofs work regardless of the cardinality of the sets involved. Since the set of pairs (x, r) is countable, it is not necessary to invoke such strong (and disputed) principles. An entirely constructive proof of the extension theorem can be given for the countable case, by induction on some enumeration of the countable set of pairs of the form (w, w'): If (w, w') is the first pair in this ordering, construct R_1 from R_0 as follows: If wR_0w' or $w'R_0w$ or both, then $R_1 = R_0$. Otherwise, $w_1R_1w_2$ if and only if either $w_1R_0w_2$, or w_1R_0w and $w'R_0w_2$. That is, the pair (w, w') is added to R_0, and the transitive closure is taken. R_2 is constructed from R_1 in the same way, using the second pair in the enumeration, etc. Finally, R is the union of R_1, R_2, ... That is, $w_1 R w_2$ if and only if there exists an m such that $w_1 R_m w_2$.

162. This is not true when the number of parties is limited to two, as shown by the following example. For any positive rational number x, let $\sigma(x)$ be the remainder when $\lfloor x \rfloor$ is divided by 3; hence $\sigma(x)$ is equal to 0, 1 or 2. Compute F(n; x, y) by using the method of the highest average, but breaking ties as follows: If $\sigma(x) = \sigma(y)$, break a tie any way. If either $\sigma(x) + 1 = \sigma(y)$, or $\sigma(x) = 2$ and $\sigma(y) = 0$, break ties in favor of the party with x votes. If either $\sigma(x) = \sigma(y) + 1$, or $\sigma(x) = 0$ and $\sigma(y) = 2$, break ties in favor of the party with y votes. This covers all possibilities and gives a well-defined, consistent and membership monotone method when there are only two parties. F is also strongly balanced. If F were a priority method, it must be defined by a strict priority relation. (See Theorem 21 below, the relevant part of the proof does not use situations with more than two parties.) This priority relation, in turn, must define a consistent and membership monotone method on situations with an arbitrary number of parties. (See Theorem 21.) This method must coincide with F for two-party situations, and will also be called F. Consider F(4; 1, 2, 3). At least one element of $T_{3,4}$ must be a possible allotment here. But for each of the 15 elements of $T_{3,4}$, one can pick out two parties, and use consistency to obtain a contradiction to the definition of F for two-party situations. In particular, if a smaller party gets more seats than a larger one, these two parties provide the contradiction. The remaining possibilities are (0, 0, 4), (0, 1, 3), (0, 2, 2) and (1, 1, 2). The two last parties give the counterexample in the first and second case. The third case would give F(2; 1, 2) $\rightarrow$ (0, 2) and the fourth would give F(3; 1, 3) $\rightarrow$ (1, 2), also contradicting the definition of F.

It follows that F is not a priority method. Neither can F be consistently extended to situations with three or more parties.

The relation R* defined from F satisfies (15.2) for k = 2; but for $k \geq 3$ examples can be found which violate (15.2). F was chosen to be strongly balanced, in order to show that the "only if" part of Theorem 21 is also wrong when there are only two parties.

The method constructed in note 98 is also consistent and membership monotone when there are two parties, and it is not a priority method.

163. This will not be proved in detail, but a sketch is given. Let x and r > 0 be given. r is said to be a <u>hurdle</u> for x if (x, s)R(x, r) for all s = 1, ..., r. 1 is always a hurdle for x. If r is a hurdle for x, the pair (x, r-1) is said to be <u>firm</u>. The meaning of this is the following, with reference to Definition 21 and the method F_0: If r is a hurdle for x while r + 1 is not, and a party with x votes gets its seat number r, then that party must immediately be awarded further seats, up to but not including the next hurdle for x. (Provided, of

course, that this many additional seats are to be awarded at all.) Only if (x, r) is firm can a party with x votes have r seats and remain at that number while seats are given to other parties. If $F_0(n; x_1, \ldots, x_k) \rightarrow (r_1, \ldots, r_k)$, at most one of the pairs (x_i, r_i) can fail to be firm.

It is easy to dispose of the case $k_2 = 2$ of Definition 8; the relevant part of the distribution of seats in the k_1-party situation can be simulated in the situation with $k_2 = 2$ parties. Then consider the case $k_1 = 2$, $k_2 > 2$. To simplify the notation, assume $F_0(n; x_1, x_2) \rightarrow (r_1, r_2)$, $F_0(n_2; x_1, \ldots, x_k) \rightarrow (s_1, \ldots, s_k)$ and $r_1 + r_2 = s_1 + s_2$. It is necessary to prove $F_0(n_2; x_1, \ldots, x_k) \rightarrow (r_1, r_2, s_3, \ldots, s_k)$. Suppose that this allocation contains two (or more) pairs which are not firm; in particular, assume that (x_1, r_1) and (x_3, s_3) are not firm. (This is essentially the only possibility.) Consistency, with $k_2 = 2$, gives $F_0(s_i+s_j; x_i, x_j) \rightarrow (s_i, s_j)$ for all i and j with $1 \leq i < j \leq 3$. Lemma 1(b) can be applied to situations with two and three parties (where F_0 is equal to F and therefore consistent) to get $F_0(s_1+s_2+s_3; x_1, x_2, x_3) \rightarrow (s_1, s_2, s_3)$. Another application of consistency to this three-party situation gives $F_0(s_1+s_2+s_3; x_1, x_2, x_3) \rightarrow (r_1, r_2, s_3)$, which is impossible if neither (x_1, r_1) nor (x_3, s_3) is firm.

Consistency of F_0 is now proved in essentially the same way as in the proof of the "if" part of Theorem 21(a). All seats from one hurdle to the next, including the former but not the latter, are lumped together. A formal proof must consider separately the case where $(x_3, s_3), \ldots, (x_k, s_k)$ are all firm, and the case where one of these pairs is not firm.

In the example of note 158, the problem was that the allotment which should have existed by consistency, contained two non-firm pairs. Here this possibility can be ruled out since F is assumed consistent on three-party situations.

164. A corresponding priority relation is given by: $(x, r)R(y, s)$ if and only if either $r > s$, or $r = s$ and $x \geq y$.

165. So far, the condition that F is weakly balanced has been used only to conclude that F is membership monotone. Therefore, the construction of R used in this proof would have been meaningful in the proof of Theorem 20. The resulting R would be a priority relation, F_0 could be constructed, and the proof of $F \subseteq F_0$ would apply. But $F_0 \subseteq F$ could fail

when F is not balanced. For example, let F be the method which gives all the seats to one party but can give the seats to any party; see the end of note 154. F is a consistent priority method, but the construction outlined above gives $F_0 = F_T \neq F$.

REFERENCES

[1] M. L. Balinski and H. P. Young: "The quota method of apportionment." American Mathematical Monthly, 82 (1975), 701-730.

[2] M. L. Balinski and H. P. Young: "Criteria for proportional representation." Research Report 76-20, International Institute for Applied Systems Analysis, 1976.

[3] M. L. Balinski and H. P. Young: "Apportionment schemes and the quota method." Professional Paper 77-3, International Institute for Applied Systems Analysis, 1977.

[4] M. L. Balinski and H. P. Young: "Quotatone apportionment methods." Research Report 77-11, International Institute for Applied Systems Analysis, 1977.

[5] M. L. Balinski and H. P. Young: "On Huntington methods of apportionment." SIAM Journal of Applied Mathematics, 33 (1977), 607-618.

[6] M. L. Balinski and H. P. Young: "The Jefferson method of apportionment." SIAM Review, 20 (1978), 278-284.

[7] G. Birkhoff: Lattice theory, 3rd ed., American Mathematical Society, 1967.

[8] G. Birkhoff: "House monotone apportionment schemes." Proceedings of the National Academy of Sciences of the USA, 73 (1976), 684-686.

[9] A. K. Erlang: "Flerfoldsvalg efter rene Partilister." Nyt Tidsskrift for Matematik, Afdeling B, 18 (1907), 82-83. (In Danish; title translates "Proportional election with pure party lists.")

[10] P. C. Fishburn: Utility theory for decision making, Wiley, New York, 1970.

[11] P. C. Fishburn: The theory of social choice, Princeton University Press, 1973.

[12] C.-E. Fröberg: "Proportionella valmetoder." Nordisk Matematisk Tidskrift, 5 (1957), 91-98. (In Swedish; summary in English on page 120, title translates "Proportional representation systems.")

[13] E. V. Huntington: "The apportionment of representatives in Congress." Transactions of the American Mathematical Society, 30 (1928), 85-110.

[14] Aa. Hylland: Metoder for valgoppgjør, Universitetet i Oslo, 1974. (In Norwegian; title translates "Electoral methods.")

[15] Aa. Hylland: "En merknad til en artikkel fra 1907 om forholdsvalg." Nordisk Matematisk Tidskrift, 23 (1975), 15-19. (In Norwegian; summary in English on page 48, title translates "A note on a paper from 1907 on proportional representation.")

[16] Aa. Hylland: "Proportional representation, methods and criteria." In preparation.

[17] E. Lakeman: How democracies vote, 4th ed., Faber and Faber, London, 1974.

[18] D. W. Rae: The political consequences of electoral laws, 2nd ed., Yale University Press, 1971.

[19] W. Rudin: Principles of mathematical analysis, McGraw-Hill, New York, 1963.

[20] A. K. Sen: Collective choice and social welfare, Holden-Day, San Francisco, 1970.

[21] E. Szpilrajn: "Sur l'extension de l'ordre partiel." Fundamenta Mathematicae, 16 (1930), 386-389.

LIST OF DEFINITIONS, THEOREMS, ETC.

Definitions

Theorems

Theorems (continued)

Theorems (continued)

Lemmas

Problem

ARBEIDSNOTATER
WORKING PAPERS

NR. NO	FORFATTER AUTHOR	TITTEL TITLE
1990/1	Svein S. Andersen og Øivind Revang	Strategisk ledelse og omstilling - Når bedriftsøkonomiske kriterier ikke strekker til
1990/2	Göran Persson	Logistik som konkurrensstrategi
1990/3	Rolf Høyer	Tilsyn med mismanagement
1990/4	Atle Midttun	Norway in the 1980s: Competitive Adaptation or structural Crisis? A Comment on Katzenstein's Small State - Flexible Adjustment Thesis
1990/5	Rolf Høyer	Informasjonsteknologi og brukermedvirkning i 80-årene: Fra ideologi til arbeidsform
1990/6	Hans Mathias Thjømøe	Beslutningsstøttesystemer i markedsføringen - hvor står vi og hvor går vi? En oversikt over ulike tilnærminger
1990/7	Rolv Petter Amdam	Før handelshøyskoler fantes. Utdanning av næringslivsledere før den annen verdenskrig
1990/8	Rolv Petter Amdam	Fra Ludvig Holberg til Finn Øien og Peter Lorange - et perspektiv på BIs historie
1990/9	Aanund Hylland	The Romanian Electroral Law of 1990: Discussion and Comments
1990/10	Nils Arne Bakke og Roland Hellberg	Relevance Lost? A Critical Discussion of Different Cost Accounting Principles in Connection with Decision Making for both Short and Long Term Production Scheduling
1990/11	Aanund Hylland	Allotment Methods: Procedures for Proportional Distribution of Indivisible Entities