

**STABILITY, COALITIONS, AND SCHISMS IN
PROPORTIONAL REPRESENTATION SYSTEMS**

M.L. Balinski* and H.P. Young*

RR-77-17

July 1977

This work was supported in part by the National Science Foundation under Grant MPS 75-07414 with the Graduate School of the City University of New York.

*Graduate School and University Center, City University of New York, and IIASA.

Research Reports provide the formal record of research conducted by the International Institute for Applied Systems Analysis. They are carefully reviewed before publication and represent, in the Institute's best judgment, competent scientific work. Views or opinions expressed therein, however, do not necessarily reflect those of the National Member Organizations supporting the Institute or of the Institute itself.

International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

PREFACE

Task 3 of the System and Decision Sciences Area, "Decision Analysis and Conflict Resolution", is devoted, in part, to the study of the question: what criteria and methods can be developed for the equitable division of resources or other given quantities among competing interests?

This paper treats a specific problem that has wide application in public sector decision making, namely, the allocation of representation to different parties, regions, or interest groups according to their size or to some other measure. In the context of allocating seats to parties this is known as the *proportional representation* problem.

The paper is directly concerned with normative political science, and as such, has policy relevance to decision makers responsible for determining methods of representation.

The objective of the paper is to analyze various proposed methods of allocation in the light of common-sense properties, and to make more precise what is meant by a "fair" or "equitable" solution. Experience has shown that if political debate can be raised from a bickering over particular numbers to an agreement on principles, a solution is more likely to be forthcoming.

In an earlier paper by the authors, *Criteria for Proportional Representation*, various reasonable properties were introduced for proportional representation methods, and detailed mathematical arguments were used to show that various of these properties could be used to characterize particular methods. In this paper the implications of these results are discussed in depth. Moreover, a new property called *uniformity* is introduced which is shown to have important implications for characterizing classes of proportional representation methods.

The methods and approaches developed in this paper, and in other allied papers, should prove to be a methodological step towards treating problems of fair division and conflict resolution found in more complex situations than the one considered here. Thus, this work represents a step in an ongoing program, albeit a step which has immediate relevance to a problem encountered in many countries.

This paper directly addresses the political science community rather than the applied mathematics or systems sciences community. Technical detail is avoided in order to concentrate on normative political issues.

SUMMARY

The following problem arises in many political contexts. A certain number of political "parties" are to split a given number of seats in a legislature according to the numbers of votes they obtained in an election. What principles should determine how these seats are to be allocated?

A basic principle dictated by both common sense and historical precedent is "house monotonicity": if the number of seats to be allocated goes *up*, then, all other factors being equal, no party's allocation should go *down*. Another basic principle is "uniformity": if two parties of given sizes divide a certain number of seats, then they should always divide this number in the same way. These properties characterize an important class of methods called the Huntington methods. A further property relevant to proportional representation is the tendency of some methods to encourage parties to merge by awarding more seats to the merged party than they would receive separately. (Similarly, some methods encourage schisms by penalizing larger parties.) This property singles out one of the Huntington methods--the so-called Jefferson, or d'Hondt method--as particularly desirable. But a third property, that of "satisfying quota" is violated by the Jefferson method. This recommends for consideration a new method, called the Quota method.

The conclusion is that agreement on the desired principles of proportional representation provides a basis for the rational selection of a method to affect allocations. Thus the axiomatic approach of mathematics provides a normative basis for the design of political systems.

Stability, Coalitions, and Schisms in
Proportional Representation Systems

INTRODUCTION

There exist wholesale numbers of possible election procedures. A basic classification of these has been made into "plurality systems" and "proportional representation (PR) systems".

In a *plurality system* an elector usually casts one vote for the candidate or the (party) list of candidates of his or her choice in some election district, and the candidate or the list receiving a majority or a plurality is elected. Such systems are based on a notion of geographical representation. Mid-nineteenth century Europe saw an increasing dissatisfaction with plurality systems as unfair to minorities, for small political parties were effectively barred from having any representation whenever their adherents were distributed throughout many single-member election districts.

This led to the idea of *proportional representation* which, in its pure form, has electors cast one vote for a party or a party list in a multimember district and then, by some rule, metes numbers of seats "proportionally" among the parties according to their respective vote totals. Of course, variants of both types of system exist, as do complex mixtures.

This paper focuses on the pure form of the *proportional representation problem*: voters cast a single vote for a party in a multimember district and the question is to determine the *just* number of representatives due each party. Exact proportionality cannot, in general, be achieved since representation must be integral. Some "rounding" must take place. Appearances to the contrary, an operational definition of exactly how to effect this rounding is not easily forthcoming and history is rich with controversies over proposed solutions and methods.

We argue that the only valid approach to the comparison of methods, and hence to the ultimate choice of a method, is through careful analysis of their *properties*. Consider an arbitrary method of allocation \tilde{M} . Does \tilde{M} have the property of always assuring a party at least its exact proportional number of seats rounded down, or at most that number of seats rounded up? Suppose that, with vote totals unchanged, the number of seats in a parliament is increased: does the method \tilde{M} always assure each party at least as much representation as was allocated to it before? Suppose that, instead of standing alone, two parties form a coalition which obtains the same vote total as the sum of the votes that would have been received by the parties separately: does the method \tilde{M} give at least as many seats to the coalition as the sum of the seats \tilde{M} would have allocated to the parties separately? Postulate an election in which some subset of parties, each standing alone, is allocated certain numbers of seats by a method \tilde{M} and that they jointly hold h seats: if \tilde{M} were used to allocate those h seats among *only* these parties with the given vote totals, would \tilde{M} arrive at the same allocation?

It is perhaps surprising -- but it is a fact -- that *no* method enjoys many of the combinations of the properties suggested by these questions, that some methods are the *only ones* which enjoy certain combinations, and that some other combinations define particular classes of methods. Below, we describe the fundamental properties of allocation methods in the context of methods of allocation which have been advanced over the years since the problem of proportional representation first arose. We point to some impossible combinations, and characterize two methods which seem particularly apt for PR systems. In particular, we introduce the notions of methods being stable, encouraging coalitions, and encouraging schisms. But the choice of appropriate properties depends upon the particular situation, nation, and heritage. Thus the moral of the paper is: politicians should not choose

numerical solutions, or even numerical methods. Rather, they should argue the merits of properties of methods, and let their conclusions in principle determine the methods and thus the numbers. Thus, axiomatization finds its political role.

BACKGROUND

Let us consider s parties, and represent the vote total of each party i by p_i , $1 \leq i \leq s$. h , a nonnegative integer, will be the total number of seats to be allocated (called the *house size*). We consider only those parties which receive votes, so that $p_i > 0$. For a given house size h , the problem is to find an *allocation for h* : namely, s nonnegative integers a_1, a_2, \dots, a_s whose sum is h . A *solution* to the proportional representation problem is a function f which gives, for all possible vote totals $\underline{p} = (p_1, p_2, \dots, p_s)$ and house size h , an allocation for h : $a_i = f_i(\underline{p}, h) \geq 0$, $\sum_i a_i = h$. A *method*, \underline{M} , is the family of all solutions produced by a specific computational rule. (This allows for the possibility of ties.)

In fact a similar problem arises under plurality systems, but in a different guise. For, usually, a country is divided into states or provinces and each single-member election district is wholly contained in one such subdivision. How many election districts or representatives should one geographical region be allocated? This problem is known as the *apportionment problem*. (Of course, geographical apportionment can arise in proportional representation systems too.) In this case the numbers p_1, p_2, \dots, p_s are interpreted as the population (or the number of voters) in the various states, and a_i is the number of seats allocated to state i . The apportionment problem has a long and colorful history, particularly in the United States [1,7,21].

For proportional representation, three principal methods seem to have been considered: Sainte-Laguë's [20], d'Hondt's

[12,13], and Hamilton's [23], the latter usually known as "la répartition au plus fort reste" [10]. In the apportionment literature Sainte-Laguë's and d'Hondt's methods have found their places under other names (see [14]), in a class of five "modern" methods [8,13] which, from about 1920 through 1974, were the ones collectively considered for apportioning the United States House of Representatives.

The five modern methods were first grouped by E.V. Huntington [15] in 1921 via an approach to allocation based on pairwise comparisons of "inequality in representation". Given vote totals $\underline{p} = (p_1, \dots, p_s)$ and an allocation $\underline{a} = (a_1, \dots, a_s)$ for h , consider the numbers p_i/a_i and a_i/p_i . These represent the number of votes per representative of party i and the number of representatives per vote of party i , respectively. If $p_i/a_i > p_j/a_j$, or $a_i/p_i < a_j/p_j$, or $a_j > a_i (p_j/p_i)$, or \dots , $(p_i/a_i)(a_j/p_j) > 1$, then party j is *better off* than party i . Given a particular measure of inequality between a pair of parties such as $|p_i/a_i - p_j/a_j|$ or $|a_i/p_i - a_j/p_j|$ it is natural to ask whether the amount of inequality can be reduced by a transfer of one seat from the better-off party to the less-well-off party. For a given measure an allocation is said to be *in equilibrium* if no transfer of a representative from a "better off" party to a "worse off" party reduces the amount of inequality between them. Of course, certain conceivable measures may not (and do not) admit equilibrium solutions for all vote-total distributions, but Huntington showed [14,15] that five measures do.

All of Huntington's methods are examples of the following type, called *divisor methods*. Let $d(a)$ be some monotone increasing function of the nonnegative integers a where $d(0) \geq 0$. The *divisor method* \underline{M} based on $d(a)$ is defined to be the set of all solutions obtained in the following manner for successive house sizes h . For $h = 0$, the allocation must be zero for every party. Given that an allocation (a_1, \dots, a_s) has been found for a house size h , an allocation for house size $h + 1$ is found by giving one more seat to a party k for which $p_k/d(a_k)$ is a maximum.

If there are several states that are tied for maximum, then several allocations may result unless some tie-breaking rule is employed.

The numbers $p_i/d(a_i)$ may be thought of as measuring the "priority" of a party with a_i seats to receive one more seat. Thus, if one more seat is to be distributed, then party i will "deserve" it more than party j if $p_i/d(a_i) > p_j/d(a_j)$. In this context, $d(a_i)$ may be thought of as some sort of "weighting" of the number of seats that party i already has.

One of the most commonly used methods for proportional representation (e.g., in Argentina, Belgium, Brazil, Finland, Israel, Liechtenstein, and the Netherlands [16]) is that of *d'Hondt*, a nineteenth century Belgian lawyer and proportional representation advocate. This method, which was actually first proposed by Thomas Jefferson [11] in 1792, has therefore been called the *Jefferson method* J [7]. In fact the same method is known in the literature variously as the method of "highest average" or "*la répartition à la plus forte moyenne*" and is described differently, though in fact it leads to the same result (see Section 4 below).

The Jefferson method is one of the five divisor methods studied by Huntington, and it uses the divisor criterion $d(a) = a + 1$. The rationalization for this particular criterion is as follows. If party i has a_i seats then p_i/a_i -- the number of votes per seat -- is a measure of how well-represented that party currently is. If one *more* seat were available for distribution, the Jefferson method proceeds by giving the extra seat to a party that would be the worst-off were *every* party to get one more seat, i.e., by giving the extra seat to a party k with the largest $p_k/(a_k+1)$.

Another plausible criterion, one might argue, would be to always give the "additional" seat to the party that is *currently* the worst off, i.e., to the party for which p_k/a_k is a maximum.

This method is known in the apportionment literature as the "method of smallest divisors" (SD), and is also one of Huntington's five. These five methods, and their corresponding divisor criteria, are shown in Table 1.

Table 1. Huntington's five methods.

Method	Divisor Criterion $d(a)$
Smallest Divisors (SD)	a
Harmonic Mean (HM)	$2a(a+1)/(2a+1)$
Equal Proportions (EP)	$(a(a+1))^{\frac{1}{2}}$
Webster (W) or Major Fractions or Sainte-Laguë Formula	$(a+\frac{1}{2})$
Jefferson (J) or Greatest Divisors or d'Hondt or plus forte moyenne	$a+1$

That these are actually all different methods is seen from the example in Table 2, where the five methods allocate 36 seats among six parties in five different ways.

Table 2.

<u>Party</u>	<u>Votes Received</u>	<u>Exact Proportionality</u>	<u>Exact</u>				
			<u>SD</u>	<u>HM</u>	<u>EP</u>	<u>W</u>	<u>J</u>
A	27,744	9.988	10	10	10	10	11
B	25,178	9.064	9	9	9	9	9
C	19,947	7.181	7	7	7	8	7
D	14,614	5.261	5	5	6	5	5
E	9,225	3.321	3	4	3	3	3
F	3,292	1.185	2	1	1	1	1
	100,000	36.000	36	36	36	36	36

Another commonly used rule, e.g., in Denmark and Norway [16], is that known as Sainte-Lagüe's method [15] (fourth in the list of Huntington's methods). This is also known in the apportionment literature as "the method of major fractions", but was in fact first suggested in embryonic form by Daniel Webster [24] in 1832 and has therefore been called the *Webster method* \tilde{W} [7]. A particular variant of \tilde{W} is the modified "method of odd numbers" used in Sweden. It is defined by: $d(0) = 7/10$ (instead of $d(0) = \frac{1}{2}$), and otherwise the divisors are identical with those of \tilde{W} , $d(a) = a + \frac{1}{2}$. The third of Huntington's methods, Equal Proportions (EP), was the one favored by Huntington, and is the method currently used to apportion the United States House of Representatives.

It is an interesting historical note that Sainte-Lagüe [20] came upon the Webster method quite independently via the idea of minimizing a total measure of the inequality of an allocation. He proposed that an allocation should minimize

$$\sum_i p_i \left(\frac{h}{p} - \frac{a_i}{p_i} \right)^2 ,$$

since in a perfect allocation $h/p = a_i/p_i$ for all i . The Webster method provides solutions which do this. In the same paper Sainte-Lagüe suggests in words (though not in symbols) that one could be interested in minimizing

$$\sum_i a_i \left(\frac{p}{h} - \frac{p_i}{a_i} \right)^2 ,$$

but that "one is led to a more complex rule". In fact, this gives precisely the method of Equal Proportions.

There is another method that is often proposed for proportional representation, and which is, seemingly, the most natural one. Although known by several names, including "*la répartition*

au plus fort reste" and "Vinton's method of 1850", it was apparently first proposed by Alexander Hamilton [23] in 1792 and has therefore been called the *Hamilton method* \underline{H} [7].

Define the *exact quota* of party j to be $q_j = q_j(p, h) = p_j h / \sum_i p_i$; it is the exactly proportional number of seats deserved by party j and the number that one would wish to allocate to j were it integral. Let $\lfloor q_j \rfloor$ denote the largest integer less than or equal to q_j . The *Hamilton method* is defined in the following way: first, give to each party j $\lfloor q_j \rfloor$ seats; then order the parties by their fractional remainders $d_j = q_j - \lfloor q_j \rfloor \geq 0$ in a priority list $d_{j_1} \geq \dots \geq d_{j_s}$. Second, give one additional seat to each of the first $h - \sum \lfloor q_j \rfloor$ parties on the list. If there are ties then there exist distinct arrangements of the priority list and, hence, possibly several solutions.

The rationale of the Hamilton method is the following. Every party j certainly deserves a number of seats at least equal to its *lower quota* $\lfloor q_j \rfloor$, hence we begin by giving each party at least this many seats. Any method such that

$$a_j \geq \lfloor q_j \rfloor = \lfloor p_j h / p \rfloor$$

for all allocations \underline{a} for h is said to *satisfy lower quota*. By the same token, no party j deserves to receive more than $\lceil q_j \rceil$ seats (where $\lceil q_j \rceil$ is the next integer larger than q_j , or if q_j itself is integer, $\lceil q_j \rceil = q_j$). $\lceil q_j \rceil$ is called the *upper quota* of state j , and a method *satisfies upper quota* if for all allocations \underline{a} for h ,

$$a_j \leq \lceil q_j \rceil = \lceil p_j h / p \rceil .$$

A method \underline{M} *satisfies quota* if it satisfies both lower and upper quota for all parties j and allocations \underline{a} . In particular, the Hamilton method satisfies quota.

The Hamilton method, or modifications of it, is currently used for PR in the parliaments of Costa Rica and Italy [16], and was formerly used in Israel. Under the name of the *Vinton method of 1850* it was also used to apportion the United States House of Representatives from 1850 to 1900.

In the United States experience with the Hamilton method a startling discovery was made: in 1881, it would have given Alabama 8 seats in a house of 299 but only 7 seats in a house of 300. "This atrocity which [mathematicians] have elected to call a 'paradox' ... this freak presents a mathematical impossibility" (Representative John C. Bell of Colorado, 8 January 1901, [9, pp. 724-725]) is a typical, albeit somewhat hysterical, politician's reaction to the property. The phenomenon, known as the *Alabama paradox*, is not an isolated quirk of the Hamilton method, but occurs frequently. Consider the vote totals of the six parties in Table 3: parties D and E are given a total of 10 seats when 37 are to be allocated, but only 8 when 38 are to be allocated. The point is not that one would necessarily be in a situation where the number of seats to be allocated increases by 1--although history is replete with houses of representatives which increase their numbers--but rather that any method exhibiting such behavior runs totally contrary to any notion of what a fair and reasonable method should do. One need only refer to the universal reaction of shock among politicians when the event occurred to gauge the *political* unacceptability of the Hamilton method. Its use in the United States has been definitely abandoned. Were a political *party* to lose a seat only because the total number of seats to be distributed had been increased, a similar hue and cry would certainly result.

Table 3.

Party	A	B	C	D	E	F	Totals
Vote Total	27,744	25,178	19,947	14,614	9,225	3,292	100,000
Exact quota 35	9.711	8.812	6.982	5.115	3.229	1.152	35
H allocation 35	10	9	7	5	3	1	35
Exact quota 36	9.998	9.064	7.181	5.261	3.321	1.185	36
H allocation 36	10	9	7	5	4	1	36
Exact quota 37	10.265	9.316	7.380	5.407	3.413	1.218	37
H allocation 37	10	9	7	6	4	1	37
Exact quota 38	10.543	9.568	7.580	5.533	3.506	1.251	38
H allocation 38	11	10	8	5	3	1	38
Exact quota 39	10.820	9.819	7.779	5.699	3.598	1.284	39
H allocation 39	11	10	8	6	3	1	39

Thus, an essential property for any fair and reasonable PR method \underline{M} is that it be *house monotone*, that is, for any \underline{M} -solution \underline{f} , no party must receive fewer seats if the number of seats to be allocated increases:

$$\underline{f}(p, h+1) \geq \underline{f}(p, h) \quad \text{for all } p \text{ and } h .$$

It was the search for house monotone methods that motivated Huntington, beginning in the early part of this century, to formulate the class of five methods exhibited in Table 1. All of these methods are house monotone, as indeed are *all* divisor methods.

STABLE METHODS

Parties in proportional representation systems are dynamic. They may group together for electoral purposes, but they may also splinter. The properties of the method used for allocating representation to parties may well have consequences for this tendency to coalesce or to splinter -- in short, for their *stability* [19]. Specifically, it is pertinent to ask: how does the number of seats allocated by a method \tilde{M} to the joint vote total of two parties coalesced into one compare with the seats allocated by \tilde{M} to the two parties separately?

This property is independent of any coalition or schism formation which might occur in a parliament *after* an election. Rather, it forms an underlying structural incentive for larger or smaller parties to form *before* elections. This type of incentive is institutional rather than psychological and cannot be expected to be directly observable or measurable in any one election. Nevertheless, it constitutes a normative basis for the design of proportional representation systems.

Consider a method \tilde{M} and a situation with vote totals p in which some party has p^* votes and is allocated a^* seats, and another party has \bar{p} votes and is allocated \bar{a} seats. Now suppose that the two parties merge into one party with a combined vote total $\bar{p} + p^*$. The method \tilde{M} is said to be *stable* if there is an \tilde{M} -allocation of seats that gives the merged party *no more* than $a^* + \bar{a} + 1$ seats and *no less* than $a^* + \bar{a} - 1$ seats. The meaning of this condition is that the allocations to parties separately should not be too different from what the parties would receive by merger, for if this is not the case then the method itself may tend to encourage party mergers or schisms, i.e., instability.

The following result may then be shown [3].

- (1) *Any divisor method with divisor criterion $d(a)$ satisfying*

$$d(a_1 + a_2) \leq d(a_1) + d(a_2) \leq d(a_1 + a_2 + 1)$$

is stable.

It may be checked that the five methods of Table 1 satisfy the condition, hence all of them are stable.

To see how stability works, consider the vote totals of the six parties in Table 2, and suppose that parties D and E merge to form one party. The Jefferson method then accords the merged party one more seat than the sum of their separate allocations. On the other hand, consider the seemingly similar divisor method based on $d(a) = a - 1$. This does not satisfy the condition (1), and the example of Table 3 shows it is unstable: the method allocates to the merged party a total of 2 seats *less* than the sum of their separate allocations (see Table 4).

There are nondivisor methods which are stable; for example it is proven in the appendix that

(2) *The Hamilton method is stable.*

Table 4. J compared with an unstable method.

Party	Vote totals before merger	Vote totals after merger	J-solution before	J-solution after	(a-1)-sol. before	(a-1)-sol. after
A	27,744	27,744	11	10	9	10
B	25,178	25,178	9	9	8	9
C	19,947	19,947	7	7	7	7
D	14,614	23,839	5	9	6	8
E	9,225		3			
F	3,292	3,292	1	1	2	2
	<hr/> 100,000	<hr/> 100,000	<hr/> 36	<hr/> 36	<hr/> 36	<hr/> 36

METHODS ENCOURAGING COALITIONS AND METHODS ENCOURAGING SCHISMS

For proportional representation it is important to ask not only whether a method is stable, but also whether it tends to encourage parties to merge or to splinter. In other words, the way in which a method allocates seats may create a subtle institutional incentive for larger or smaller parties to evolve, depending on whether the merger of parties tends to result in a net gain or a net loss of seats. For *political* stability it would usually be considered desirable to have methods of allocation that encourage parties to merge, by assuring that this would never result in a net loss of seats, and might in fact result in an increase.

To make these ideas precise, consider any method \tilde{M} and any situation \underline{p} in which some party has \bar{p} votes and is allocated \bar{a} seats by \tilde{M} and some other party has p^* votes and is allocated a^* seat by \tilde{M} . Suppose that the star- and bar-parties merge to create a party with a total of $p^* + \bar{p}$ votes. We say that \tilde{M} *encourages coalitions* if in any such situation there is an \tilde{M} -allocation giving at least $a^* + \bar{a}$ seats to the coalesced party. Similarly, if there is an \tilde{M} -allocation giving at most $a^* + \bar{a}$ seats to the coalesced party then \tilde{M} is said to *encourage schisms*.

Consider, for example, the divisor method with divisor criterion $d(a) = a + 6$ applied for 36 seats to the vote totals given in Table 2, and compare the result with that of the Jefferson method (see Table 5). If parties B and C merge, then they receive *four* more seats under the $(a + 6)$ -method, and one more under the Jefferson method. Thus under both methods there is an *incentive* for parties to merge. However, the $(a + 6)$ -method is unstable. Moreover, it will be noticed that, while the $(a + 6)$ -method works even more strongly than \tilde{J} to encourage coalitions it does so at the cost of seriously penalizing the smaller parties.

Table 5. Example of encouraging coalitions.

Party	Vote totals before merger	Vote totals after merger	Exact quotas before	Exact quotas after	\underline{J} -sol. before	\underline{J} -sol. after	(a+6)-solution before	(a+6)-solution after
A	27,744	27,744	9.988	9.988	11	10	13	11
B	25,178	45,125	9.064	16.245	9	17	11	22
C	19,947		7.181		7		7	
D	14,614	14,614	5.261	5.261	5	5	4	3
E	9,225	9,225	3.321	3.321	3	3	1	0
F	3,292	3,292	1.185	1.185	1	1	0	0
	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
	100,000	100,000	36.000	36.000	36	36	36	36

Thus, *none* of the three smallest parties gets even its lower quota under the (a + 6)-method, whereas all of them do so under \underline{J} . Since any party certainly has a very strong claim in equity to be allocated at least its lower quota, the Jefferson method seems to be the superior one of the two. In fact, we have the following result.

- (3) *The Jefferson method is the only divisor method that satisfies lower quota and encourages coalitions.*

This is a consequence of more general results proved in [3,6].

Viewed in this light the Jefferson method presents strong credentials for being adopted in a proportional representation system. Sainte-Laguë appears to have realized the tendency of \underline{J} to encourage coalitions, but he gave no proofs and his statement has the curiosity of referring to a comparison: "In comparing the two rules, one can show that the d'Hondt rule (\underline{J}) favors

the grouping of parties which, by coalescing, may receive more seats; whereas the method of least squares (\tilde{W}) favors neither groupings nor schisms" [20, p.378]. A unique characterization of \tilde{W} may, however, be given in terms of its "rounding" properties (see [2]).

The method of smallest divisors is, in a certain sense, "symmetric" with \tilde{J} . Thus we have the following result (which is also a consequence of theorems in [3,6]).

- (4) *The method of smallest divisors is the only divisor method that satisfies upper quota and encourages schisms.*

UNIFORM METHODS

A "method" of allocation brings to mind the dictionary definition: a systematic means or manner of procedure. It conveys a sense of regularity. What is a valid mathematical interpretation of this meaning? Here we introduce and explain a notion which captures this sense.

Consider a two-party problem in which each party has precisely the same number of votes and there is an even number of seats to allocate. We say that a method is *balanced* if it has an allocation giving the same number of seats to each party. (An equivalent definition is given in [3].) It is an obvious truism that every method ever proposed or considered is balanced. In the sequel we will consider only balanced methods--except to point to the existence of "unbalanced" ones -- and so will drop any further (redundant) mention of this property.

Divisor methods (which are of course balanced) are house monotone. But they also have another property--"uniformity" --which is a logical transcription of the intuitive notion of "method". Consider the allocation problem of Table 2, and suppose, for example, that the problem is modified as follows: parties E and F are disqualified, so the total vote is 87,483

and there are 32 seats amongst the remaining states. If the Webster method \underline{W} were now used to allocate a house of 32 seats to parties A, B, C, and D, then one would naturally expect that the number of seats would be exactly the same as those in Table 2, namely, 10, 9, 8, and 5, respectively. That is, dropping certain parties and the seats allocated to them should not change the distribution of seats among the parties that remain. The solution might be said to be "secession-proof". One actually would expect more of a "method": *whenever* four parties having the vote totals of A, B, C, and D are to share 32 seats among them, then these seats should always be shared in the same way (except allowing for ties). This expectation of "uniformity" is fulfilled by the Webster method, all divisor methods and, in fact, by a vastly larger class of methods.

To be mathematically precise, consider an arbitrary method \underline{M} and two different problems with vote totals $(\underline{p}, \underline{q})$ and $(\underline{p}, \underline{q}')$, each problem containing a set of parties with identical vote totals \underline{p} . Suppose that $(\underline{a}, \underline{b})$ is an \underline{M} -allocation for $(\underline{p}, \underline{q})$, that $(\underline{a}', \underline{b}')$ is an \underline{M} -allocation for $(\underline{p}, \underline{q}')$, and $\sum a_i = \sum a'_i$. Then \underline{M} is *uniform* if $(\underline{a}', \underline{b})$ is also an \underline{M} -allocation for $(\underline{p}, \underline{q})$, and $(\underline{a}, \underline{b}')$ also an \underline{M} -allocation for $(\underline{p}, \underline{q}')$.

Not all methods are uniform: the Hamilton method is not. For, consider the example of Table 3. The Hamilton method uniquely divides 10 seats among D, E, and F by the allocation (5,4,1) in one instance (when the total house is 36), whereas it uniquely allocates them (6,3,1) in another instance (when the total house is 39). On the other hand, divisor methods are uniform (see (6) below) since they proceed by comparing the priorities of parties pair by pair through evaluation of $p/d(a)$, so the order in assigning seats to a subset of parties need not change in the presence of other parties. This observation immediately suggests a generalization.

Let $r(p, a)$ be any real-valued function of two real variables called a *rank index* (possibly including $\pm \infty$ for certain values of pa and a). Given a rank index, a (generalized) *Huntington*

method [3] of allocation \tilde{M} is the set of all solutions \tilde{f} obtained recursively as follows:

$$(i) \quad f_i(\underline{p}, 0) = 0 \quad , \quad 1 \leq i \leq s \quad .$$

(ii) If $a_i = f_i(\underline{p}, h)$ is an \tilde{M} -allocation for h , and k is some one party for which

$$r(\underline{p}_k, a_k) \geq r(\underline{p}_i, a_i) \quad \text{for} \quad 1 \leq i \leq s \quad ,$$

then

$$f_k(\underline{p}, h+1) = a_k + 1 \quad , \quad f_i(\underline{p}, h+1) = a_i \quad \text{for} \quad i \neq k.$$

Since we consider only balanced methods, we must have $r(\underline{p}, a - 1) \geq r(\underline{p}, a)$ for all \underline{p} and a , because otherwise an allocation to the two-party problem $(\underline{p}, \underline{p})$ of form $(a-1, a+1)$ would result. Huntington methods are a direct generalization of divisor methods and are house monotone and uniform for the same reasons. They admit a local characterization of allocations.

(5) \underline{a} is an allocation for the Huntington method \tilde{M} based on $r(\underline{p}, a)$ if and only if (with $r(\underline{p}, -1) = \infty$)

$$\max_i r(\underline{p}_i, a_i) \leq \min_i r(\underline{p}_i, a_i - 1) \quad .$$

This is immediate by definition. It applies to divisor methods; thus, for example, \underline{a} is a Jefferson allocation if and only if

$$\max_i P_i / (a_i + 1) \leq \min_i P_i / a_i \quad .$$

One application of this result is to show that various methods that go by different names and descriptions in the literature are actually the same method. For example, the method of *la répartition à la plus forte moyenne* -- usually ascribed to d'Hondt -- has the following description [17]. First, give to each state its lower quota $\lfloor p_i h / p \rfloor = a_i$; the number of seats remaining to be distributed is then $h - \sum_i a_i$. For these remaining seats, proceed as for \tilde{J} , (that is, as for the Huntington method with $r(\underline{p}, a) = p / (a+1)$). But this must give exactly the same

answer as \tilde{J} , since as we have noted above (3), any \tilde{J} -solution automatically satisfies lower quota.

D'Hondt actually made the following proposal: "... to allocate discrete entities proportionally among several numbers, it is necessary to divide these numbers by a common divisor, producing quotients whose sum is equal to the number of entities to be allocated" [13, p.22]. Thus, according to d'Hondt, we are to find a "divisor" λ such that the integer parts of the numbers p_i/λ sum to h , that is, $\sum_i \lfloor p_i/\lambda \rfloor = h$. To see that this is the same as \tilde{J} , set $a_i = \lfloor p_i/\lambda \rfloor$ and notice that, by definition of the integer part,

$$a_i + 1 > p_i/\lambda \geq a_i .$$

Thus

$$p_i/(a_i+1) < \lambda \leq p_i/a_i \quad \text{for all } i ,$$

that is,

$$\max_i p_i/(a_i+1) \leq \min_i p_i/a_i ,$$

so by (5) (a_1, \dots, a_s) is a \tilde{J} apportionment for h . The converse is established similarly.

But exactly the same idea was proposed by Thomas Jefferson in a letter to George Washington nearly a century earlier (1792): "for *representatives* there can be no such common ratio, or divisor which ... will divide them exactly without a remainder or fraction. I answer then ... that representatives [must be divided] as *nearly* as the *nearest* ratio will admit; and the fractions must be neglected" [11, p. 463]. For this reason the divisor method \tilde{J} must be credited to Jefferson.

That Huntington methods constitute a very general class may be seen from the following result (which follows from (7) below and the main theorem of [4]).

- (6) A (balanced) \tilde{M} is house monotone and uniform if and only if it is a Huntington method based on a rank index $r(p,a)$ which is nonincreasing in a .

House monotonicity and uniformity are sufficient to imply the existence of a rank index $r(p,a)$. This is a strong implication. It permits a host of pertinent conclusions. It should be noted that there exist rather dubious non-balanced Huntington methods characterized by rank indices such as $r(p,a) = a/p$: this method gives all seats to the party which receives the first seat.

The statement (6) permits a strengthening of some of the previous results. Thus (3) can be stated: there is exactly one uniform, house monotone method satisfying lower quota and encouraging coalitions, namely, the Jefferson method. And (4) can be reformulated: there is exactly one uniform, house monotone method satisfying upper quota and encouraging schisms, namely, the method of smallest divisors.

There is another way of viewing uniformity. Consider an arbitrary house monotone method \tilde{M} . The behavior of \tilde{M} can be specified simply by saying how one gets an allocation for a house with $h+1$ seats, *given* the allocation for a house with h seats; that is, all that is needed to be known is which state has priority in receiving the next seat. An arbitrary method \tilde{M} is *consistent* if the decision as to which party of any pair most deserves the extra seat when the house size is increased by 1, depends upon *only* the vote totals and the seats already allocated to those parties singly, and not on any other data (such as the house size, number of other parties, or vote totals of other parties). The example of Table 3 shows that the Hamilton method is not consistent. Parties D and E receive 5 and 3 seats, respectively, for $h = 35$ and $h = 38$, but uniquely receive 5 and 4 for $h = 36$ and uniquely 6 and 3 for $h = 39$. However, it is immediately clear that all Huntington methods are consistent. Indeed, consistency and uniformity are essentially the same:

- (7) *A (balanced) house monotone method \tilde{M} is uniform if and only if it is consistent.*

For the proof of this result see the Appendix.

Unfortunately,

- (8) *There is no uniform, house monotone method which satisfies quota.*

If a method is house monotone and uniform then it is a Huntington method (see Appendix). But the Jefferson method is the only Huntington method that satisfies lower quota and the method of smallest divisors is the only method that satisfies upper quota [6,7]. Since these methods are different, there can be no uniform, house monotone method satisfying quota.

QUOTATONE METHODS

Three properties appear to dominate in importance -- house monotonicity, uniformity, and satisfying quota -- but they cannot be satisfied by any one method. This raises the question: is there some "minimal" weakening of these conditions which admits the existence of a method?

The need for house monotonicity has already been demonstrated. Allocations which do not satisfy quota -- that is, which are not the result of rounding the exact proportional share of representation due to a party either up and down -- seem to violate common sense and have proven to be politically subject to attack (see, for example, [19,24]). A party receiving 40 seats when its proportional share is 37.34 seems unreasonably well served; if its true share were instead 43.34 its leadership would justifiably complain. Satisfying quota seems essential. Note that any method satisfying quota is *almost stable* in the sense that if any two parties with a^* and \bar{a} coalesce, then the coalesced party receives b seats where $a^* + \bar{a} - 2 \leq b \leq a^* + \bar{a} + 2$.

Given these considerations it is reasonable to attempt to weaken the uniformity or consistency condition. Specifically, we say that a method \underline{M} is *quota consistent* if it is consistent subject to the need to satisfy upper quota, that is, consistent unless this imperative is in conflict with satisfying upper quota, in which case it cedes to the latter stronger imperative. Then, it can be affirmed ([1,7]).

- (9) *There exists a unique allocation method \underline{Q} , the quota method, that is house monotone, quota consistent, and satisfies quota.*

The *quota method* \underline{Q} is related to the Jefferson method, and may be described as the set of allocations computed recursively as follows:

(i) $f_i(\underline{p}, 0) = 0$, $1 \leq i \leq s$.

- (ii) Let $a_i = f_i(\underline{p}, h)$ be an allocation for h and let $E(h+1)$ be the set of parties which can be given an extra seat without violating upper quota at house size $h + 1$. If $k \in E(h+1)$ is some one party for which

$$p_k(a_k + 1) \geq p_i/(a_i + 1) \text{ for all } i \in E(h+1)$$

then

$$f_k(\underline{p}, h+1) = a_k + 1 \text{ , } f_i(\underline{p}, h+1) = a_i \text{ for } i \neq k \text{ .}$$

Statement (9) gives powerful reasons for acceptance of the quota method \underline{Q} in problems of proportional representation. In practice one finds that \underline{Q} has a tendency to produce solutions that round up the exact quotas of large parties more often than those of small parties. This seems reasonable for the application of \underline{Q} to proportional representation systems in that it inferentially asks for a "large" vote before according any representation at all. Notice, however, that no large party is allowed more seats than its upper quota.

One can simply drop the uniformity or consistency condition and ask about the class of *all* methods that are house monotone and satisfy quota. These are called *quotatone methods*. They are describable and have been characterized in several different ways [5,22]. There are a multiplicity of allocation methods which are quotatone and all operate as follows. In a house $h = 0$ all parties have no representation. Given an allocation $\tilde{f}(p, h) = \tilde{a}$ for h , the characterization specifies a set of parties which are "eligible" to receive an extra seat, eligible meaning that this can be done in a manner which will guarantee quota and house monotonicity. Some choice mechanism (e.g., a rank index) can be imposed to determine which party receives the seat at $h + 1$ if there is more than one eligible party. The difficulty is that the set of eligible parties has no rational direct interpretation other than that just given, and that its computation is conceptually complex (although easily accomplished by computer).

In the class of quotatone methods the quota method \underline{Q} enjoys an especially natural position [5]; moreover, its eligibility set is easily understood and computed since it simply asks that the upper quota condition not be violated. It is the unique method which devolves from a minimal weakening of the consistency condition.

CONCLUSIONS

This paper has addressed the problem of the allocation of integral representation to parties having vote totals in a proportional representation system. The principal point is that methods of allocation should not be chosen by bickering over numbers, nor, indeed, through ad hoc claims of various mechanical procedures, but rather through analysis of the properties of methods. The issue is to decide upon a method whose qualitative properties are equitable for the situation at hand. This analysis commends one of two methods: the quota method \underline{Q} , or the Jefferson method \underline{J} .

The Jefferson method claims recognition because it is house monotone, uniform, guarantees lower quota, and encourages coalitions. Specifically, encouraging coalitions would seem to be precisely the type of stability desired for a body politic operating a proportional representation system. However, a major defect of \underline{J} is that it fails the seemingly most common-sense test of satisfying quota. The Quota method merits recognition because it does satisfy quota, is consistent subject to that property, and is house monotone. It is "almost stable" (see [3]), but does not necessarily encourage coalitions.

The choice between these methods of allocation should be made in terms of which *criteria* are viewed as most important for the situation in question.

APPENDIX

We first prove that the Hamilton method is stable (see (2)).

Consider any two parties, say $i = 1, 2$, and suppose that in a particular problem party i has an exact quota $q_i = n_i + r_i$, $n_i \geq 0$ and integer, and $0 \leq r_i < 1$, and let a_i be their allocations at h . Then for the problem in which parties 1 and 2 form a coalition, its exact quota for h is $q_1 + q_2 = n_1 + n_2 + r_1 + r_2$. Let b be the number of seats given the coalition by the Hamilton method.

We consider several cases. First, if $b = n_1 + n_2$ then $r_1 + r_2 < 1$, implying that the same total number of parties is rounded up in both problems. If $r_1, r_2 > 0$ then it must be that $a_i = n_i$ ($i = 1, 2$). For otherwise one of the two parties would have a remainder r_i high enough in the list to warrant an extra seat while $r_1 + r_2 > r_1, r_2$ is not high enough, which cannot be. If $r_1 = 0$ then $a_1 = n_1$ and $a_2 = n_2$ or $a_2 = n_2 + 1$; in either case the criterion for stability is satisfied.

If $b = n_1 + n_2 + 1$, then since $a_i = n_i$ or $n_i + 1$ there is nothing more to show. If $b = n_1 + n_2 + 2$, then $r_1 + r_2 > 1$. Suppose stability is not satisfied, i.e., that $a_i = n_i$ ($i = 1, 2$). Then for some party $k \neq 1, 2$, $r_1 + r_2 - 1 \geq r_k$ while $r_1 \leq r_k$ and $r_2 \leq r_k$. Thus $2r_k - 1 \geq r_k$ and $r_k \geq 1$, a contradiction.

Next, we show that a balanced house monotone method \tilde{M} is uniform if and only if it is consistent (see (7)).

First, assume that \tilde{M} is house monotone and consistent. The main theorem of [4] states that this implies \tilde{M} is a Huntington method based on some rank index $r(p, a)$. If, in addition, \tilde{M} is balanced then $r(p, a-1) \geq r(p, a)$ and \tilde{a} is an \tilde{M} -apportionment if and only if the min-max inequality of (5) is satisfied. To prove uniformity suppose that (\tilde{a}, \tilde{b}) is an \tilde{M} -apportionment for (\tilde{p}, \tilde{q}) and that (\tilde{a}', \tilde{b}') is one for (\tilde{p}, \tilde{q}') where $\sum_i a_i = \sum_i a'_i$, $\tilde{a} \neq \tilde{a}'$. Then, by (5),

$$\max_{i,j} \{r(p_i, a_i), r(q_j, b_j)\} \leq \min_{i,j} \{r(p_i, a_i-1), r(q_j, b_j-1)\} ,$$

(10) and

$$\max_{i,j} \{r(p_i, a'_i), r(q'_j, b'_j)\} \leq \min_{i,j} \{r(p_i, a'_i-1), r(q'_j, b'_j-1)\} .$$

It must be shown that $(\underline{a}', \underline{b})$ is an \underline{M} -apportionment for $(\underline{p}, \underline{q})$ and $(\underline{a}, \underline{b}')$ is an \underline{M} -apportionment for $(\underline{p}, \underline{q}')$. Since $\sum a_i = \sum a'_i$, $\underline{a} \neq \underline{a}'$ implies that there exist parties k and ℓ with $a_k \geq a'_k + 1$ and $a_\ell \leq a'_\ell - 1$. Therefore

$$r(p_\ell, a_\ell) \leq r(p_k, a_k - 1) \leq r(p_k, a'_k) \leq r(p_\ell, a'_\ell - 1) \leq r(p_\ell, a_\ell) ,$$

the first and third inequalities following from (10), the second and fourth following from $r(p, a)$ being nonincreasing in a_i . Thus every inequality is an equation and the max-min inequalities (10) must both be satisfied as equations having a common value, say λ . Therefore,

$$\max_{i,j} \{r(p_i, a'_i), r(q_j, b_j)\} = \lambda = \min_{i,j} \{r(p_i, a'_i-1), r(q_j, b_j-1)\}$$

and

$$\max_{i,j} \{r(p_i, a_i), r(q'_j, b'_j)\} = \lambda = \min_{i,j} \{r(p_i, a_i-1), r(q'_j, b'_j-1)\}$$

which establishes, by the test (5), the desired result.

Second, assume that \underline{M} is house monotone and uniform. (Here it is not necessary to assume that \underline{M} is balanced.) Consider two problems, (p_1, p_2, q) and (p_1, p_2, q') , and suppose that the first has an \underline{M} -apportionment $(a_1, a_2, \underline{b})$ and the second $(a_1, a_2, \underline{b}')$, that the first has at the next house the apportionment $(a_1+1, a_2, \underline{b})$ whereas the second has at the next house $(a_1, a_2+1, \underline{b}')$, a seeming "switch" in priorities. Then, by uniformity, the first also has an apportionment $(a_1, a_2+1, \underline{b})$ and the second also has $(a_1+1, a_2, \underline{b})$, thus showing that there is a "tie" in priorities and that \underline{M} is consistent.

REFERENCES

- [1] Balinski, M.L., and H.P. Young, A New Method for Congressional Apportionment, *Proceedings of the National Academy of Sciences, U.S.A.*, 71 (1974), 4602-4606.
- [2] Balinski, M.L., and H.P. Young, Apportionment Schemes and the Quota Method, *American Mathematical Monthly* (forthcoming).
- [3] Balinski, M.L., and H.P. Young, *Criteria for Proportional Representation*, RR 46-20, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976; *Operations Research* (forthcoming).
- [4] Balinski, M.L., and H.P. Young, *On Huntington Methods of Apportionment*, RR 77-2, International Institute of Applied Systems Analysis, Laxenburg, Austria, 1977; *SIAM Journal on Applied Mathematics C* (forthcoming).
- [5] Balinski, M.L., and H.P. Young, *Quotatone Methods of Apportionment*, RR 77-11, International Institute of Applied Systems Analysis, Laxenburg, Austria, 1977; *Mathematics of Operations Research* (forthcoming).
- [6] Balinski, M.L., and H.P. Young, The Jefferson Method of Apportionment, *SIAM Review* (forthcoming).
- [7] Balinski, M.L., and H.P. Young, The Quota Method of Apportionment, *American Mathematical Monthly*, 82 (1975), 701-730.
- [8] Bliss, G.A., E.W. Brown, L.P. Eisenhart, and R. Pearl, Report to the President of the National Academy of Sciences, 9 February 1929, in *Congressional Record*, 70th Congress, 2nd Session, 70 (1929), 4966-4967.
- [9] *Congressional Record*, 56th Congress, 2nd Session, vol. XXXIV, Washington, D.C., 1901.
- [10] Cotteret, J.M., and C. Emeri, *Les systèmes électoraux*, Presses Universitaires de France, Paris, 1970.
- [11] Ford, Paul Leicester, ed., *The Works of Thomas Jefferson*, Vol. VI, G.P. Putnam, New York, 1904, 460-471.
- [12] d'Hondt, V., *La représentation proportionnelle des partis par un électeur*, Ghent 1878.
- [13] d'Hondt, V., *Système pratique et raisonné de représentation proportionnelle*, Muquardt, Brussels, 1882.
- [14] Huntington, E.V., The Apportionment of Representatives in Congress, *Transactions of the American Mathematical Society*, 30 (1928), 85-110.

- [15] Huntington, E.V., The Mathematical Theory of the Apportionment of Representatives, *Proceedings of the National Academy of Sciences, U.S.A.*, 7 (1921), 123-127.
- [16] Interparliamentary Union, ed., *Parliaments of the World*, Macmillan, London, 1976.
- [17] *La Grande Encyclopédie*, Librairie Larousse, 7 (1973), "Election", 4189-4197.
- [18] Morse, M., J. von Neumann, and L.P. Eisenhart, Report to the President of the National Academy of Sciences, 28 May, 1948.
- [19] Stein, Rokkan, Elections: Electoral Systems, *International Encyclopedia of the Social Sciences*, Macmillan and the Free Lance Press, New York, 1968, 6-21.
- [20] Sainte-Laguë, La représentation et la méthode des moindres carrés, *Comptes Rendus de l'Académie des Sciences*, 151 (1910), 377-378.
- [21] Schmeckebier, L.F., *Congressional Apportionment*, The Brookings Institution, Washington, D.C., 1941.
- [22] Still, J.W., *A New Class of House Monotone Apportionment Methods*, Yale University, New Haven, Conn., 1976 (mimeographed).
- [23] Syrett, Harold C., ed., *The Papers of Alexander Hamilton, Vol. XI (February 1792-June 1792)*, Columbia University Press, New York, 1966, 228-230.
- [24] *The Writings and Speeches of Daniel Webster, Vol. VI*, National Edition, Little, Brown and Company, Boston, Mass., 1903, 102-123.