THE THEORY OF APPORTIONMENT

M.L. Balinski
H.P. Young

September 1980
WP-80-131

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

A key problem area at IIASA is the study of how goods and resources -- as well as 'bads' such as costs, pollution, and risks -- can or should be shared among different nations, groups, or individuals. This raises the question of what is meant by a fair division -- and, if this question can be answered at all, how fair divisions can be achieved. One of the situations studied in the System and Decision Sciences Area was how to allocate or "apportion" discrete entities in proportion to predetermined claims, a problem which encompasses many situations including for example the apportionment of political rperesentation among different regions and constituencies. The result of this study was the development of a general theory to deal with such problems, together with concrete criteria of fairness which will hopefully prove useful to analyzing larger classes of problems.

THE THEORY OF APPORTIONMENT

M.L. Balinski and H.P. Young

1. THE PROBLEM

A widely held ideal of fair representation is representation in proportion to some numerical criterion. In the United States each state receives seats in the House of Representatives proportionally to its population, but in any case is assured of at least one seat. In France each department is given a number of deputies in like manner, but is assured of at least two. In the European Parliament each country is also accorded representation in proportion to its population, but is assured a minimum number of seats that ranges between 6 and 36 depending upon the country in question.

A specific *problem of apportionment* is given by a vector of "populations" $\underset{\sim}{p} = (p_1, \ldots, p_s)$, an integer "size of house" $h \geqq 0$, and a vector of "minimum requirements" $\underset{\sim}{r} = (r_1, \ldots, r_s) \geqq 0$. The p_i are positive integers and the r_i non-negative integers. Usually the minimum requirements are a common number r. For the United States r = 1 and for France r = 2. However, Canada and the European Parliament are instances where minimum requirements differ. Occasionally, maximum limits may also be imposed on the number of seats allowed to each state, as for example the U.S. Constitution's stated limit of one per thirty thousand, though this is not an issue today. An *apportionment of* h *among* s is, therefore, a vector $\underset{\sim}{a} = (a_1, \ldots, a_s)$ of non-negative integers $a_i \geqq r_i$ that sum to h. In the sequel we will explicitly treat $\underset{\sim}{r} = \underset{\sim}{0}$ and for the most part leave as exercises for the reader the verification of the general cases.

The question which we address is: what is a fair method for determining apportionments? This is done in the framework and the language of allocating representation among geographical regions. However, many other problems have the same form. In proportional representation the problem is to allocate seats proportionally to party vote totals (Cotteret and Emeri (1970). In manpower planning a problem is to allocate jobs in proportion to certain characteristics of the labor pool (Mayberry, 1978). Service facilities -- courts, judges, or hospitals -- may need to be allotted to areas in proportion to the numbers of people to be served. In reporting statistical findings there is the problem of making tables of rounded percentages add up to 100 percent. Any problem in which h objects are to be allocated in non-negative integers proportionally to some numerical criterion belongs to this class, and the theory below applies to it. Some of the principles we discuss are particularly relevant to the regional representation problem while other principles may be more telling for other applications (Balinski and Young 1978b, 1979a).

Example 1.1. Consider the apportionment problem with populations $\underset{\sim}{p}$ = (27 744, 25 178, 19 951, 14 610, 9 225, 3 292) and h = 36. Let $q_i = p_i h/(\sum_j p_j)$ be the *quota* of state i, so that $\underset{\sim}{q}$ = (9.988, 9.064, 7.182, 5.261, 3.321, 1.185) represents the vector of "fair shares". What integer apportionment *should* each state receive?

Example 1.2. Members of the European Parliament are now elected by direct universal suffrage in each of the nine countries that constitute the European Economic Community (E.E.C.).

The apportionment of seats was decided upon in April 1976 after a period of intense negotiation. By previously agreed-upon convention, the apportionment was to be proportional to the populations of the respective countries and yet assure each country at least the number of seats it held in the previous Parliament. The populations of the nine countries were estimated to be (in thousands): Germany 62 041, United Kingdom 56 056, Italy 55 361, France 53 780, Netherlands 13 450, Belgium 9 772, Denmark 5 052, Ireland 3 086 and Luxembourg 357. The total population was 258 955 000. The vector of minimum requirements was $\underset{\sim}{r}$ = (36,36, 36,36,14,14,10,10,6). The negotiation produced the apportionment $\underset{\sim}{a}$ = (81,81,81,81,25,24,16,15,6). Compare this solution with the quotas.

2. ELEMENTARY PRINCIPLES

We seek a method of apportionment, that is, a rule which for every s-vector $\underset{\sim}{p} > 0$ and integer $h \geqq 0$ gives an apportionment of h among s. A single-valued function does not suffice as the concept for a method. For suppose two states with identical populations are to share an odd number of seats 2a+1. There are only two natural solutions, (a,a+1) and (a+1,a), but there is no basis in terms of the proportional ideal for preferring one to the other. Any method that is fair must admit both possibilities as solutions.

Formally, then, define a *method* to be a multiple valued function M, consisting of a set of apportionments of h among s for each s-vector $\underset{\sim}{p} > 0$ and integer $h \geqq 0$. A *particular M-solution* is a single-valued function f with $f(\underset{\sim}{p},h) = \underset{\sim}{a} \in M(\underset{\sim}{p},h)$. A particular M-solution breaks every "tie" in some arbitrary fashion, e.g., it might choose (a,a+1) in the case of two equal states sharing 2a+1 seats. Ties can also arise in more interesting ways that depend on the logic of the particular method used.

The ideal of proportionality immediately imposes several elementary properties that a method should enjoy. If all populations change by the same proportion then, since there is no change in the proportional shares of the states, there should be no change in the set of apportionments. Specifically, M is *homogeneous* if the M-apportionments for $\underset{\sim}{p}$ and h are the same as the M-apportionments for $\lambda\underset{\sim}{p}$ and h, for any positive rational number λ. This means that any positive s-vector of rational numbers may be considered as "populations".

Proportionality concerns the size of populations, not their names or other characteristics. Therefore, permuting the populations to obtain a "new" problem should only result in apportionments which are permuted in the same way. Methods with this property are called *symmetric*.

Proportionality means that whenever a problem can be solved perfectly in integers then it must be. M is *weakly proportional* if whenever an apportionment $\underset{\sim}{a}$ is proportional to $\underset{\sim}{p}$, then $\underset{\sim}{a}$ is the unique M-apportionment for $\underset{\sim}{p} > \underset{\sim}{0}$ when $h = \sum a_i$. Moreover, as the house size grows solutions should increasingly approach the ideal of proportionality. So if $\underset{\sim}{b}'$ is an M-apportionment for $\underset{\sim}{p} > \underset{\sim}{0}$ and $\underset{\sim}{b}$ is integer and proportional to $\underset{\sim}{b}'$ with $\sum b_i < \sum b_i'$ then $\underset{\sim}{b}$ should be the unique apportionment in $M(\underset{\sim}{p}, \sum b_i)$. This means that if M splits 6 seats between two states 3 and 3, then it must split 4 seats between the same two in no way other than 2 and 2. A method that satisfies this condition and is also weakly proportional is called *proportional*. Although all reasonable methods are proportional, much of the theory is developed using the weaker notion.

Ties--where a method gives several different apportionments for the same problem--arise naturally not only from considerations of symmetry but also from changing populations. One expects that as populations change more and more in some direction a point (i.e. an s-vector $\underset{\sim}{p}$) is eventually reached where the method changes apportionments, otherwise the ideal of proportionality could not be met. These natural tie points depend, of course, upon the method that is used. One way of describing a tie point $\underset{\sim}{p}^*$ is to say that arbitrarily small perturbations about it can produce different apportionments: a slight increase in one state's population relative to an other's may result in one apportionment, while a slight decrease may result in another apportionment. Such tie points $\underset{\sim}{p}^*$ may involve irrational numbers. Hence it is important to extend the concept of method to all real populations $\underset{\sim}{p} \in \mathbb{R}^s$, $\underset{\sim}{p} > \underset{\sim}{0}$. We say that a method is *complete* if whenever $\underset{\sim}{p}^n \to \underset{\sim}{p} > \underset{\sim}{0}$ and $\underset{\sim}{a} \in M(\underset{\sim}{p}^n, h)$ for every n, then $\underset{\sim}{a} \in M(\underset{\sim}{p}, h)$. A method M is *completed* by letting $\underset{\sim}{a} \in M(\underset{\sim}{p})$ for $\underset{\sim}{p} \in \mathbb{R}^s$ if and only if there is a sequence of rational s-vectors $\underset{\sim}{p}^n$ converging to $\underset{\sim}{p}$ such that $\underset{\sim}{a} \in M(\underset{\sim}{p}^n)$ for all n. Any natural view of proportional allocation carries with it the idea of completeness; however some consequences of the theory also hold for methods that are not complete.

From this point on, methods of apportionment will always be assumed to be homogeneous, symmetric, weakly proportional, and complete, unless stated otherwise. These are the rock-bottom requirements that must be satisfied by any method that is worthy of consideration.

Proposition 2.1. The methods of Hamilton, Jefferson, Lowndes, Webster, Adams, Dean, and Hill are homogeneous, symmetric, weakly proportional, and complete.

Proposition 2.2. The method of Hamilton is proportional.

Proposition 2.3. The completion of a method is complete, and inherits the properties of homogeneity, symmetry, and weak proportionality.

Proposition 2.4. In the presence of minimum requirements -- or of minimum and maximum requirements -- it is natural to generalize symmetry by attaching the requirements to the particular states (e.g., recall the European Parliament). Thus a method is *symmetric* if permuting the populations results in permuting the apportionments in the same way provided the requirements remain satisfied. Homogeneity and completeness hold in the case of minimum requirements without any further modification. Show that weak proportionality and proportionality can be modified by simply imposing the requirements as constraints.

3. TRADITIONAL APPROACHES

This section reviews the different types of methods that have traditionally been followed. The claims for these methods have typically been based on the computational procedures they employ, in other words on *ad hoc* considerations. Moreover it often happens that these differing computational approaches represent the same method in different guises.

A most natural approach is to compute the quotas and round in the usual way. But this does not always work (e.g., the example of Section 1). Hamilton's method is one way around the difficulty. Another is to choose an ideal district size or *divisor* x, to compute the *quotients* of each state $q_i^x = p_i/x$, and to round these according to some rule. The proposal of Webster was to round these in the usual way: remainders above one-half are rounded up, remainders below one-half are rounded down, and a remainder of exactly one-half may be rounded either up or down -- it is a natural tie point. In general, for any real number z, let [z] denote the integer closest to z. If the fractional part of z is one-half then [z] has two possible values. The *method of Webster* is

$$W(\underset{\sim}{p},h) = \{\underset{\sim}{a} : a_i = [p_i/x] \;, \textstyle\sum_i a_i = h \text{ for some choice of } x\} \;.$$

If there are states having quotients with a remainder of one-half then all possible values of $[p_i/x]$ are admitted that sum to h.

Instead of "ordinary" rounding the same approach may be used with rounding of quotients taken at other points, as was for example proposed by Jefferson and Adams. In general, any rounding procedure may be described by specifying a dividing point d(a) in each interval of quotients [a,a+1] for each nonnegative integer a.

For any real number z a d-*rounding* of z, $[z]_d$, is an integer a such that $d(a-1) \leqq z \leqq d(a)$, which is unique unless z = d(a), in which case it takes on either of the values a or a+1. To avoid more than two-way ties it is required that d(a) < d(a+1). Any monotone increasing d(a) defined for all integers $a \geqq 0$ and satisfying $a \leqq d(a) \leqq a+1$ is called a *divisor criterion*. The *divisor method based on* d is

$$M(\underset{\sim}{p},h) = \{\underset{\sim}{a} : a_i = [p_i/x]_d \text{ and } \textstyle\sum_i a_i = h \text{ for some } x\} \quad .$$

To accommodate the case d(0) = 0 the possibility x = ∞ is allowed. The five traditional methods are described as divisor methods in Table 3.1.

Method :	Adams	Dean	Hill	Webster	Jefferson
d(a) :	a	$a(a+1)/(a+\frac{1}{2})$	$\sqrt{a(a+1)}$	$a+\frac{1}{2}$	a+1

Table 3.1 The Five Traditional Divisor Methods

An alternate but equivalent description is that $\underset{\sim}{a}$ is an M-apportionment if and only if there exists an x such that for all $a_i > 0$, $p_i/d(a_i-1) \geqq x \geqq p_i/d(a_i)$ and for all $a_i = 0$, $x \geqq p_i/d(a_i)$. So the divisor method based on d can also be described in terms of a min-max inequality:

$$M(\underset{\sim}{p},h) = \left\{ \underset{\sim}{a} : \min_{a_i>0} p_i/d(a_i-1) \geqq \max_{a_j\geqq 0} p_j/d(a_j) \ , \sum a_i = h \right\} .$$

Proposition 3.1. Each of the traditional five divisor methods gives a different apportionment for the problem of Example 1.1.

Proposition 3.2. Every divisor method M has particular solutions that avoid the Alabama paradox.

Proposition 3.3. The divisor method M based on d may also be described recursively as follows: (i) $f(\underset{\sim}{p},0) = \underset{\sim}{0}$, (ii) suppose $\underset{\sim}{a} \in M(\underset{\sim}{p},h)$ and k is some state satisfying $p_k/d(a_k) = \max_i p_i/d(a_i)$. Then $\underset{\sim}{b} \in M(\underset{\sim}{p},h+1)$, with $b_k = a_k+1$ and $b_i = a_i$ for $i \neq k$.

Proposition 3.4. A divisor method M based on d for problems with both minimum and maximum requirements $\underset{\sim}{r} \leqq \underset{\sim}{r}^+$ may be described as follows

$$M(\underset{\sim}{p},h) = \{\underset{\sim}{a} : a_i = \text{mid}(r_i, r_i^+, [p_i/x]_d) \text{ and } \textstyle\sum_i a_i = h \text{ for some } x\}$$

where mid(u,v,w) for any three unordered reals satisfying $u \leqq v \leqq w$ is v. Derive the corresponding min-max and recursive description.

There are an infinite number of different divisor methods. How is one to choose among them? An ingenious approach to this question, first suggested by Joseph Hill and fully carried out by E.V. Huntington (1921, 1928), was to make *pairwise comparisons* of state's representations. "Between any two states, there will practically always be a certain inequality which gives one of the states a slight advantage over the other. A transfer of one representative from the more favored state to the less favored state will ordinarily reverse the *sign* of this inequality, so that the more favored state now becomes the less favored, and vice versa. Whether such a transfer should be made or not depends on whether the 'amount of inequality' between the two states after the transfer is less or greater than it was before; if the 'amount of inequality' is reduced by the transfer, it is obvious that the transfer should be made. The fundamental question therefore at once presents itself, as to how the 'amount of inequality' between two states is to be measured" (Huntington, 1928).

Let states i and j, having populations p_i and p_j, be apportioned a_i and a_j seats respectively. It is unambiguous to say that i *is favored relative to* j if and only if $a_i/p_i > a_j/p_j$. One natural measure of the inequality between i and j is therefore $|a_i/p_i - a_j/p_j|$.

Huntington's argument is that if this inequality can be reduced by a transfer of seats between i and j, then such a transfer should be made. In particular, if $|(a_i-1)/p_i - (a_j+1)/p_j| < |a_i/p_i - a_j/p_j|$, then i should give up one seat to j. Of course it is quite conceivable that *every* apportionment is *unstable*, i.e. admits such transfers. Remarkably enough this is not the case. An apportionment admits *no* such transfers if for all pairs i and j with $a_i/p_i \geqq a_j/p_j$

$$a_i/p_i - a_j/p_j \leqq (a_j+1)/p_j - (a_i-1)/p_i$$

or

$$p_i/(a_i-\tfrac{1}{2}) \geqq p_j/(a_j+\tfrac{1}{2}) \qquad (3.1)$$

Therefore such an $\tilde{a}$ must be a Webster method apportionment. Conversely, every Webster method apportionment satisfies (3.1), hence satisfies the transfer test. In particular this pairwise comparison approach has produced a house monotone method!

Unfortunately for this logic the statement that i is favored relative to j can be expressed in many different ways. The inequality $a_i/p_i > a_j/p_j$ can be rearranged by cross-multiplication in $2^4 = 16$ different ways. Hence to measure the inequality between states i and j it would be equally valid to consider the inequalities between the numbers p_i/a_i and p_j/a_j, or between a_i and $a_j p_i/p_j$, or p_i and $p_j a_i/a_j$, or $p_i a_j/p_j a_i$ and 1, or p_i/p_j and a_i/a_j, etc.

Not every measure of inequality gives stable apportionments: for some measures there exist problems for which every apportionment can be improved upon by some transfer. Huntington showed that, except for four such "unworkable" measures, all others resulted in the methods of either Adams, Dean, Hill, Webster, or Jefferson. Examples of tests that lead to these methods are given in Table 3.2. Huntington argued that it is not the absolute

Method :	Adams	Dean	Hill	Webster	Jefferson
Test : (for $a_i/p_i \geq a_j/p_j$)	$a_i - a_j(p_i/p_j)$	$p_j/a_j - p_i/a_i$	$\dfrac{a_i/p_i}{a_j/p_j} - 1$	$a_i/p_i - a_j/p_j$	$a_i(p_j/p_i) - a_j$

Table 3.2 Pairwise Comparison Tests for Five Traditional Methods

difference that should be used in measuring the inequality between two numbers y and z, but the relative difference $|y-z|/\min(y,z)$, and he observed that the relative differences in all 16 cases are the same: $a_i p_j / a_j p_i - 1$. All relative differences yield the method of Hill, or what Huntington called the method of equal proportions. This is a neat argument, yet it boils down to a question of preference for one among several competing tests of inequality.

Proposition 3.5. The test $|a_i/a_j - p_i/p_j|$ does not always yield stable apportionments. Use the three state example with $\underset{\sim}{p} = (762, 534, 304)$ and $h = 16$ to show that no stable solution exists.

Proposition 3.6. Modify the pairwise comparison approach to accommodate minimum and maximum requirements.

A favorite approach of operations research analysts is constrained optimization. Not surprisingly it has been advocated for apportionment. The variables in the problem are $\underset{\sim}{a} = (a_1, \ldots, a_s)$ and the constraints are that $\underset{\sim}{a}$ be nonnegative and integer with $\sum_i a_i = h$. The question that remains is: what function should be optimized?

Ideally one would like to have the a_i "close to" the quotas $q_i = p_i h/p$, where $p = \sum_i p_i$ is the total population. One plausible choice is to minimize $\sum_i |a_i - q_i|$, or perhaps instead $\sum_i (a_i - q_i)^2$. In either case the *method of Hamilton* solves the problem: first, give every state its *lower quota** $\lfloor q_i \rfloor$; second, give the remaining

*$\lfloor z \rfloor$ is the largest integer less than or equal to z.

$\sum_i q_i - \lfloor q_i \rfloor$ seats one each to states having the largest remainders, $q_i - \lfloor q_i \rfloor$.

The "error" inherent in a trial apportionment can, of course, be measured in other ways. $a_i \neq q_i$ means that the average district size in state i, p_i/a_i, is different from the average national district size, p/h. So, perhaps, it might be reasonable to minimize $\sum_i |p_i/a_i - p/h|$ or $\sum_i (p_i/a_i - p/h)^2$. These yield two different methods, neither of which is Hamilton's. Alternatively, and just as reasonably, one might choose to minimize $\sum_i |a_i/p_i - h/p|$ or $\sum_i (a_i/p_i - h/p)^2$, or other variations on the theme.

In 1910 Sainte-Laguë argued -- as did F.W. Owens in 1921 -- that if individuals are considered the basic elements whose shares are to be made as nearly equal as possible, then the error should be measured by $\sum_i p_i (a_i/p_i - h/p)^2$. The method of Webster, it turns out, minimizes this function.

To see this note that

$$\sum p_i (a_i/p_i - h/p)^2 = \sum a_i^2/p_i - (2h/p)\sum a_i + (h^2/p^2)\sum p_i$$
$$= \sum a_i^2/p_i - h^2/p \quad .$$

Thus the constrained optimization problem is equivalent to minimizing $\sum a_i^2/p_i$ when $\sum a_i = h$, $a_i \geq 0$ integer. If $\underset{\sim}{a}$ is optimal then for all a_i, a_j, with $i \neq j$ and $a_i > 0$, a transfer from i to j cannot improve the objective, that is,

$$(a_i-1)^2/p_i + (a_j+1)^2/p_j \geq a_i^2/p_i + a_j^2/p_j$$

which is the same as

$$p_i/(a_i-\tfrac{1}{2}) \geq p_j/(a_j+\tfrac{1}{2}) \quad .$$

Therefore, $\underset{\sim}{a}$ optimal implies

$$\min_{a_i>0} p_i/(a_i-\tfrac{1}{2}) \geqq \max_{a_j\geqq 0} p_j/(a_j+\tfrac{1}{2}) \quad ,$$

which is the min-max inequality that characterizes Webster apportionments.

Conversely, suppose that $\underset{\sim}{a}$ satisfies the Webster min-max inequality or, what is the same thing,

$$(2a_i+1)/p_i \geqq (2a_j-1)/p_j \text{ for all } a_i \geqq 0 \text{ and } a_j > 0 \quad .$$

If $\underset{\sim}{b}$ is some apportionment different from $\underset{\sim}{a}$, let $S^+ = \{i : b_i > a_i\}$, $S^- = \{j : b_j < a_j\}$ and let $b_i = a_i + \delta_i$ for $i \in S^+$, $b_j = a_j - \lambda_j$ for $j \in S^-$. Then $\sum_{S^+} \delta_i = \sum_{S^-} \lambda_j = \alpha > 0$ and, by the above inequalities,

$$(3.2) \qquad (2a_i+\delta_i)/p_i \geqq (2a_j-\lambda_j)/p_j \text{ for all } i \in S^+ \text{ and } j \in S^- \quad .$$

Now we can see that

$$\textstyle\sum_i b_i^2/p_i - \sum_i a_i^2/p_j = \sum_{S^+}(2a_i+\delta_i)\delta_i/p_i - \sum_{S^-}(2a_j-\lambda_j)\lambda_j/p_j \geqq 0$$

since the last term is simply the sum of α differences between the left and right hand sides of (3.2). Therefore $\underset{\sim}{b}$ gives to the objective function a value that can be no smaller than that of $\underset{\sim}{a}$, showing that $\underset{\sim}{a}$ must be a minimizing solution.

If instead the average district sizes are taken as the basic elements to be made as equal as possible, the natural measure of statistical error to minimize would be $\sum_i a_i(p_i/a_i - p/h)^2$. Sainte-Laguë (1910) suggested this possibility and cryptically remarked "one is led to a more complex rule"; this rule turns out to be Hill's method, as Huntington (1928) later showed.

The total error inherent in an apportionment could be small, while the error for some particular state might be unreasonably large. The objective might therefore be formulated in terms of making the worst error for any state as small as possible. There are, again, many different ways of realizing this idea. One such approach, advocated by Burt and Harris (1963) is the objective $\min_{\underset{\sim}{a}} \max_{i,j} |p_i/a_i - p_j/a_j|$. Why not then take instead $\min_{\underset{\sim}{a}} \max_{i,j} |a_i/p_i - a_j/p_j|$? In a slightly different spirit one might choose $\min_{\underset{\sim}{a}} \max_{i} |a_i - q_i|$, or $\min_{\underset{\sim}{a}} \max_{i} |p_i/a_i - p/h|$.

A still different point of view is to consider a state's situation by itself, neither comparing it to another nor to any fixed standard. $\text{Min}_{\underset{\sim}{a}} \max_{i} p_i/a_i$ makes the least advantaged state as advantaged as possible. It is solved by the method of Adams. $\text{Min}_{\underset{\sim}{a}} \max_{i} a_i/p_i$ makes the most advantaged state as little advantaged as possible. It is solved by the method of Jefferson.

The moral of this tale is that one cannot choose objective functions with impunity, despite current practices in applied mathematics. The choice of an objective is, by and large an *ad hoc* affair. The same is true of the other traditional approaches that have been used: Why advocate one divisor d(a) rather than another? Why adopt one measure of pairwise inequality rather than another? Why choose one objective function rather than another? Of much deeper significance than the formulas that are used are the *properties* they enjoy.

Proposition 3.7. (Birkhoff (1976). Hamilton apportionments minimize $\sum |a_i - q_i|$, $\sum (a_i - q_i)^2$ and, actually, any ℓ_p norm of $\underset{\sim}{a} - \underset{\sim}{q}$.

Proposition 3.8. Hill apportionments minimize $\sum a_i(p_i/a_i - p/h)^2$.

Proposition 3.9. The methods defined by $\min_{\tilde{a}} \max_{i,j} |p_i/a_i - p_j/a_j|$ and $\min_{\tilde{a}} \max_{i} |a_i - q_i|$ both admit the Alabama paradox.

Proposition 3.10. Jefferson apportionments solve $\min_{\tilde{a}} \max_{i} a_i/p_i$ and Adams apportionments solve $\min_{\tilde{a}} \max_{i} p_i/a_i$.

Proposition 3.11. The constrained optimization approach can be modified to accommodate both minimum requirements and maximum requirements. The modifications may be made so as to lead to solutions that are consistent with the parallel modifications used for the previous approaches.

4. PRINCIPLES: POPULATION MONOTONICITY

History and common sense have provided the principles we need to sift through the vast number of different numerical apportionment schemes and determine which are appropriate to the problem. A few fundamental principles suffice: consistency with changes in populations, avoiding the Alabama paradox, lack of bias, and staying within the quota. The interplay of these four simple ideas provides a logical framework with which to judge the merits of different methods.

What methods should be seriously considered? The view of the National Academy of Sciences Committees was that "there are only five methods that require consideration at this time" -- namely the five traditional divisor methods that have kept recurring throughout the two hundred year history of the problem and were shown by Huntington to be variations on the single theme of pairwise comparisons. The single most important criterion applied by the Academy Committees to judge between these five methods was bias. Their conclusion was that Hill's method was the least biased. But a careful analysis of historical data shows that this conclusion is wrong: Hill's method is consistently biased toward the small states, while Webster's method is apparently unbiased and is the only one of the five that is so. In other words, a straightforward empirical analysis of an historically important class of methods points to Webster's as the preferred one, and little or no theory is needed to reach this conclusion.

The foregoing argument is simple but limited in scope. What about other methods? In particular, what about the divisor methods, of which the five are but special examples? To study this infinite class with respect to bias requires theoretical models that are treated in the next section. The conclusion, however, is the same: different models of bias all point to Webster's as the only method in the class that is unbiased.

But why should the analysis be restricted to divisor methods? After all, they represent but one computational approach out of many. The reason lies not in their computational attractiveness--many methods, including Hamilton's, could be said to be more attractive computationally. The reason is more fundamental: they are the only methods that are consistent with changing data. This section is devoted to establishing this result.

Of the various parameters affecting apportionment--populations, house size, and number of states, -- the former is constantly in flux, while the last two typically change less frequently. It is essential that a method be consistent with changes in all three of these parameters, and most particularly with changes in populations. If over the short term both the house size h and number of seats s are assumed to be fixed then it suffices to have a *partial method* $M^*(\underset{\sim}{p})$, which gives a set of apportionments of h for every s-vector $\underset{\sim}{p} > \underset{\sim}{0}$. M^* should behave monotonically in populations: roughly speaking, states that increase in size should get more, while those that decrease should get less. Formally this desire can be interpreted in several different ways.

One approach to population monotonicity would be the usual mathematical definition:

(4.1) *If* p_i *increases and all* $p_j (j \neq i)$ *remain the same, then* i's *apportionment does not decrease.*

This notion was proposed as early as 1907 by Erlang and has been studied by Hylland (1975), (1978). The difficulty with this definition is that it is not relevant to the problem in an applied sense, since such comparisons scarcely ever occur in practice. Populations change dynamically, and any useful definition of population monotonicity must reflect this fact.

An alternate definition that seems more appealing at first sight and that takes dynamic changes into account is the following.

(4.2) *If a state's quota increases then its apportionment does not decrease.*

This notion is called *strong population monotonicity.* Unfortunately it is too strong.

Theorem 4.1. *For* $s \geq 3$ *and* $h \neq 0$, $h \neq s$, *no partial method satisfies strong population monotonicity.*

Given M^*, the minimum number of seats a state ever gets over all populations $\underset{\sim}{p}$ is denoted by $\bar{a}$, the maximum number by $\bar{\bar{a}}$. For many (but not all) methods $\bar{a} = 0$ and $\bar{\bar{a}} = h$ (see Proposition 4.2 below).

Proof of Theorem 4.1. Fix $s \geq 3$ and h different from 0 and s and suppose that M^* is strongly population monotone. By homogeneity it suffices to restrict $M^*(\underset{\sim}{p})$ to the set $\bar{P}$ of populations $\underset{\sim}{p}$ whose sum is h, i.e. to the quotas.

If $h = 1$, $(1,0,\ldots,0)$ is an apportionment for $(1/s,1/s,\ldots,1/s)$. It follows that whenever $p_i > 1/s$ then $a_i \geq 1$. But then for any small enough $\varepsilon > 0$ $\underset{\sim}{p} = ((1+\varepsilon)/s,(1+\varepsilon)/s,(1-2\varepsilon)/s,1/s,\ldots,1/s) \in \bar{P}$ and for any apportionment $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$ $a_1 \geq 1$, $a_2 \geq 1$ which implies $h \geq 2$, a contradiction.

Next suppose that $1 < h < s$. Consider any $\underset{\sim}{p} \in \overline{P}$ such that $p_1 > p_2 > \ldots > p_s$ and $h/s < p_i < h/(s-1)$ for $i = 1,\ldots,h-1$ while $p_h < h/s$. Every apportionment $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$ must satisfy $a_1 \geqq a_2 \geqq \ldots \geqq a_s$.

Choose rational $\varepsilon > 0$ small enough such that $p_1 + \varepsilon < h/(s-1)$ and let $\underset{\sim}{p}' = (p_1+\varepsilon,\ldots,p_1+\varepsilon,p_s') \in \overline{P}$ where $p_s' < h/s$. Each of the first $s-1$ states gets at least a_1 seats for any apportionment $\underset{\sim}{b} \in M^*(\underset{\sim}{p}')$. Thus $h = \sum b_i \geqq a_1(s-1)$, which is a contradiction unless $h = s-1$ and $a_1 = 1$. We may conclude that $a_1 = 1$ whenever $\overline{p}_1$ is arbitrarily close to $h/(s-1) = 1$ and $a_s = 0$ whenever $\overline{p}_s$ is arbitrarily close to h/s. But then $(1-(s-1)\varepsilon,(h-1)/(s-1)+\varepsilon,\ldots,$ $(h-1)/s-1+\varepsilon)$ has apportionment $(1,0,\ldots,0)$ which sums to $1 < h$, a contradiction.

Finally consider the case $h > s$. By weak proportionality, $\bar{a} = 0$ or 1, and $\bar{\bar{a}} \geqq h - s + 1 \geqq 2$. Define the sets

$$P_{\bar{a}} = \{p \in (0,h) : a_1 = \bar{a} \text{ whenever } \underset{\sim}{a} \in M^*(\underset{\sim}{p}) \text{ and } p_1 = p\} \quad ,$$

$$P_{\bar{\bar{a}}} = \{p \in (0,h) : a_1 = \bar{\bar{a}} \text{ whenever } \underset{\sim}{a} \in M^*(\underset{\sim}{p}) \text{ and } p_1 = p\} \quad .$$

By definition of $\bar{a}$ there is a $\underset{\sim}{p} \in \overline{P}$ amd $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$ with $a_1 = \bar{a}$ hence $p \in P_{\bar{a}}$ for every $0 < p < p_i$. Moreover for any $p \in P_{\bar{a}}$ and $0 < p' < p$, $p' \in P_{\bar{a}}$. Therefore $P_{\bar{a}}$ is an interval such that $g\ell b(P_{\bar{a}}) = 0$, lub $(P_{\bar{a}}) = \bar{q} > 0$. Similarly $P_{\bar{\bar{a}}}$ is an interval with lub $(P_{\bar{\bar{a}}}) = \bar{\bar{q}} < h$, $g\ell b(P_{\bar{\bar{a}}}) = h$.

Choose rational $\varepsilon > 0$ such that $\varepsilon < \bar{q}$ and $(s-1)\varepsilon + \bar{\bar{q}} < h$. By definition of $\bar{q}$ and $\bar{\bar{q}}$, $\underset{\sim}{p}' = (h-(s-1)\varepsilon,\varepsilon,\varepsilon,\ldots,\varepsilon)$ has the unique apportionment $(\bar{\bar{a}},\bar{a},\bar{a},\ldots,\bar{a})$. Hence $\bar{\bar{a}} + (s-1)\bar{a} = h$. Letting ε approach $\bar{q}$ from below, the inequality $(s-1)\varepsilon + \bar{\bar{q}} \leqq h$ must always be satisfied, since otherwise state 1 would receive fewer than

$\bar{\bar{a}}$ seats while all the others receive $\bar{a}$. Therefore $(s-1)\bar{q} + \bar{\bar{q}} \leq h$. Hence $(\bar{\bar{q}}+\gamma, \bar{q}, \ldots, \bar{q}) \in \bar{P}$ for some $\gamma \geq 0$. Define $\underset{\sim}{p}'' = (\bar{\bar{q}}+\gamma + \frac{(s-2)\delta}{2}, \bar{q} + \frac{(s-2)\delta}{2}, \bar{p}-\delta, \ldots, \bar{q}-\delta) \in \bar{P}$. $\underset{\sim}{p}''$ has apportionment $\underset{\sim}{a} = (\bar{\bar{a}}, a_2, \bar{a}, \ldots, \bar{a})$ where $a_2 > \bar{a}$, a contradiction since then $\sum a_i > h$. □

Proposition 4.1. There exist counterexamples to the theorem when $h = 0$ and $h = s$.

Proposition 4.2. Weak proportionality implies $\bar{a} = 0$ or 1.

A more satisfactory and natural approach to population monotonicity is to consider the *relative changes* in the population of different states, and require that if state i's population increases *relative to* j's then i should not get less seats and j more (unless there is a tie). More exactly, a partial method M^* is *population monotone* if for every two s-vectors $\underset{\sim}{p}, \underset{\sim}{p}' > \underset{\sim}{0}$ and corresponding M^*-apportionments $\underset{\sim}{a}$ and $\underset{\sim}{a}'$, and for all $i < j$,

$$(4.3) \qquad p_i'/p_j' \geq p_i/p_j \text{ implies } \begin{cases} a_i' \geq a_i \text{ or } a_j' \leq a_j, \\ \qquad\text{or} \\ p_i'/p_j' = p_i/p_j \text{ and } a_i', a_j' \text{ can be substituted for } a_i, a_j \text{ in } \underset{\sim}{a}. \end{cases}$$

A partial method M^* is a *partial divisor method* if for some monotone increasing function $d(a)$

$$(4.4) \qquad M^*(\underset{\sim}{p}) = \{\underset{\sim}{a} \geq \underset{\sim}{0} : \textstyle\sum a_i = h, \ \min\limits_{a_i > 0} p_i/d(a_i - 1) \geq \max\limits_{a_i \geq 0} p_i/d(a_i)\}.$$

Note that $d(a)$ may not be a divisor criterion in the strict sense, since it is not assumed to satisfy $a \leq d(a) \leq a+1$.

Theorem 4.2. *Let* $h \geq s \geq 2$, $s \neq 3$. *The partial method* M^* *is a population monotone method if and only if it is a partial divisor method for* (s,h).

Proposition 4.3. Every partial divisor method is population monotone, even when $s = 3$.

The proof of the converse in the case $s = 2$ is relatively simple and intuitive. Fix $h \geq s = 2$. Given a population monotone method M^* we shall show the existence of a monotone increasing function $d(a)$ such that for every $\underset{\sim}{p} = (p_1,p_2) > \underset{\sim}{0}$, $\underset{\sim}{a} = (a_1,a_2) \in M^*(\underset{\sim}{p})$ if and only if $a_1,a_2 \geq 0$, $a_1 + a_2 = h$, and

$$\min_{a_i>0} p_i/d(a_i - 1) \geq \max_{a_i\geq 0} p_i/d(a_i) \quad .$$

Equivalently,

$$(4.5) \qquad d(a_1 - 1)/d(a_2) \leq p_1/p_2 \leq d(a_1)/d(a_2 - 1) \text{ if } a_1,a_2 > 0 \text{ .}$$

$$(4.6) \qquad p_1/p_2 \leq d(0)/d(h-1) \quad \text{if } a_1 = 0, \quad a_2 = h \quad ,$$

and

$$(4.7) \qquad d(h-1)/d(0) \leq p_1/p_2 \quad \text{if } a_1 = h, \quad a_2 = 0 \quad .$$

Let $\overline{P}$ be the set of *normalized* populations $P = \{\underset{\sim}{p} > \underset{\sim}{0}: p_1+p_2=h\}$, and for each $a_1,a_2 \geq 0$ $a_1 + a_2 = h$, let $\overline{P}(\underset{\sim}{a})$ be the set of populations $\underset{\sim}{p} \in \overline{P}$ such that $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$. By population monotonicity, each $\overline{P}(\underset{\sim}{a})$ is an *interval* of the line $\overline{P}$. By weak proportionality, $\underset{\sim}{a} \in \overline{P}(\underset{\sim}{a})$ whenever $\underset{\sim}{a} > \underset{\sim}{0}$. Moreover, since $\underset{\sim}{a}$ is the *unique* apportionment when $\underset{\sim}{p} = \underset{\sim}{a} > \underset{\sim}{0}$, completeness implies that $\underset{\sim}{a}$ is in the *interior* of the interval $\overline{P}(\underset{\sim}{a})$. Completeness also implies that $P(\underset{\sim}{a})$ is a *closed* interval whenever $\underset{\sim}{a} > \underset{\sim}{1}$; the intervals $\overline{P}(1,h-1)$ and $\overline{P}(h-1,1)$ are either closed or half-open; and the $\overline{P}(0,h)$

and $\overline{P}(h,0)$ are either half-open (since zero populations are not admitted) or empty. Finally, the intervals can overlap only at their endpoints, since otherwise population monotonicity would be violated. Thus the situation is like that shown in Figure 4.1 for the case $h = 6$.

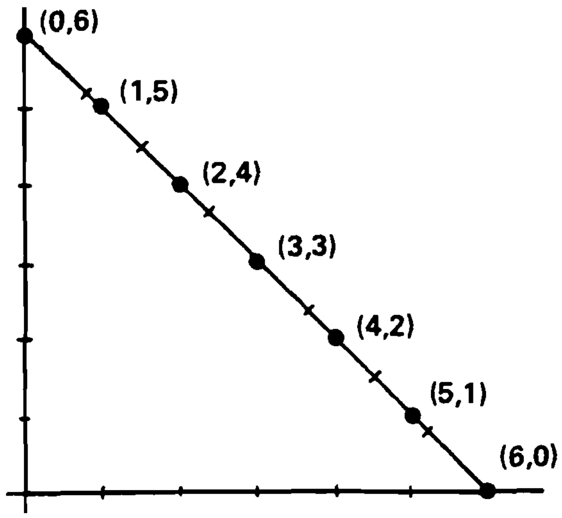


Figure 4.1 Intervals Defining a Population Monotone Partial Method on Two States

To define the divisor criterion $d(a)$, simply let $(d(a_1 - 1), d(a_2))$ be the left-hand endpoint of the interval $\overline{P}(a_1,a_2)$ and $(d(a_1),d(a_2 - 1))$ the right-hand endpoint, for all $a_1 \geqq a_2 > 0$, $a_1 + a_2 = h$. This defines $d(a)$ for $0 \leqq a \leqq h$, and $d(a)$ is evidently monotone increasing in a. In fact $d(a)$ also satisfies $a \leqq d(a) \leqq a + 1$.

For the case $s = 3$ the result does not hold, as will be shown presently by a counterexample.

For the case $s \geqq 4$, several definitions and lemmas are needed. For every $\underset{\sim}{a} \geqq \underset{\sim}{0}$, $\sum a_i = h$, let $P(\underset{\sim}{a}) = \{\underset{\sim}{p} > \underset{\sim}{0} : \underset{\sim}{a} \in M^*(\underset{\sim}{p})\}$.

Lemma 4.1. $P(\underset{\sim}{a})$ *is convex.*

Proof. Let $\underset{\sim}{p}, \underset{\sim}{p}' \in P(\underset{\sim}{a})$ but suppose that $\bar{\underset{\sim}{p}} = \lambda\underset{\sim}{p} + (1-\lambda)\underset{\sim}{p}' \notin P(\underset{\sim}{a})$ for some λ, $0 < \lambda < 1$. Let $\bar{\underset{\sim}{a}} \in M^*(\bar{\underset{\sim}{p}})$ differ from $\underset{\sim}{a}$ in a *minimal* number of coordinates. By hypothesis $\bar{\underset{\sim}{a}} \neq \underset{\sim}{a}$, so choose $i \neq j$ with $a_i < \bar{a}_i$, $a_j > \bar{a}_j$. By population monotonicity $p_i/p_j < \bar{p}_i/\bar{p}_j$; the inequality is strict since otherwise a_i, a_j could be substituted into $\bar{\underset{\sim}{a}}$ for $\bar{a}_i, \bar{a}_j$, contradicting minimality.

Similarly, $p_i'/p_j' < \bar{p}_i/\bar{p}_j$. Thus

$$\lambda p_i < \lambda(\bar{p}_i/\bar{p}_j)p_j \quad ,$$

$$(1-\lambda)p_i' < (1-\lambda)(\bar{p}_i/\bar{p}_j)p_j' \quad ,$$

so

$$\bar{p}_i = \lambda p_i + (1-\lambda)p_i' < \bar{p}_i/\bar{p}_j(\lambda p_j + (1-\lambda)p_j') = \bar{p}_i \quad ,$$

a contradiction. Hence $P(\underset{\sim}{a})$ is convex. □

As before let $\bar{a}$ be the minimum number of seats any state receives and $\bar{\bar{a}}$ the maximum number over all populations $\underset{\sim}{p}$ and $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$.

Lemma 4.2. *There exists* $\underset{\sim}{p}^* > \underset{\sim}{0}$ *such that* $\underset{\sim}{p}^*$ *has an* M^*-*apportionment of form* $(\bar{a}, \bar{\bar{a}}, a_3, \ldots, a_s)$.

Proof. Choose some $\underset{\sim}{p}'$ having an M^*-apportionment of form $\underset{\sim}{a} = (\bar{a}, a_2, \ldots, a_s)$ and $\underset{\sim}{p}''$ having an M^*-apportionment of form $\underset{\sim}{b} = (b_1, \bar{\bar{a}}, b_3, \ldots, b_s)$. Choose α sufficiently large that $\alpha p_2'/p_1' > p_2''/p_j''$ for $1 \leq j \leq s$ and let $\underset{\sim}{p}^* = (p_1', \alpha p_2', \ldots, p_s')$. For every $\underset{\sim}{c} \in M^*(\underset{\sim}{p}^*)$ population monotonicity implies that $c_2 \geq a_2$, $c_1 = \bar{a}$ and $c_j \leq a_j$ for $j \geq 3$. If $c_2 < \bar{\bar{a}}$ then since $\sum_1^s c_i = \sum_1^s b_i$ and $b_1 \geq c_1 = \bar{a}$ there exists $j \geq 3$ with $b_j < c_j$. Thus $b_2 = \bar{\bar{a}} > c_2$ and $b_j < c_j$ but $p_2^*/p_j^* > p_2''/p_j''$, contradicting population monotonicity. Therefore $c_2 = \bar{\bar{a}}$ and $\underset{\sim}{p}^*$ has the desired property. □

This particular $\underset{\sim}{p}^*$ will be used in the proof of the final lemma.

A partial method M^* may admit several different apportionments for a fixed population. When this occurs the subset of states T which receive different numbers of seats in different apportionments are said to be *tied*. Fix $\underset{\sim}{p} > 0$ and for each i let $\bar{a}_i(\underset{\sim}{p}) = \min a_i$ over all $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$.

Lemma 4.3. If M^* is population monotone, then $M^*(\underset{\sim}{p}) =$

$$\{\underset{\sim}{a} \geq \underset{\sim}{0} : \sum a_i = h \text{ and } a_i = \bar{a}_i(\underset{\sim}{p}) \text{ for all } i \notin T(\underset{\sim}{p})$$
$$a_i = \bar{a}_i(\underset{\sim}{p}) \text{ or } \bar{a}_i(\underset{\sim}{p}) + 1 \text{ for all } i \in T(\underset{\sim}{p})\} \quad .$$

Proof. Fix $\underset{\sim}{p} > \underset{\sim}{0}$ and let $T = T(\underset{\sim}{p})$, $\bar{a}_i = \bar{a}_i(\underset{\sim}{p})$ be as defined above. Choose an arbitrary apportionment $\hat{\underset{\sim}{a}} \in M^*(\underset{\sim}{p})$ and fix it for the remainder of the argument. For distinct i, j $\in$ T write $i \to j$ if there is some $\underset{\sim}{a} \in \underset{\sim}{M}^*(p)$ for which $a_i < \hat{a}_i$, $a_j > \hat{a}_j$. By population monotonicity, $a_i + a_j = \hat{a}_i + \hat{a}_j$ and $(\hat{a}_1, \ldots, a_i, \ldots, a_j, \ldots, \hat{a}_s) \in M^*(\underset{\sim}{p})$. That is $i \to j$ means that an alternate apportionment to $\hat{\underset{\sim}{a}}$ can be found by "switching" some number of seats from i to j.

(4.8) *If $i \to j$, $k \to \ell$ and $i \neq \ell$, then $i \to \ell$; moreover in each case the same number of seats are switched.*

If $i = k$ or $j = \ell$ the result is trivial. Otherwise, suppose that $i < j \leq k < \ell$. Since $k \to \ell$ there exists an apportionment of form $\underset{\sim}{b} = (\hat{a}_1, \ldots, \hat{a}_k - \beta, \ldots, \hat{a}_\ell + \beta, \ldots, \hat{a}_s) \in M^*(\underset{\sim}{p})$. Since $i \to j$ there is some $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$ such that $a_i < \hat{a}_i$ and $a_j > \hat{a}_j$. It follows that a_i, a_j may be substituted into $\underset{\sim}{b}$ to obtain $(\hat{a}_1, \ldots, a_i, \ldots, a_j, \ldots, \hat{a}_\ell + \beta, \ldots, a_s) \in M^*(\underset{\sim}{p})$. Thus $i \to \ell$ and $a_i + \hat{a}_\ell + \beta = \hat{a}_i + a_\ell$, so i and ℓ also switch β seats.

Since for every $i \in T$ there must exist some $j \in T$ with $i \to j$ or $j \to i$, T may be partitioned into two classses A and B such that $i \to j$ for every $i \in A$ and $j \in B$. Moreover every switch involves exactly β seats, $\beta \geqq 1$. If $\beta = 1$ the characterization of the lemma follows immediately. The proof is completed by showing that $\beta \geqq 2$ leads to a contradiction. We consider two cases:

<u>Case 1</u> $\hat{a}_i \geqq 1$ for all i.

Let $i \to j$ and consider the s-tuple $\underset{\sim}{c} = (\hat{a}_1, \ldots, \hat{a}_i - 1, \ldots, \hat{a}_j + 1, \ldots, \hat{a}_s) > \underset{\sim}{0}$, where $\hat{a}_i - 1 > \hat{a}_i - \beta \geqq 0$. By hypothesis $\underset{\sim}{c} \geqq \underset{\sim}{0}$, so by weak proportionality $M^*(\underset{\sim}{c}) = \{\underset{\sim}{c}\}$. Now $\underset{\sim}{\hat{a}} \in M^*(\underset{\sim}{p})$ and $(\hat{a}_1, \ldots, \hat{a}_i - \beta, \ldots, \hat{a}_j + \beta, \ldots, \hat{a}_s) \in M^*(\underset{\sim}{p})$, hence by population monotonicity $c_i/c_j = p_i/p_j$. But then $(\hat{a}_1, \ldots, \hat{a}_i - \beta, \ldots, \hat{a}_j + \beta, \ldots, \hat{a}_s)$ is also an M^*-apportionment for $\underset{\sim}{c}$, contradicting weak proportionality.

<u>Case 2</u> $\hat{a}_j = 0$ for some j.

Let p_j be smallest for all j such that $\hat{a}_j = 0$. If $j \notin T$, increase p_j (always staying in the set $P(\underset{\sim}{\hat{a}})$) until at some point $\underset{\sim}{p}'$ j enters the tied class. This must eventually happen because $a_j = 0$ cannot hold when j is the largest state. $M^*(\underset{\sim}{p}) \subsetneq M^*(\underset{\sim}{p}')$ because the populations of all tied states stayed the same. Therefore for some i, $(\hat{a}_1, \ldots, \hat{a}_i - \beta, \ldots, \hat{a}_j + \beta, \ldots, \hat{a}_s) \in M^*(\underset{\sim}{p}')$, where

$\hat{a}_i - 1 > \hat{a}_i - \beta \geq 0$. If $\hat{a}_k \geq 1$ for all $k \neq i,j$, then by weak proportionality $\underset{\sim}{c} = (\hat{a}_1,\ldots,\hat{a}_i-1,\ldots,\hat{a}_j+1,\ldots,\hat{a}_s)$ is the unique M^*-apportionment when $\underset{\sim}{p} = \underset{\sim}{c}$, and a contradiction is obtained by comparing $\underset{\sim}{p}$ and $\underset{\sim}{c}$ as in the preceding case. If $\hat{a}_k = 0$ for some $k \neq j$ then by choice of j, k must also be in T. Hence there exist $\underset{\sim}{a}$, $\underset{\sim}{a}' \in M^*(\underset{\sim}{p}')$ such that $a_j = 0$, $a_k = \beta$, $a_j' = \beta$, $a_k' = 0$. But by weak proportionality there is an M^*-apportionment $\underset{\sim}{a}''$ of form $a_j'' = a_k'' = 1$ (this uses the assumption that $h \geq s$). A contradiction is obtained as before by comparing $\underset{\sim}{p}$ and $\underset{\sim}{a}''$. □

If S is a subset of tied states at $M^*(\underset{\sim}{p})$ and each state in S gets either a_i or a_i+1 seats at $\underset{\sim}{p}$ we say $(\underset{\sim}{p}_S;\underset{\sim}{a}_S)$ is a *tie* and write $t(\underset{\sim}{p}_S;\underset{\sim}{a}_S)$. In particular if $t(p_1,p_2;a_1,a_2)$ in some problem $\underset{\sim}{p}$ then by the preceding lemma $\underset{\sim}{p}$ has M^*-apportionments of form $(a_1+1,a_2,b_3,\ldots,b_s)$ and $(a_1,a_2+1,b_3,\ldots,b_s)$. Define Π to be the set of all pairs (a,b) that occur in an M^*-apportionment for some $\underset{\sim}{p}$.

Lemma 4.4. *If* $(a,b) \in \Pi$ *and* $\bar{\bar{a}} \geq a > 1$, $\bar{\bar{a}} > b \geq 1$, *then there exist* $p',p'',p''' > 0$ *such that* $t(p',p'',p''';\bar{a},a-1,b)$.

Proof. The first step is to find some $\underset{\sim}{p}$ with an apportionment of form $\underset{\sim}{a} = (\bar{a},a,b,c_4,\ldots,c_s)$.

If $\bar{a} = 1$ choose $a_i \geq 1$ for all i and let $\underset{\sim}{p} = \underset{\sim}{a}$.

Suppose that $\bar{a} = 0$. Since $(a,b) \in \Pi$ there is some apportionment $\underset{\sim}{b}$ with $b_i = a > 0$, $b_j = b > 0$. If some $b_k = 0$ $(k \neq i,j)$ then a permutation of $\underset{\sim}{b}$ yields the desired $\underset{\sim}{a}$. Otherwise $b_k \geq 1$ for all k and by weak proportionality there exists an M^*-apportionment of form $\underset{\sim}{a}' = (1,a,b,\ldots) \geq (1,1,\ldots,1)$. Now $(1,1,1,\ldots,h-s+1)$ is an apportionment by weak proportionality, hence $\bar{\bar{a}} \geq h-s+1$. Since $a > 1$ and $b \geq 1$, there exists $k \geq 4$ such that $a_k' < \bar{\bar{a}}$.

Beginning with $\underset{\sim}{p}' = \underset{\sim}{a}' \in P(\underset{\sim}{a}')$, decrease p_1' always staying in $P(\underset{\sim}{a}')$ until a point $\underset{\sim}{p}'' \in P(\underset{\sim}{a}')$ is reached for which $1 \to j$ for some j. Such a point exists because if p_1' were decreased until $p_1'/p_k' < p_1^*/p_k^*$, where $\underset{\sim}{p}^*$ is the vector of Lemma 4.2 then state 1 would get $\bar{a} = 0$ seats. If $j \geqq 4$ then an apportionment of the desired form (0,a,b,...) exists for $\underset{\sim}{p}''$.

If j = 2 or 3, begin at $\underset{\sim}{p}''$ and decrease states 1, 2, 3 proportionally (by a common factor α) until a point $\underset{\sim}{p}''' \in M^*(\underset{\sim}{a}')$ is reached for which $i \to \ell$, $i \leqq 3$, $\ell \geqq 4$. Such a point exists by virtue of Lemma 2 and the fact that $a_k' < \bar{\bar{a}}$. By Lemma 3, $1 \to \ell$, establishing that an apportionment of form (0,a,b,...) exists for $\underset{\sim}{p}'''$.

Let then $\underset{\sim}{a} = (\bar{a},a,b,c_4,\ldots,c_s)$ and choose any $\underset{\sim}{p} \in M^*(\underset{\sim}{a})$. Let T_1 consist of state 2 and all other states (if any) tied with state 2 at $\underset{\sim}{p}$. Beginning at $\underset{\sim}{p}$, decrease all states in T_1 proportionally until at some point $\underset{\sim}{p}^1 \in M^*(\underset{\sim}{a})$, $2 \to j$ for some $j \notin T_1$. At this point $\underset{\sim}{p}^1$ let $T_2 \supsetneqq T_1$ be the class of tied states. Decrease all states in T_2 proportionally until the tied class again increases at point $\underset{\sim}{p}^2$, etc. The process terminates at some $\underset{\sim}{p}^n$, where all states satisfying $a_i < \bar{\bar{a}}$ are tied. In particular, states 1 and 3 are tied at $\underset{\sim}{p}^n$. It follows from Lemma 4.2 that

$$(\bar{a},a,b,c_4,\ldots,c_s),(\bar{a}+1,a-1,b,c_4,\ldots,c_s),(\bar{a},a-1,b+1,c_4,\ldots,c_s) \in M^*(\underset{\sim}{p}^n). \quad \square$$

The idea of the proof of Theorem 4.2 is the following. If in some problem a state having $\bar{a}$ seats is tied with a state having a seats, i.e. if $t(\bar{p},p;\bar{a},a)$, define $d(a) = p/\bar{p}$. d(a) is well-defined because if also $t(\bar{p}',p';\bar{a},a)$ then population monotonicity implies that $p/\bar{p} = p'/\bar{p}'$. It is an easy exercise to show that if d(a) and d(b) are defined then a > b implies d(a) > d(b), i.e. d is monotone increasing.

Now let $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$ and choose any $i \neq j$ for which $\bar{\bar{a}} \geqq a_i > 1$ and $\bar{\bar{a}} > a_j \geqq 1$. By Lemma 4.4 $t(p',p'',p''';\bar{a},a_i-1,a_j)$ for some $p',p'',p''' > 0$. Comparing this with the apportionment $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$, it follows from population monotonicity that $p''/p''' \leqq p_i/p_j$. But $p'''/p' = d(a_j)$ and $p''/p' = d(a_i-1)$ whence $d(a_i-1)/d(a_j) \leqq p_i/p_j$ and

$$(4.10) \qquad p_i/d(a_i-1) \geqq p_j/d(a_j) \quad .$$

It remains to show (4.10) when $a_j = \bar{\bar{a}}$ and/or $a_i = 1$. Now $d(\bar{\bar{a}})$ has not yet been defined, hence set $d(\bar{\bar{a}}) = \infty$ and (4.10) holds. If $a_i = 1$ there are two cases to consider. In case $\bar{a} = 0$ then d(0) is defined and equals 1, so (4.10) says that $p_j/p_i \leqq d(a_j)$, which is an immediate consequence of the definition of $d(a_j)$ and population monotonicity. Otherwise $\bar{a} = 1$ and d(0) is not yet defined. In this case set d(0) = 0 and again (4.10) holds. Thus (4.10) holds in every case. Therefore, since $d(a_i) > d(a_i-1)$ for all i we can write

$$(4.11) \qquad \underset{\sim}{a} \in M^*(\underset{\sim}{p}) \text{ implies } \min_{a_i>0} p_i/d(a_i-1) \geqq \max_{a_i\geqq 0} p_i/d(a_i) \quad .$$

Conversely let $\underset{\sim}{a}$ satisfy (4.11) for some $\underset{\sim}{p}$. Since d(a) is strictly monotone increasing, $\underset{\sim}{p}$ may be wiggled slightly to obtain some $\underset{\sim}{p}'$ such that the min max inequality holds strictly. For such a $\underset{\sim}{p}'$ the only apportionment satisfying the inequality is $\underset{\sim}{a}$, so by (4.11) $\underset{\sim}{a}$ is the unique apportionment for $\underset{\sim}{p}'$. Now construct a sequence of such $\underset{\sim}{p}'$ converging to $\underset{\sim}{p}$ and conclude by completeness that $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$. □

The theorem fails when there are only three states. The reason is that the proof depends on constructing a sufficiently rich collection of 3-way ties, which cannot be done when there is no fourth state to take up the slack. In general, let $e(a,a') = p/p'$ if $t(p,p';a,a')$; thus $e(a,a')$ is the ratio at which a state having a+1 seats would first give up a seat to a state having a' seats as the former decreases and the latter increases in population. If M^* is a partial divisor method then $e(a,a') = d(a)/d(a')$ and the following multiplicative rule must hold among all pairs on which e is defined:

$$(4.12) \qquad e(a',a'')e(a'',a''') = e(a',a''') \quad .$$

Conversely, if (4.12) holds we can define $d(a) = e(\bar{a},a)$ and immediately derive the min max inequality from population monotonicity.

(4.12) can be established by constructing 3-way ties of form $t(p',p'',p''';a',a'',a''')$ as in Lemma 4.4, but for fixed h the construction only works if there is at least one more state to absorb the other seats.

To illustrate what can go wrong when $s = 3$, consider the case $h = 7$. A 3-way tie will involve either one or two seats being shifted around. Hence $t(\underset{\sim}{p};\underset{\sim}{a})$ means that $\sum a_i = 5$ or $\sum a_i = 6$ so there are (up to order) the following ten possibilities for $\underset{\sim}{a}$:

$\sum a_i = 6$	$\sum a_i = 5$
(0,0,6)	(0,0,5)
(0,1,5)	(0,1,4)
(0,2,4)	(0,2,3)
(0,3,3)	(1,1,3)
(1,2,3)	(1,2,2)

Each of these ten triples produces one dependency in the variables e(a,b) of the form (4.12). Also we must have e(b,a) = 1/e(a,b) and e(a,a) = 1. Now choose e(a,0) = 2a+1 as in Webster's method. The ten dependencies are redundant and only determine four additional values:

$$e(1,4) = e(1,0)e(0,4) = 1/3$$
$$e(1,5) = e(1,0)e(0,5) = 3/11$$
$$e(2,3) = e(2,0)e(0,3) = 5/7$$
$$e(2,4) = e(2,0)e(0,4) = 5/9$$

There remain e(1,2) and e(1,3), which are related by the expression e(1,3) = e(1,2)e(2,3) = 5e(1,2)/7. To be a divisor method e(1,2) must also satisfy e(1,2) = e(1,0) e(0,2) = 3/5. But for $h = 7$, the triple (0,1,2) does not occur as a tie, so no such dependency is imposed. In fact if e(1,2) is chosen close to but not equal to 3/5 then M^* will be population monotone but not a divisor method. An example with e(1,2) = 4/5 is shown in Figure 4.1.

Proposition 4.4. For $s = 3$ and $3 \leq h \leq 6$, every population monotone partial method is a partial divisor method.

Proposition 4.5 If $1 < h < s$, define M^* such that the largest state always gets h seats and the rest get zero seats. M^* is population monotone but technically is not a partial divisor method, because it cannot be represented by a monotone increasing d(a).

Proposition 4.6 Theorem 4.2 also holds in the case of arbitrary minimum requirements $\underset{\sim}{r}$ and $h \geqq \max (s, \sum r_i)$ $s \neq 3$. Modify the proof by first defining $\bar{a}_i$ and $\bar{\bar{a}}_i$ to be the minimum and maximum apportionments ever received by state i among all $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$, and then proving the analog of Lemma 4.3. To establish the analog of Lemma 4.4 suppose without loss of generality that state 1 has the smallest $\bar{a}_i$, and state 2 the next smallest. Show that if $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$ and $\bar{\bar{a}}_i \geqq a_i > \max \{\bar{a}_i, 1\}$, $\bar{\bar{a}}_j > a_j \geqq \max \{\bar{a}_j, 1\}$, $i, j \geqq 2$, then a 3-way tie of form $t(p', p'', p'''; \bar{a}_1, a_i - 1, a_j)$ can be constructed. If $t(\bar{p}, p; \bar{a}_1, a)$ define $d(a) = p/\bar{p}$. This defines d(a) for all relevant $a, \bar{\bar{a}} > a \geqq \bar{a}_2$. For $\bar{a}_1 < a < \bar{a}_2$ construct a tie of form $t(p, p'; a, \bar{a}_2)$ and define $d(a) = (p/p')d(\bar{a}_2)$. Finally, let $d(\bar{\bar{a}}) = \infty$ and $d(0) = 0$ if $\bar{a}_1 > r_1 = 0$. Now check the well-definition of d(a) and show that the min max inequality follows from population monotonicity.

Proposition 4.7 For given minimum *and* maximum requirements $r \leqq r^+$ Theorem 4.2 may fail even when $s \geqq 4$.

Suppose that M is a method such that every restriction to some fixed s and h is population monotone. By Theorem 4.2, every such restriction is a partial divisor method; however the associated functions d(a) may depend on s and h. For example, M might be Adams' method when h-s is odd and Jefferson's when h-s is even. Of course, such a method is ridiculous, since it does not give

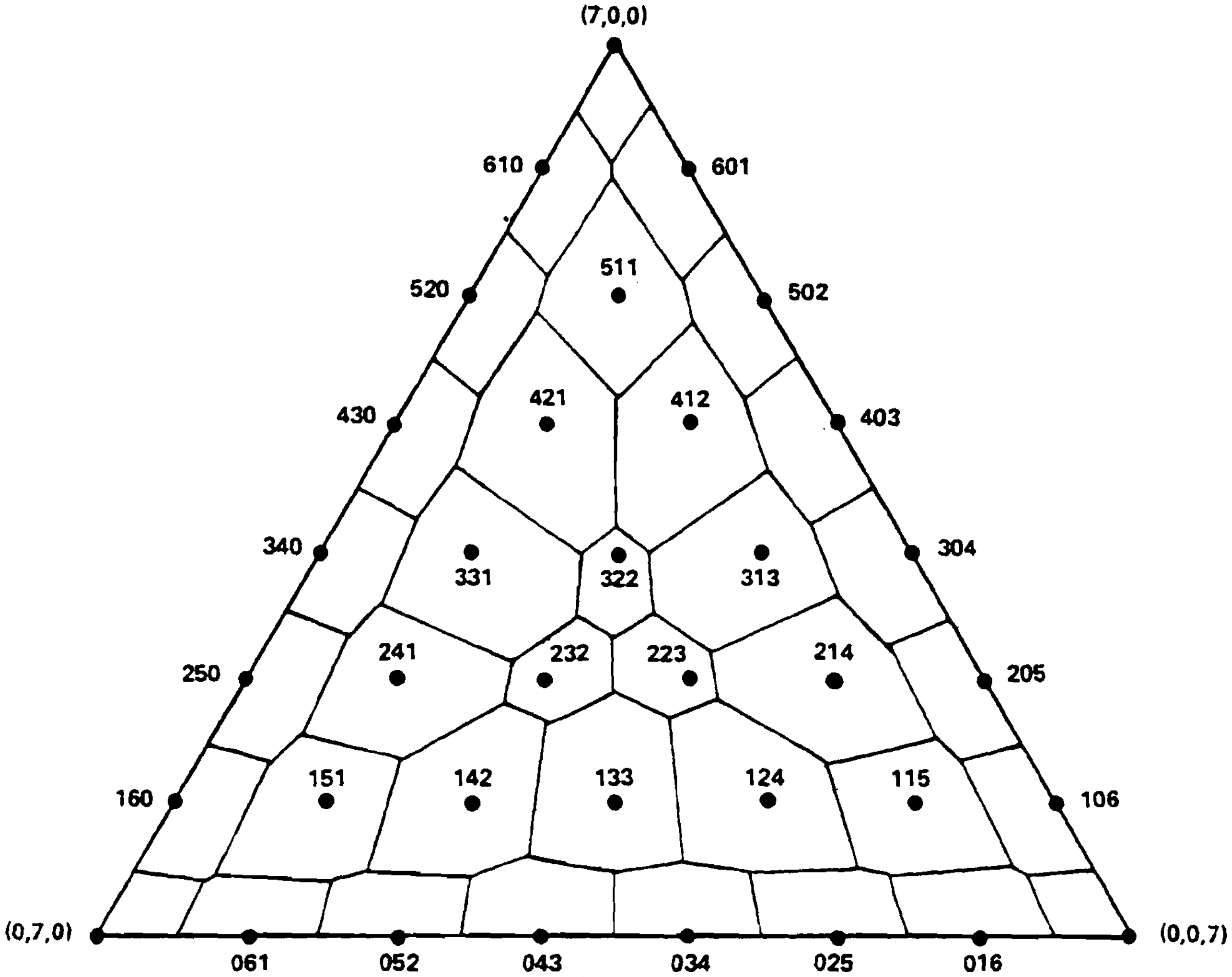


Figure 4.2 A Population Monotone Partial Method on Three States that is Not a Divisor Method

consistent results when the house size or number of states changes. Such absurdities are avoided by extending the concept of population monotonicity to allow comparisons between *any* two problems, including ones with different s and h.

A method M is *population monotone* if for any two vectors $\underset{\sim}{p}$, $\underset{\sim}{p}' > \underset{\sim}{0}$ and $\underset{\sim}{a} \in M(\underset{\sim}{p},h)$, $\underset{\sim}{a}' \in M(\underset{\sim}{p}',h')$

$$(4.13)\quad p'_{i'}/p'_{j'} \geqq p_i/p_j \text{ implies } \begin{cases} a'_{i'} \geqq a_i \text{ or } a'_{j'} \leqq a_j, \\ \text{or} \\ p'_{i'}/p'_{j'} = p_i/p_j \text{ and } a'_{i'},\ a'_{j'}, \\ \text{can be substituted for } a_i, a_j \text{ in } \underset{\sim}{a}. \end{cases}$$

<u>*Theorem 4.3.*</u> *A method is population monotone if and only if it is a divisor method.*

<u>*Corollary 4.3.1*</u> *If a method is population monotone then it is house monotone.*

Theorem 4.3 is proved by showing that a single divisor criterion applies to all 2-state problems (use Theorem 4.2) and then showing that the same divisor applies to all problems, including those with s = 3. The theorem still holds when (4.13) is modified to allow for minimum and maximum requirements.

Population monotonicity says that as the conditions of a problem change -- as populations shift, as the size of house expands or contracts, and states join or secede -- apportionments should respond accordingly, i.e., they should not move contrary to the relative changes of states' populations. This is an elementary requirement for any scheme of fair representation, and the only methods that satisfy it are the divisor methods.

It is significant that other fairness ideas follow automatically. Any method that is population monotone avoids the Alabama paradox. Another fairness concept implied by population monotonicity is "uniformity", which says (roughly) that any apportionment should still be valid when restricted to any subgroup of states. This concept is treated in more detail in Section 8.

5. BIAS

The decisive conclusion of the preceding section is that the only realistic candidates for methods of apportionment are the divisor methods. The question then becomes: which of the infinite number of divisor methods should be chosen?

Historical precedent -- both the debate of 1832 and the twentieth century controversy over the methods of Hill and Webster -- is that a method is unacceptable if it has a persistent bias either in favor of the large or of the small states. There are several ways to measure "bias", and there are different probabilistic models by which a tendency toward bias can be revealed theoretically. While no one particular definition or model provides absolute proof that a method is biased or not, the weight of the evidence over a variety of definitions and models is persuasive that Webster's is the only divisor method that is unbiased. Analysis of the historical statistics confirms the theory with experimental fact.

Even a casual inspection of the five "traditional" methods applied to examples shows that in the order Adams (A), Dean (D), Hill (H), Webster (W), and Jefferson (J), they tend increasingly to favor the larger states. This is revealed in Example 1.1 as well as in the more convincing 19 examples of U.S. history. This observation may be made precise. A method M' *favors small states relative to* M if, for every M-apportionment $\underset{\sim}{a}$ and M'-apportionment $\underset{\sim}{a}'$ for $\underset{\sim}{p}$ and h,

$$p_i < p_j \text{ implies either } a_i' \geqq a_i \text{ or } a_j' \leqq a_j \;.$$

Theorem 5.1. *If* M *and* M' *are divisor methods with divisor criteria* d(a) *and* d'(a) *satisfying* $d'(a)/d'(b) > d(a)/d(b)$ *for all integers* $a > b \geq 0$, *then* M' *favors small states relative to* M.

Proof of Theorem 5.1. Suppose by way of contradiction, that for some $\underset{\sim}{a} \in M(\underset{\sim}{p},h)$ and $\underset{\sim}{a}' \in M'(\underset{\sim}{p},h)$, $p_i < p_j$, $a_i' < a_i$, and $a_j' > a_j$. By population monotonicity $a_i' < a_i \leq a_j < a_j'$, so $a_j' - 1 > a_i' \geq 0$ and $d(a_j' - 1) \geq 1$ because $a \leq d(a) \leq a+1$ for all a.

The min-max inequality for $\underset{\sim}{a}'$ implies $p_j/d'(a_j' - 1) \geq p_i/d'(a_i')$ from which, by the preceding, it follows that $d'(a_i') > 0$. Hence,

$$p_j/p_i \geq d'(a_j' - 1)/d'(a_i') > d(a_j' - 1)/d(a_i') \geq d(a_j)/d(a_i - 1) ,$$

the last by monotonicity of d(a). Thus, $p_j/d(a_j) > p_i/d(a_i - 1)$, contradicting the min-max inequality for $\underset{\sim}{a}$. □

Write M' > M if M' favors small states relative to M.

Proposition 5.1. A > D > H > W > J .

Bias has, however, an absolute as well as a relative meaning. Any apportionment that gives a_1 and a_2 seats respectively to states having populations $p_1 > p_2 > 0$ *favors the larger state over the smaller state* if $a_1/p_1 > a_2/p_2$ and *favors the smaller state over the larger state* if $a_1/p_1 < a_2/p_2$. It may be asked whether over many pairs of states a method tends more often to favor the larger over the smaller state or *vice versa*.

Different meanings may be attached to "many pairs". One simple approach is to consider a pair of populations $p_1 > p_2 > 0$.

Two states having these populations could divide any number of seats h between them and, since we are considering only divisor methods, the way in which they would share h seats is determined independently of any other states that may be part of the same problem. Since p_1 and p_2 are rational (or integer) there is a smallest "perfect" house size h^* where both states have integer quotas. A reasonable idea then is to count, for any method M and pair of populations (p_1,p_2), the number $S(p_1,p_2)$ of apportionments favoring the smaller state and the number $L(p_1,p_2)$ of apportionments favoring the larger state, over all M-apportionments (a_1,a_2) such that $a_1 + a_2 \leqq h^*$. The method M is *pairwise unbiased on populations* if for every pair of populations (p_1,p_2), $L(p_1,p_2) = S(p_1,p_2)$.

Proposition 5.2. Webster's is the only divisor method that is pairwise unbiased on populations.

This approach to defining and measuring bias has the merit that no assumption need be made about the distribution of populations: *whatever* the pair of populations, Webster's is the unique unbiased divisor method. However, as an empirical test the approach has limitations, for the number of seats shared by a pair of states will typically be much smaller than h^*.

A more realistic approach is to consider a pair of integer apportionments $a_1 > a_2 > 0$ and ask: if the populations (p_1,p_2) have the M-apportionment (a_1,a_2) how likely is it that the small state (state 2) is favored? Note that the population monotonicity of M guarantees $p_1 \geqq p_2$ since $a_1 > a_2$.

Take as a probabilistic model that the populations $(p_1,p_2) = \underset{\sim}{p} > \underset{\sim}{0}$ are uniformly distributed in the positive quadrant. Given integers $\underset{\sim}{a} = (a_1,a_2) > \underset{\sim}{0}$ and a method M, the set of $\underset{\sim}{p}$'s for which $\underset{\sim}{a}$ is an M-apportionment is unbounded. To define a bounded subset of the sample space in a natural way we use the fact that M is a divisor method.

Choose any $x > 0$ representing a hypothetical district size, let $\underset{\sim}{a} > \underset{\sim}{0}$ be an apportionment, and define $R_x(\underset{\sim}{a})$ to be the set of all populations $\underset{\sim}{p} > \underset{\sim}{0}$ which yield the M-apportionment $\underset{\sim}{a}$ using the divisor x:

$$R_x(\underset{\sim}{a}) = \{\underset{\sim}{p} > \underset{\sim}{0} : d(a_i) \geqq p_i/x \geqq d(a_i-1)\}$$

where by convention $d(-1) = 0$. Each region $R_x(\underset{\sim}{a})$ is a rectangle containing the point $\underset{\sim}{a}$ and having sides of length $d(a_1) - d(a_1-1)$ and $d(a_2) - d(a_2-1)$. Figure 5.1 shows these for Dean's method, and Figure 5.2 for Webster's ($x = 1$ in both cases).

Figure 5.3 shows them for the somewhat bizarre divisor method that is defined as follows:

$$d(2a) = 2a + 1/5, \quad d(2a+1) = 2a + 9/5 .$$

A divisor method M is *pairwise unbiased on apportionments* if for *every* $a_1 > a_2 > 0$ and every $x > 0$ the probability that state 1 is favored over state 2 equals the probability that state 2 is favored over state 1, given that $(p_1,p_2) \in R_x(\underset{\sim}{a})$. Note that the probability that one state is favored over the other is independent of the choice of x, so that if a method satisfies the definition for some x then it does for all x.

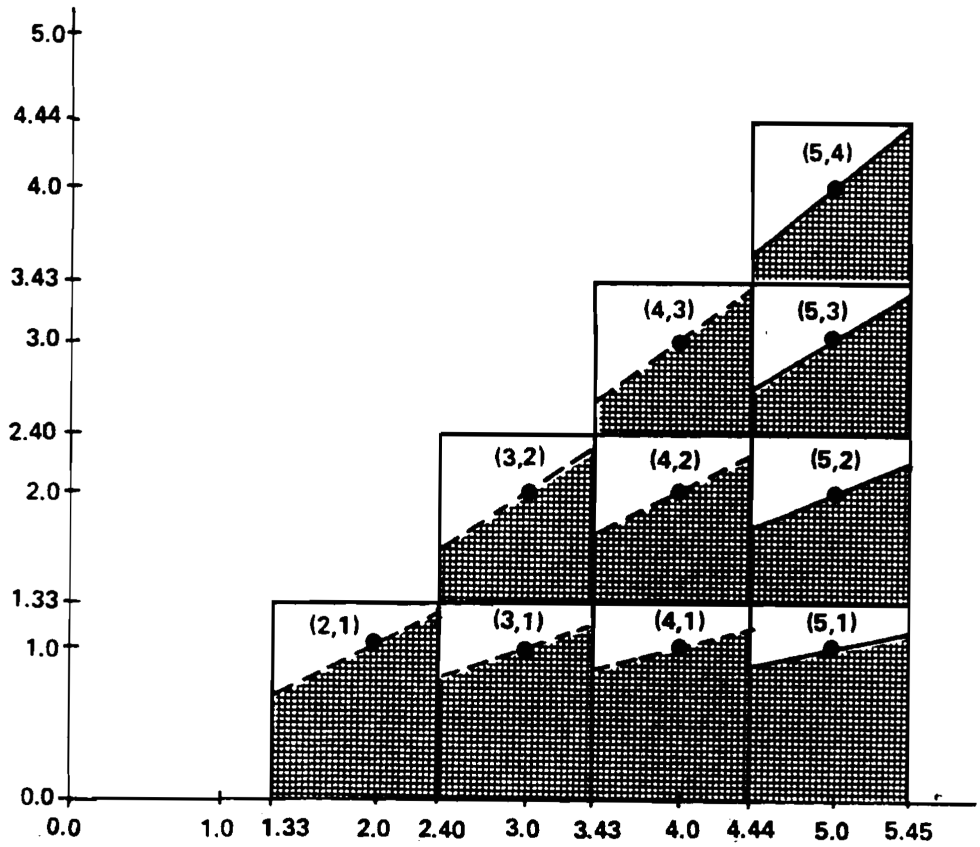


Figure 5.1 Regions Favoring Small and Large States -- Dean's Method

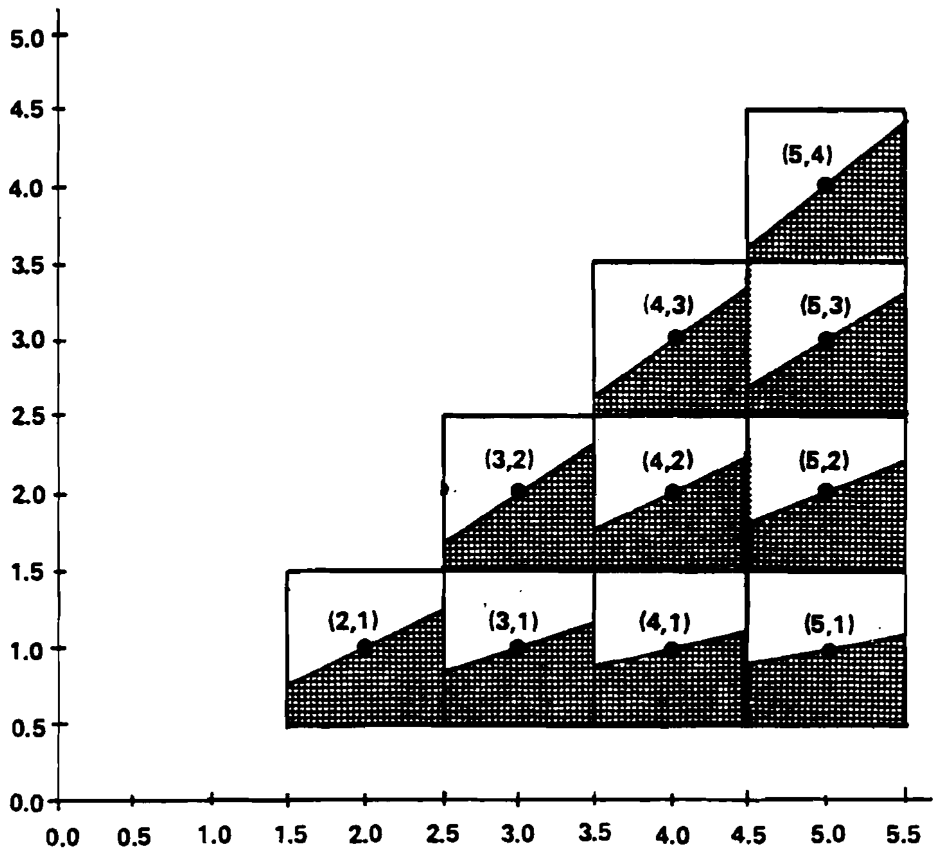
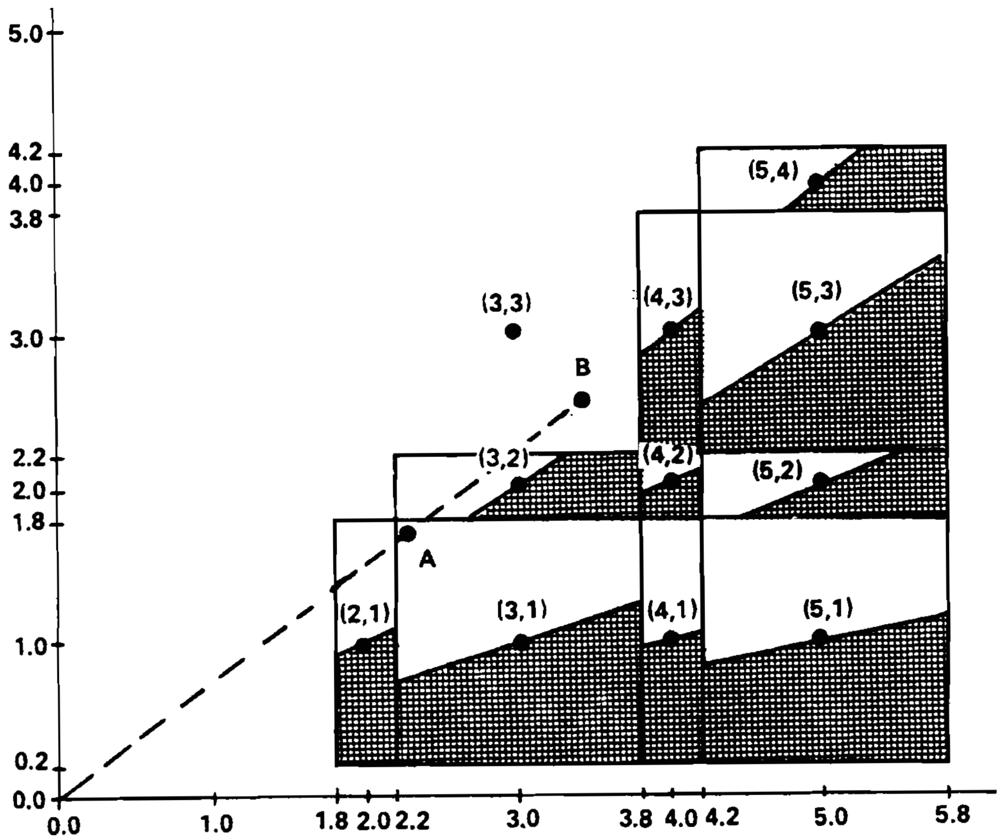


Figure 5.2 Regions Favoring Small and Large States -- Webster's Method



<u>Figure 5.3</u> Regions Favoring Small and Large States -- A Non-Proportional Method

For each pair $a_1 > a_2 > 0$ the shaded area in the figures show those populations that favor the smaller state. It is quite evident that Dean's method is systematically biased toward smaller states, whereas both Webster's method and the "1/5 - 4/5" method are unbiased. However the latter method is a strange one; indeed it is not proportional. For example if the populations are in the ratio represented by the ray OAB, they will split 6 seats as 3 and 3, but 4 seats as 3 and 1. Such methods are not reasonable, since they do not adhere to the ideal of allocating seats in proportion to numbers as near as may be.

Proposition 5.3. The theoretical model developed above may be used to compute the probability that the small state is favored for each of the five traditional methods, for any given apportionment (e.g., (4.2)).

Theorem 5.2. Webster's is the unique proportional divisor method that is pairwise unbiased on apportionments.

Proof of Theorem 5.2. Given a divisor method M and $a_1 > a_2 > 0$, the ray defined by $a_1/p_1 = a_2/p_2$ divides $R_x(\underset{\sim}{a})$ into the populations favoring state 1 and those favoring state 2. These two sets have equal measure, i.e., the ray *bisects* the rectangle, if and only if it passes through its center, $\underset{\sim}{c} = (c_1, c_2)$. Since

$$c_1 = x(d(a_1-1) + d(a_1))/2 \quad \text{and} \quad c_2 = x(d(a_2-1) + d(a_2))/2 \;,$$

M is pairwise unbiased if and only if it satisfies the following condition:

$$(5.1) \qquad \frac{d(a_1-1) + d(a_1)}{d(a_2-1) + d(a_2)} = \frac{a_1}{a_2} \quad \text{for all } a_1 > a_2 > 0 \;.$$

In particular, W is pairwise unbiased on apportionments. Conversely, suppose that M satisfies (5.1). Set $a_1 = a \geq 2$, $a_2 = 1$. Then

$$(d(a-1) + d(a))/a = d(0) + d(1)$$

and, since $a \leq d(a) \leq a+1$ for all a, letting $a \to \infty$ implies

$$d(0) + d(1) = 2 \quad .$$

From the two preceding relations it follows that

(5.2) $\quad d(2a) = 2a + d(0)$ and $d(2a+1) = 2a + 2 - d(0)$.

Proposition 5.4. Proportionality implies that

(5.3) $\quad d(b)/d(b-1) \geq d(b+1)/d(b)$ for all $b > 0$.

Substituting $b = 2a+1$ and using (5.2) gives

$$\frac{2a+2-d(0)}{2a+d(0)} \geq \frac{2a+2+d(0)}{2a+2-d(0)}$$

or $4a+4 \geq (8a+6)d(0)$. Letting $a \to \infty$ shows that $d(0) \leq 1/2$. Substituting instead $b = 2a$ into (5.3) we find $(8a+2)d(0) \geq 4a$ and so $d(0) \geq 1/2$. Therefore $d(0) = 1/2$ and we see that M is Webster's method. □

Other probabilistic models of pairwise bias point to the same conclusion. For example, one might assume that all *normalized populations* $p_1 + p_2 = 1$ are equally likely. Given a divisor method M, the populations for which some particular M-apportionment

(a_1,a_2), $a_1 > a_2 > 0$, is the solution for h can be represented as a line segment. It is convenient to renormalize so that $p_1 + p_2 = h$; then (a_1,a_2) itself is contained in the segment, and the smaller state is favored in the part of segment lying to the right of (a_1,a_2). Thus in this model M will be pairwise unbiased on apportionments if, for every $a_1 > a_2 > 0$, (a_1,a_2) is the midpoint of its segment. This is manifestly true for Webster, and indeed it may be shown that *Webster's is the unique divisor method that is pairwise unbiased on apportionments in this model.*

The proof follows by observing that $(a_1,a_2) > \underset{\sim}{0}$ is the midpoint of its segment if and only if

$$\frac{a_1}{a_1 + a_2} = \left[\frac{d(a_1-1)}{d(a_1-1) + d(a_2)}\right] + \left[\frac{d(a_1)}{d(a_2-1) + d(a_1)}\right] \Big/ 2 .$$

With $a_1 = a$, $a_2 = a+1$ we obtain the recursion $d(a+1) = (2a+3)d(a-1)/(2a-1)$, from which

$$d(2a) = (4a+1)d(0), \quad \text{and} \quad d(2a+1) = (4a/3+1)d(1) \quad .$$

Now use $2a \leqq d(2a) \leqq 2a+1$ and let $a \to \infty$ to conclude that $d(0) = 1/2$. Similarly conclude that $d(1) = 3/2$. Hence d(a) is the Webster divisor criterion. Note that only weak proportionality was needed to obtain this result.

The preceding theoretical results suggest the following empirical test for bias. Given a divisor method M, let $\mathscr{P}$ be a collection of 2-state *sample problems* $(p_1,p_2;a_1,a_2)$, where $a_1 > a_2 > 0$ and (a_1,a_2) is an M-apportionment for (p_1,p_2). One way of obtaining such a sample is to select from a larger problem

$(p_1,\ldots,p_s;a_1,\ldots,a_s)$ all pairs i,j such that $a_i > a_j > 0$ and consider each such $(p_i,p_j;a_i,a_j)$ as a 2-state problem. This exploits the fact that divisor methods are "uniform", i.e. the way in which a divisor method shares $a_i + a_j$ seats between two states is determined independently of the populations of the other states (see Section 8 for a discussion of uniformity). The *bias ratio* is the percentage of pairs $(p;a) \in \mathcal{P}$ for which the small state is favored. One would expect that for sufficiently large samples the bias ratio for Webster is close to 50%, whereas the ratio for other proportional methods is significantly different from 50%.

Analysis of United States historical data fulfills this expectation. However, the Congressional apportionment problem requires that a minimum of one seat be given to each state no matter how small its population. This requirement introduces an additional built-in favoritism toward the small states. In order to analyze the actual apportionments that would have resulted by different methods applied to the nineteen historical problems, and yet eliminate this "artificial" bias toward the small, each method was applied with the minimum requirement of 1, but the counts have been made after leaving out all states having a quota less than .5. The *bias ratio over U.S. history* for any method M was computed by counting, for each census year, the number of pairs of states in which the smaller state would have been favored by this method and dividing by the total number of pairs over all nineteen problems. The results are given in Table 5.1.

	Adams	Dean	Hill	Webster	Jefferson
Bias ratio	77.2%	56.6%	54.4%	51.3%	24.7%

Table 5.1 Bias ratio over U.S. history.

The theoretical bias results also hold in the presence of minimum and maximum requirements. It is natural to ask only that a method be unbiased whenever the requirements are not binding. If, for example, $\underset{\sim}{r} = (r_1,\ldots,r_s)$ is a set of minimum requirements for states with populations $\underset{\sim}{p} = (p_1,\ldots,p_s)$, then state i is *free* at h using the method M if $a_i > r_i$ for *some* $\underset{\sim}{a} \in M(\underset{\sim}{p},h)$, and otherwise is *bound*. This definition parallels the previous analysis where $\underset{\sim}{r} = \underset{\sim}{0}$ and the arguments were confined to $\underset{\sim}{a} > \underset{\sim}{0}$. The theorems and proofs go through with little or no adjustment.

Problems involving more than two states can also be analyzed for bias by comparing how the larger states and smaller states fare as *groups*. Given a possible apportionment $\underset{\sim}{a} = (a_1,\ldots,a_s) > \underset{\sim}{0}$ for populations $\underset{\sim}{p} = (p_1,\ldots,p_s)$, let L and S be any two disjoint sets of states such that $a_i > a_j$ for every $i \in L$ and $j \in S$. L is a set of *larger states* and S a set of *smaller states*. The apportionment $\underset{\sim}{a} > \underset{\sim}{0}$ *favors the smaller states* if $\sum_S a_i/\sum_S p_i > \sum_L a_j/\sum_L p_j$ and *favors the larger states* if $\sum_L a_j/\sum_L p_j > \sum_S a_i/\sum_S p_i$.

A divisor method M is *unbiased* if for every $\underset{\sim}{a} > \underset{\sim}{0}$ and divisor $x > 0$, and for any disjoint sets of larger and smaller states L and S, the probability given $\underset{\sim}{p} \in R_x(\underset{\sim}{a})$ that $(\underset{\sim}{p};\underset{\sim}{a})$ favors the smaller states equals the probability that it favors the larger states. Note that the condition holds for some x if and only if it holds for all x.

Theorem 5.3. *Webster's is the unique unbiased proportional divisor method.*

Proof of Theorem 5.3. Let M be a proportional divisor method. Given $\underset{\sim}{a} > \underset{\sim}{0}$ and $x > 0$, let L and S be any disjoint sets of larger and smaller states respectively. The hyperplane $\sum_L a_i/\sum_L p_i = \sum_S a_i/\sum_S p_i$ divides $R_x(\underset{\sim}{a})$ into the populations favoring L and those favoring S. These two sets have equal measure, i.e., the hyperplane bisects the rectangular solid $R_x(\underset{\sim}{a})$, if and only if it passes through the center of $R_x(\underset{\sim}{a})$. The center $\underset{\sim}{c}$ has coordinates $c_i = [d(a_i-1) + d(a_i)]/2$; hence M is unbiased if and only if for all $\underset{\sim}{a} > \underset{\sim}{0}$

$$(5.4) \qquad \sum_L a_i/(\sum_L d(a_i-1) + d(a_i)) = \sum_S a_i/(\sum_S d(a_i-1) + d(a_i)) \quad .$$

Conversely, if M satisfies (5.4) for all $\underset{\sim}{a} > \underset{\sim}{0}$, then, in particular, it satisfies it whenever $\underset{\sim}{a} = (a+1,a)$, $a > 0$. From this it may be concluded as in the proof of Theorem 5.2 that M is Webster's method. □

Note that if it is desired to argue with the number of states s *fixed*, $s \geqq 3$ then let $t = \lceil \frac{s}{2} \rceil + 1$ and consider instead the apportionment $a_1 = a_2 \ldots = a_t = a+1$, $a_{t+1} = \ldots = a_s = a \, (a>0)$. Then (5.4) becomes

$$(a+1)/(d(a) + d(a+1)) = a/(d(a-1) + d(a))$$

from which it follows as in the proof of Theorem 5.2 that M must be Webster's method.

This model of bias can be used to provide further convincing empirical evidence that Webster's is the preferred method. A test for a method M is to take the nineteen cases of United States history (p;a), where a is an M-apportionment for p, including minimum requirements of 1, and for each such problem eliminate states with quota less than .5. Divide the remaining states into three approximately equal classes: large (L), middle, and small (S), where the middle class takes up the extras if the number of remaining states is not divisible by three. Counting the number of problems in which the small states are favored one finds: Adams 19 (always), Dean 14, Hill 12, Webster 7, Jefferson 0 (never).

Considerably more insight is obtained by calculating, for each apportionment, the percentage difference β between $k_S = \sum_S a_i / \sum_S p_i$ and $k_L = \sum_L a_j / \sum_L p_j$, that is, $\beta = 100(k_S - k_L)/\min(k_S, k_L)$. β is the *percentage bias* of the apportionment: if positive the small states are favored, if negative the large states are favored. Contrasting the apportionments by each of the five traditional methods over the nineteen U.S. censuses, one finds that in *every* case the percentage bias of Webster's method is smaller in absolute value than that of Hill and, indeed, than that of any of the other methods. The average of the percentage biases of each method over the nineteen problems is given in Table 5.2. A graph of the cumulative average bias of each method may be found in Chapter 3 (Figure 5).

	Adams	Dean	Hill	Webster	Jefferson
Average bias	18.31%	5.19%	3.38%	0.26%	-18.78%

Table 5.2. Average Percentage Bias over 19 U.S. Problems

This model is also convenient for estimating the probability of bias when the number of states s is large (e.g., s = 50). Choose $\underset{\sim}{a} > \underset{\sim}{0}$, let L and S be classes of larger and smaller states respectively, and define $a_L = \sum_L a_j$ and $a_S = \sum_S a_i$. Fix $x > 0$ and let $R = R_x(\underset{\sim}{a})$. The quotients $q_i = p_i/x$ are distributed independently and uniformly in the intervals $d(a_i-1) \leq q_i \leq d(a_i)$. The mean m_i and variance σ_i of q_i are therefore

$$m_i = (d(a_i) + d(a_i-1))/2 \quad \text{and} \quad \sigma_i = (d(a_i) - d(a_i-1))^2/12 \ .$$

Define the random variable $X = \sum_S q_i - (a_S/a_L)\sum_L q_j$, where $\underset{\sim}{q} \in R$. X is a sum of independent random variables so its mean and variance are given by

$$m = \sum_S m_i - (a_S/a_L)\sum_L m_j \quad \text{and} \quad \sigma^2 = \sum_S \sigma_i^2 + (a_S/a_L)^2 \sum_L \sigma_j^2 \ .$$

If the number of states in the set $L \cup S$ is large, the central limit theorem implies that $\overline{X} = (X-m)/\sigma$ is approximately normally distributed with mean 0 and variance 1. So the probability that the small states are favored is $Pr(X \leq 0) = Pr(\overline{X} \leq -m/\sigma)$, which is estimated using standard tables of the normal distribution.

To illustrate, take the *actual* 1970 U.S. apportionment (see Hill apportionment, 1970). The one-third (i.e., 16) largest states have a total of $a_L = 300$ seats, and the one-third (16) smallest states have $a_S = 27$ seats. $|L \cup S| = 32$, a reasonably large sample for using the normal distribution as an approximation. To conform with our previous historical analysis we assume no quotient can go below .5. For Hill's method we obtain: $\sum_S m_i = 26.0861$, $\sum_L m_j = 299.8650$, $\sum_S \sigma_i^2 = 1.3075$, $\sum_L \sigma_j^2 = 1.3342$ so $m = -2.4017$, $\sigma^2 = 1.2565$, and $m/\sigma = -.7886$. Therefore $Pr(X \leq 0) = Pr(\overline{X} \leq .7886) = 78.48\%$. Thus, the odds that Hill favors the small is almost 4 to 1. The results for the five traditional methods are given in Table 5.3.

Adams	Dean	Hill	Webster	Jefferson
100.0%	93.9%	78.5%	50.0%	0.0%

Table 5.3. Estimated probability of favoring the small states for each of the five methods on the basis of the actual 1970 apportionment.

A similar calculation gives the expected percentage bias for each method. Let $X = \sum_S q_i/a_S$ and $Y = \sum_L q_i/a_L$. Then $E(X/Y) \approx E(X)/E(Y) + Var(Y)E(X)/E(Y)^3$, the approximation coming from the first terms of a Taylor expansion. E(X/Y) represents the expected ratio by which the large states are overrepresented relative to the small. So, the reciprocal minus 1 gives an estimate of β. For Hill's method, using the previous calculations, we obtain:

$$\begin{aligned} E(X/Y) &= (26.0861/27)/(299.8650/300) \\ &+ (1.3342/300^2)(26.0861/27)/(299.8650/300)^3 \\ &= .9666 \quad , \end{aligned}$$

so $\beta = -1 + 1/.9666 = .0346$. Therefore the expected percentage bias in favor of the small by the method of Hill, given the 1970 apportionment, is 3.46%. The results for the five traditional methods are given in Table 5.4

Adams	Dean	Hill	Webster	Jefferson
28.2%	7.0%	3.5%	0.0%	-20.8%

Table 5.4 Expected Percentage Bias of Five Methods on the basis of the actual 1970 apportionment.

These expected figures are in remarkably close agreement with the actual historical averages of Table 5.2. This is especially so since the statistics of U.S. history contain rather peculiar dependencies which have to do with the way in which new states were admitted to the Union and the choice of house size that was made. Had there always been 50 states and 435 seats, one would expect the historical bias percentages to be even closer to those of Table 5.4.

Alternative probabilistic models for the joint distribution of the populations could certainly be chosen; the fact is that no one model is definitive or "most realistic". One obvious choice is to assume that all normalized populations $\sum p_i = 1$ are equally likely. In this model it is not possible to prove that the Webster method is perfectly unbiased. The difficulty results from the integer nature of the problem together with the fact that in the normalized model the number of degrees of freedom is reduced by one. In spite of this, any predicted bias of Webster's method derivable from such a model is so small as to be empirically indistinguishable from the hypothesis of perfect unbias when the number of states is large.

One of the advantages of the probabilistic model we have taken is that it can be applied locally, without assuming gross changes in the relative populations of the states. Thus a method is unbiased only if it is unbiased over *every* set of populations that would yield a pre-specified distribution of seats. Other local assumptions may give slightly different numerical results -- and may rule out a mathematically perfect method -- but the empirical result will not change. The agreement between the theory and the historical statistics seems conclusive that Webster's is the least biased divisor method.

Proposition 5.5. The statements of each of the Theorems 5.1, 5.2, and 5.3 may be modified to account for minimum requirements.

6. STAYING WITHIN THE QUOTA

It seems extremely natural to require that no state's apportionment should deviate from its quota by one or more seats; in other words, no state should get less than its quota rounded down, nor more than its quota rounded up. This property is called *staying within the quota.* The primitive desire to stay within the quota clearly motivated Hamilton's method and it was a key point in Webster's critique of Jefferson's method. Webster's method is closer than any other divisor method to the ideal of staying within the quota. Nevertheless, it does not invariably stay within the quota, as the invented examples of Table 6 show. Such examples turn out to be very rare: in practice it is extremely improbable that Webster's method would ever violate quota. Moreover these same examples suggest that staying within the quota may not be such a reasonable idea after all. For example, to give d 26 instead of 25 seats in Table 6 would mean taking a seat from one of the smaller states a, b, or c. Such a transfer would penalize the per capita representation of the small state much more -- in both absolute and relative terms -- than state d is penalized by getting one less than its lower quota. Similar remarks argue for state D getting more than its upper quota in the second example in Table 6. It can be argued that staying within the quota is not really compatible with the idea of proportionality at all, since it allows a much greater variance in the per capita representation of smaller states than it does for larger states. This basic incompatibility between staying within the quota and proportionality is most clearly seen by the following "impossibility theorem", which says that *no* method can be population monotone and stay within the quota.

Theorem 6.1 *There exists no partial method with* $s \geq 4$ *and* $h \geq s+3$ *that is population monotone and stays within the quota.*

Proof of Theorem 6.1. Fix $s \geq 4$, $h \geq s+3$ and suppose M^* is a partial method that is population monotone and stays within the quota. Consider the populations $\underset{\sim}{p} = (5+\varepsilon, 2/3, 2/3, 2/3-\varepsilon, b_5, \ldots, b_s)$ where $b_5, \ldots, b_s$ are any *positive* integers whose sum is h-7 and $\varepsilon > 0$ is some small rational number. Since $\sum p_i = h$, p_i is i's exact quota for all i. Choose $\underset{\sim}{a} \in M^*(\underset{\sim}{p})$. By the quota assumption, $a_1 \geq 5$ and $a_i = b_i$ for all $i \geq 5$. Therefore $a_2 + a_3 + a_4 = h-5 - (h-7) = 2$, so at least one of states 2, 3, 4 gets 0 seats.

Now consider the populations $\underset{\sim}{p}' = (4-\varepsilon, 2-\varepsilon/2, 1/2+\varepsilon/2, 1/2+\varepsilon, b_5, \ldots, b_s)$. Again $\sum p_i' = h$ so p_i' is the exact quota of i. Choose $\underset{\sim}{a}' \in M^*(\underset{\sim}{p}')$. By quota, $a_1' \leq 4$, $a_2' \leq 2$ and $a_i' = b_i$ for $i \geq 5$. Hence either state 3 or state 4 gets one seat. By population monotonicity it must be state 4. Therefore $a_4' > a_4$ while $a_1' < a_1$ so by population monotonicity $p_1'/p_4' < p_1/p_4$, that is $(4-\varepsilon)/(1/2+\varepsilon) < (5+\varepsilon)/(2/3-\varepsilon)$. Simplifying this becomes $\varepsilon > 1/61$, which is false for sufficiently small ε. □

Corollary 6.1 *No population monotone method, i.e. no divisor method, stays within the quota for every problem.*

Proposition 6.1. Every divisor method stays within the quota for all two-state problems.

Proposition 6.2. Webster's method stays within the quota for all three-state problems.

Proposition 6.3. For any divisor method different from Webster's, there is a three-state problem in which the apportionment does not stay within the quota. (Let d(a) be a divisor criterion different than Webster's and choose a such that $d(a) \neq a+1/2$. Let $e = (a+1/2) - d(a)$ and for every integer b construct the problem with populations $\underset{\sim}{p} = (a+1/2, d(a)+e/2, b+1+e/2)$ and house size $h = 2a+b+2$. For sufficiently large b the answer does not stay within the quota.) It follows from this and the preceding exercise that Webster's method is the *unique* divisor method that stays within the quota for all three-state problems.

Population monotonicity is consistent with partial ways of satisfying quota. M *stays above lower quota* if $a_i \geq \lfloor q_i \rfloor$ for all apportionments $\underset{\sim}{a}$ and M *stays below upper quota* if $a_i \leq \lceil q_i \rceil$ for all $\underset{\sim}{a}$.

Proposition 6.4. Jefferson's method is the unique population monotone method that stays above lower quota and Adams' method is the unique population monotone method that stays below upper quota.

Note that since Jefferson's method is not the same as Adams' this gives another proof that no population monotone method satisfies quota. Another partial quota concept that is satisfied by all divisor methods is the following.

Proposition 6.5. No *divisor method* apportionment simultaneously violates upper quota on one state and violates lower quota on another.

While no population monotone method stays within the quota *all* of the time, there are population monotone methods that stay within the quota "almost" all of the time; moreover the best from this standpoint is Webster's method. The tendency of different divisor methods to violate quota can be tested using the same basic model used to test for bias. The idea is to fix a hypothetical apportionment $\underset{\sim}{a}$ and a method M, and ask how likely it is that some state violates quota given this particular distribution of seats. Choose an arbitrary but fixed divisor x and let $\underset{\sim}{p}(\underset{\sim}{a})$ be the set of all populations $\underset{\sim}{p}$ for which $\underset{\sim}{a}$ is a resulting M-apportionment using the divisor x. The simplest and most natural case is to assume that the populations p_i, and hence the quotients p_i/x, are independently and uniformly distributed. The probability that some state i violates lower quota can, in principle, be computed from the expression

$$Pr[a_i+1 \leqq p_i h/\sum p_j] = Pr[(h-a_i-1)p_i/x - (a_i+1)\sum_{j\neq i} p_j/x \geqq 0] .$$

Similarly, the probability that state i violates upper quota can be computed from the expression

$$Pr[a_i-1 \geqq p_i h/\sum p_i] = Pr[(h-a_i+1)p_i/x - (a_i-1)\sum_{j\neq 1} p_j/x \leqq 0] .$$

The probability that no state violates quota is at most the sum of all of these probabilities. However for methods like Webster's these probabilities are extremely small, hence difficult to estimate theoretically. A more practical approach is to estimate the results numerically using a Monte Carlo simulation.

For a problem in which there are many small states, i.e. in which many of the quotients p_i/x are close to zero, methods like Hill's that automatically give every state 1 seat are very likely to violate lower quota on the large states because too many seats will have been used up on the small ones. Thus, the property of automatically giving one seat to every state is a serious defect from the point of view of staying within the quota. In 1970 there were 6 states that received only 1 seat using Hill's method; *a priori* these states could be arbitrarily small. However in fact no state had a quota less than .5. To avoid the unrealistic assumption of very small states, it was assumed in estimating the future likelihood of violating quota that no state's quotient would be less than .5. Thus the probability of violating quota was estimated numerically for each of the five traditional methods by choosing $\underset{\sim}{a}$ to be the actual distribution of seats in 1970, and choosing the quotients p_i/x independently from a uniform distribution on the interval $\min\{.5, d(a_i-1)\} \leqq p_i/x \leqq d(a_i)$. The results are given in Chapter 3, Table 6. Clearly Webster's is superior to the others in regard to staying within the quota.

In fact Webster did not advocate that a method should always stay within the quota, but asked for something slightly weaker. Namely, he said, it should not be possible to take a seat from one state and give it to another and simultaneously bring *both* of them nearer to their quotas.

That is, there should be no states i and j such that

$$(6.1) \qquad q_i - (a_i - 1) < a_i - q_i \quad \text{and} \quad a_j + 1 - q_j > q_j - a_j \quad .$$

Another way of saying the same thing is that no state can be brought closer to its quota without moving another state further from its quota. Any method with this property is said to be *near quota*. This is similar to the idea of Pareto optimality in economics. It should be further noted that Webster's idea is independent of whether near quota is interpreted in absolute or relative terms. In relative terms it would say that no state can be brought closer to its quota *on a percentage basis* without moving another state further from its quota *on a percentage basis*. In other words for no states i and j do we have

$$(6.2) \qquad 1 - \frac{a_i - 1}{q_i} < \frac{a_i}{q_i} - 1 \quad \text{and} \quad \frac{a_j + 1}{q_j} - 1 > 1 - \frac{a_j}{q_j}$$

which is clearly equivalent to (6.1).

The fallacy of Hill's and Huntington's relative difference approach can now be clearly seen: it fails to take into account that there is an *absolute* standard against which the allocation to any state should be compared -- namely the quota. Compared to this standard both the relative and the absolute measures of difference lead to the same result -- Webster's method.

Theorem 6.2 *Webster's method is the unique population monotone method that is near quota, interpreted absolutely or relatively.*

Proof of Theorem 6.2. If a is not near quota, that is if (6.1) or (6.2) holds for some i and j then rearranging we have

$$1 < 2(a_i - q_i) \quad \text{and} \quad 1 > 2(q_j - a_j)$$

or

$$a_i + 1/2 < q_i \quad \text{and} \quad a_j - 1/2 > q_j$$

or

$$q_i/(a_i + 1/2) > q_j/(a_j - 1/2) \quad .$$

Hence the min max inequality for Webster's method is violated, so $\underset{\sim}{a}$ could not be a Webster apportionment. Therefore Webster's method is near quota.

Conversely let M be a population monotone method (i.e. a divisor method) different from Webster's. Then there exists a 2-state problem (p_1,p_2) in which the M-apportionment is uniquely (a_1+1,a_2) whereas the W-apportionment is uniquely (a_1,a_2+1). By the latter, $p_2/(a_2+1/2) > p_1/(a_1+1/2)$. At $h = a_1 + a_2 + 1$ the quota of state 1 is

$$q_1 = \frac{p_1 h}{p_1 + p_2} = \frac{p_1(a_1+1/2 + a_2+1/2)}{p_1 + p_2} < \frac{p_1(a_1+1/2) + p_2(a_1+1/2)}{p_1 + p_2} = a_1 + 1/2 .$$

State 2's quota is

$$q_2 = (a_1 + a_2 + 1) - q_1 > a_2 + 1/2 .$$

Therefore the M-apportionment (a_1+1,a_2) is not near quota. □

In the case of minimum requirements the notion of quota must be modified because the true quotas of some states may be less than their requirements. Hence it may not be possible to satisfy the requirements *and* have enough left over to give the other states even their lower quotas. Similar problems arise in the case of maximum requirements. The "modified quota" of a state is its proportional share subject to the requirements of *all* states being met. More precisely, let $\underset{\sim}{r} \leqq \underset{\sim}{r}^+$ be minimum and maximum requirements and let $\underset{\sim}{q}$ be the "ordinary quotas." Choose a multiplier t such that

$$(6.3) \qquad \sum_1^s \text{mid}\,\{r_i, tq_i, r_i^+\} = h .$$

"Mid" again stands for the middle in value of the three arguments. (6.3) uniquely determines t, and the resulting values $\bar{q}_i = \text{mid}\{r_i,tq_i,r_i^+\}$ are called the *modified quotas*. The *modified upper* *quota* of state i is $\lceil \bar{q}_i \rceil$ and the *modified lower quota* is $\lfloor \bar{q}_i \rfloor$. The method M *stays within the quota* if for all requirements $\underset{\sim}{r} \leqq \underset{\sim}{r}^+$ and M-apportionments $\underset{\sim}{a}$, $\lfloor \bar{q}_i \rfloor \leqq a_i \leqq \lceil \bar{q}_i \rceil$ for all i. The definitions of near quota, staying above lower quota, staying below upper quota can be similarly extended and Theorem 6.2 and Propositions 6.1 - 6.5 hold as stated.

7. STAYING WITHIN THE QUOTA AND HOUSE MONOTONICITY

The preceding section shows that no population monotone method stays within the quota. House monotonicity is a weaker property, implied by population monotonicity but not implying it. Up to this point the only methods discussed that stay within the quota are Hamilton-like and violate both population monotonicity and house monotonicity. Is it possible that there are other types of methods that stay within the quota and are also house monotone? The answer is affirmative.

The simplest of these methods is the so-called quota method. (Balinski and Young (1975)). The Quota method can also be generalized to obtain *all* house monotone methods that stay within the quota. In fact, all apportionments arising from such methods can be described by a system of inequalities, much as the divisor methods can be described by min max inequalities. This approach can be used to adapt traditional methods like those of Adams, Dean, Hill, Webster, Jefferson, and Hamilton so that they always stay within the quota. Unfortunately these adaptations suffer acutely from the population paradox, so are not to be recommended.

Satisfying quota can be described in a convenient analytical way as follows: a_i satisfies lower quota for $\underset{\sim}{p}$ and h if and only if $a_i+1 > q_i = p_i h/p$ and satisfies upper quota if and only if $a_i-1 < q_i = p_i h/p$, where $p = \sum p_j$. Hence $\underset{\sim}{a}$ satisfies quota if and only if

$$(7.1) \qquad p_i/(a_i+1) < p/h < p_i/(a_i-1) \quad \text{for all } i.$$

The most straightforward way of constructing a house monotone apportionment solution f is to define it recursively on h for every given $\underset{\sim}{p}$. For any apportionment $\underset{\sim}{a}$ of h let $U(\underset{\sim}{p},\underset{\sim}{a})$ be the set of states that are *eligible* to receive one more seat without exceeding their upper quotas at house size h+1; thus $U(\underset{\sim}{p},\underset{\sim}{a}) = \{i : a_i < p_i(h+1)/p\}$. Clearly $U(\underset{\sim}{p},\underset{\sim}{a}) \neq \phi$ for every $\underset{\sim}{p}$ and $\underset{\sim}{a}$. The house monotone solutions generated by the following recursive procedure defines the *Quota method* Q.

Theorem 7.1. *Every solution* f *defined as follows is house monotone and stays within the quota:*

(i) $f(\underset{\sim}{p},0) = \underset{\sim}{0}$

(ii) if $f(\underset{\sim}{p},h) = \underset{\sim}{a}$ then $f_k(\underset{\sim}{p},h+1) = a_k+1$ for some state $k \in U(\underset{\sim}{a})$ satisfying $p_k/(a_k+1) \geqq p_i(a_i+1)$ for all $i \in U(\underset{\sim}{p},\underset{\sim}{a})$ and $f_j(\underset{\sim}{p},h+1) = a_j$ for all $j \neq k$.

Proof of Theorem 7.1. It is immediate from the definition that every such f is house monotone and that it stays below upper quota.

Suppose, by way of contradiction, that f violates lower quota for some problem $(\underset{\sim}{p},h)$. As $\underset{\sim}{p}$ will be fixed for the remainder of the proof and f is single-valued, write $\underset{\sim}{a}^{h'} = f(\underset{\sim}{p},h')$ for every $h' \geqq 0$. Also, let $p = \sum p_i$. Since some state, say state 1, is below lower quota at h, the set $S = \{i : a_i^h > p_i h/p\}$ is nonempty. For each i let $h_i \leqq h$ be the least house size for which $a_i^{h_i} = a_i^h$. Choose $j \in S$ such that $h_j = \max_{i \in S} h_i$; thus j is the last state in S to have reached the apportionment it has at h. Now j did not receive the h^{th} seat, because state 1 was also eligible to get the h^{th} seat and had a higher priority, i.e., (7.2) $$a_1^h + 1 \leqq p_1 h/p \quad ,$$

(7.3) $$a_j^h > p_j h/p \quad ,$$

so

(7.4) $$p_j/a_j^h < p/h < p_1/(a_1+1) \quad .$$

Hence $h_j < h$. Let T be the set of states that receive seats between h_j+1 and h inclusive. By definition of T, $T \cap S = \phi$, that is

(7.5) $$a_i^h \leqq p_i h/p \quad \text{for all } i \in T \quad .$$

Therefore using (7.3),

$$p_i/(a_i^{h_j}+1) \geqq p_i/a_i^h \geqq p/h > p_j/a_j^h = p_j/a_j^{h_j} \text{ for all } i \in T.$$

From this and the definition of f it follows that no member of T was eligible to get the h_j^{th} seat:

(7.6) $$a_i^{h_j-1} = a_i^{h_j} \geqq p_i h_j/p \quad \text{for all } i \in T \quad .$$

Subtracting (7.6) from (7.5) and summing over T,

(7.7) $$h - h_j = \sum_T (a_i^h - a_i^{h_j}) \leqq (h - h_j) \sum_T p_i/p \quad .$$

Dividing by $h - h_j > 0$, conclude that $\sum_T p_i \geqq \sum_1^S p = p$, a contradiction, since $j \notin T$ and $p_j > 0$. □

Proposition 7.1. Given requirements $\underset{\sim}{r} \leqq \underset{\sim}{r}^+$ define $U(\underset{\sim}{p},\underset{\sim}{a}) = \{i : a_i < \bar{q}(h+1)\}$ whenever $\underset{\sim}{a}$ is an apportionment for h. Every solution f defined recursively as follows is house monotone and stays within the quota:

(i) $f(\underset{\sim}{p}, \sum_1^s r_i) = \underset{\sim}{r}$

ii) if $f(\underset{\sim}{p},h) = \underset{\sim}{a}$ and $h < \sum r_i^+$, then $f_k(\underset{\sim}{p},h+1) = a_k + 1$ for some state $k \in U(\underset{\sim}{p},\underset{\sim}{a})$ satisfying $p_k/(a_k+1) \geqq p_i/(a_i+1)$ for all $i \in U(\underset{\sim}{p},\underset{\sim}{a})$ and $f_j(\underset{\sim}{p},h+1) = a_j$ for all $j \neq k$.

The argument parallels that of Theorem 7.1 and leads to the analog of (7.6) $a_i^h \leqq \overline{q}_i(h) = thp_i/p$ for all $i \in T$. Next deduce that no member of T was eligible to get the h_jth seat. Since every member of T gets at least one seat between h_j and h, this means that $r_i^+ > a_i^{h_j-1} = a_i^{h_j} \geqq \overline{q}(h_j) = t_j h_j p_i/p$ for all $i \in T$. From these two expressions we obtain

$$(th - t_j h_j)\sum_T p_i/p \geqq \sum_T a_i^h - a_i^{h_j} = h - h_j \geqq \sum \overline{q}_i(h) - \overline{q}(h_j) \geqq (th - t_j h_j)\sum_T p_i/p$$

and a contradiction, since $j \notin T$.

Any house monotone method may be described recursively on the house size h by defining which states are eligible to get the "next" seat given the current apportionment. The family $\overline{\underset{\sim}{a}}$ of all apportionment solutions $\underset{\sim}{f}$ that are house monotone and stay within the quota may be described in this way using a suitable notion of eligibility.

Fix $\underset{\sim}{p} > \underset{\sim}{0}$, and $\underset{\sim}{a}$ an apportionment of h. For each integer $\alpha \geqq 1$ let $S_\alpha = S_\alpha(\underset{\sim}{p},\underset{\sim}{a})$ be the set of states i such that $\lfloor p_i(h+\alpha)/p \rfloor > a_i$, and let $\overline{\alpha} = \overline{\alpha}(\underset{\sim}{p},\underset{\sim}{a})$ be the least $\alpha \geqq 1$ such that $\sum_{S_\alpha} (\lfloor p_i(h+\alpha)/p \rfloor - a_i) \geqq \alpha$. Define $L(\underset{\sim}{p},\underset{\sim}{a}) = S_{\overline{\alpha}}$ or if no such $\overline{\alpha}$ exists, let $L(\underset{\sim}{p},\underset{\sim}{a})$ be the set of all states.

To construct a house-monotone sequence of apportionments continuing from $\underset{\sim}{a}$ that stay within the quota, it is clearly necessary that the $(h+1)^{st}$ seat be given to some state in $L(\underset{\sim}{p},\underset{\sim}{a})$ since otherwise at house size $h + \overline{\alpha}$ some state in $L(\underset{\sim}{p},\underset{\sim}{a})$ would have fallen below its lower quota. It is also clearly necessary that the $(h+1)^{st}$ seat be given to some state in $U(\underset{\sim}{p},\underset{\sim}{a})$, or else the upper quota would be violated at h+1. It turns out that these two conditions are both necessary *and* sufficient to determine which states are eligible to get the $(h+1)^{st}$ seat.

Theorem 7.2. *f is a house monotone solution satisfying quota if and only if for each* $\underset{\sim}{p} > \underset{\sim}{0}$ *f is constructed recursively as follows:*

(i) $f(\underset{\sim}{p},0) = 0$

(ii) if $f(\underset{\sim}{p},h) = \underset{\sim}{a}$ then $f(\underset{\sim}{p},h+1)$ is found by giving a_i+1 seats to some one state $i \in L(\underset{\sim}{p},\underset{\sim}{a}) \cap U(p,a)$ and a_j seats to each $j \neq i$.

In this recursion the eligibility class $L(\underset{\sim}{p},\underset{\sim}{a}) \cap U(\underset{\sim}{p},\underset{\sim}{a})$ *is never empty.*

Proof of Theorem 7.2. It has already been noted that every $f \in \overline{Q}$ must be defined as in (i) and (ii). Conversely, if f is defined as in (i) and (ii) then it is obviously house monotone and stays below upper quota. Suppose that for some $\underset{\sim}{p}$ and h, $\underset{\sim}{a} = f(\underset{\sim}{p},h)$ violates lower quota. Let $\underset{\sim}{0},\underset{\sim}{a}^1,\ldots,\underset{\sim}{a}^h$ be the apportionments by f up to h, and suppose that $\lfloor p_k h/p \rfloor > a_k^h$. Then $k \in S_1 = S_1(\underset{\sim}{p},\underset{\sim}{a}^{h-1})$ and $\sum_{S_1} (\lfloor p_i h/p \rfloor - a_i^{h-1}) \geqq \sum_{S_1} (\lfloor p_i h/p \rfloor - a_i^h) \geqq 1$, so $\overline{\alpha}(\underset{\sim}{p},\underset{\sim}{a}^{h-1}) = 1$ and $S_1 = L(\underset{\sim}{p},\underset{\sim}{a}^{h-1})$. Therefore the h^{th} seat was given to some state in S_1, but not to k, so in fact $\sum_{S_1} (\lfloor p_i h/p \rfloor - a_i^{h-1}) > \sum_{S_1} (p_i h/p - a_i^h)$ and $\sum_{S_1} (\lfloor p_i h/p \rfloor - a_i^{h-1}) \geqq 2$.

For each i, $p_i h/p > a_i^{h-1}$ implies $p_i h/p > a_i^{h-2}$, so $S_2 = S_2(\underset{\sim}{p},\underset{\sim}{a}^{h-2}) \supseteq S_1(\underset{\sim}{p},\underset{\sim}{a}^{h-1})$. Therefore $\sum_{S_2} (\lfloor p_i h/p \rfloor - a_i) \geqq \sum_{S_1} (\lfloor p_i h/p \rfloor - a_i) \geqq 2$ whence $\bar{\alpha}(\underset{\sim}{p},\underset{\sim}{a}^{h-2}) \leqq 2$. Since $p_i(h-2+\bar{\alpha})/p > a_i^{h-2}$ implies $p_i h/p > a_i^{h-2}$ it follows that $L(\underset{\sim}{p},\underset{\sim}{a}^{h-2}) \subseteq S_2$. Therefore the $(h-1)^{st}$ seat was given to some state in S_2 but not in S_1, whence $\sum_{S_2} (\lfloor p_i h/p \rfloor - a^{h-2}) > \sum_{S_1} (\lfloor p_i h/p \rfloor - a_i^{h-2}) \geqq 2$ and $\sum_{S_2} (\lfloor p_i h/p \rfloor - a_i^{h-2}) \geqq 3$. Continuing in this manner it follows that for $S_h = S_h(p,0)$ $\sum_{S_h} (\lfloor p_i h/p \rfloor - 0) \geqq h+1$, which is impossible.

It remains to show that the eligibility class $L(\underset{\sim}{p},\underset{\sim}{a}) \cap U(\underset{\sim}{p},\underset{\sim}{a})$ must be nonempty at each step of the recursion. If $L = L(\underset{\sim}{p},\underset{\sim}{a})$ is the set of all states the result holds. Otherwise there is an $\bar{\alpha} = \bar{\alpha}(\underset{\sim}{p},\underset{\sim}{a}) \geqq 1$ such that $\sum_{L} (\lfloor p_i(h+\bar{\alpha})/p \rfloor - a_i) \geqq \bar{\alpha}$, where $\lfloor p_i(h+\bar{\alpha})/p \rfloor > a_i$ for each $i \in L$. If $\bar{\alpha} = 1$ this says that $p_i(h+1)/p \geqq a_i + 1$, so $i \in L \subseteq U$. If $\bar{\alpha} > 1$ but $L \cap U = \phi$, then $p_i(h+1)/p \leqq a_i$ for all $i \in L$ and so

$$p_i(h+\bar{\alpha})/p - a_i \leqq p_i(h+\bar{\alpha})/p - p_i(h+1)/p = (p_i/p)(\bar{\alpha}-1) \quad .$$

By definition of L, $\sum_{L} (p_i(h+\bar{\alpha})/p - a_i) \geqq \bar{\alpha}$, hence we obtain $\bar{\alpha} \leqq (\sum_{L} p_i/p)(\bar{\alpha}-1)$ and $\sum_{L} p_i/p \geqq \bar{\alpha}/(\bar{\alpha}-1) > 1$ (since $\bar{\alpha} > 1$), a contradiction. □

To compute the set $L(\underset{\sim}{p},\underset{\sim}{a})$ it is not necessary to consider infinitely many α. In fact, if $\alpha > \max_i \left\lceil \frac{a_i - p_i h/p}{p_i/p} \right\rceil$, then for every i, $p_i(h+\alpha)/p > a_i$ and

$$\sum_1^s (\lfloor p_i(h+\alpha)/p \rfloor - a_i) < \sum_1^s (p_i(h+\alpha)/p - a_i) = h + \alpha - \sum a_i = \alpha$$

showing that $\bar{\alpha}(\underset{\sim}{p},\underset{\sim}{a}) \leqq \max \left\lceil \frac{a_i - p_i h/p}{p_i/p} \right\rceil$ if it is defined at all.

Let $\underset{\sim}{a} \in \bar{Q}(\underset{\sim}{p},h)$ if and only if there is a house monotone solution f satisfying quota with $\underset{\sim}{a} = f(\underset{\sim}{p},h)$. The Quota method Q is a sub-method of $\bar{Q}$ and is determined by a more stringent criterion of eligibility. But this eligibility criterion is also very simple to compute, as it only "looks ahead" by one seat. Given $\underset{\sim}{a} \in Q(\underset{\sim}{p},h)$, it is only necessary to find those states for which $p_i(h+1)/p > a_i$; the eligible class at $\underset{\sim}{a}$ is then

$$\{i: p_i/a_i > p/(h+1) \text{ and } p_i/(a_i+1) \geqq p_j/(a_j+1) \text{ whenever } p_j/a_j > p/(h+1)\}.$$

As in the case of rank-index methods, it is also possible to give a "local" characterization of all $\bar{Q}$-apportionments for a given $\underset{\sim}{p}$ and h by a system of inequalities.

Theorem 7.3. $\underset{\sim}{a} \in \bar{Q}(\underset{\sim}{p},h)$ if and only if $\underset{\sim}{a} \geqq \underset{\sim}{0}$, $\sum a_i = h$ and

$$\sum_i \max\,(a_i, \lfloor p_i(h+\alpha)/p \rfloor) \leqq h+\alpha,$$

$$\sum_i \min\,(a_i, \lceil p_i(h-\beta)/p \rceil) \geqq h-\beta, \text{ for all } \alpha,\beta \geqq 0.$$

The conditions are feasible for every $\underset{\sim}{p}$ *and* h.

As in the computation of $L(\underset{\sim}{p},\underset{\sim}{a})$ it is easy to verify that if the inequalities hold for

$$0 \leqq \alpha \leqq \max_i \left\lceil \frac{a_i - p_i h/p}{p_i/p} \right\rceil \text{ and } 0 \leqq \beta \leqq \max_j \left\lceil \frac{p_j h/p - a_j}{p_j/p} \right\rceil,$$

then they hold for *all* $\alpha\ \beta \geqq 0$.

The conditions are transparently necessary if $\underset{\sim}{a} = f(\underset{\sim}{p},h)$ for some house monotone f that stays within the quota. Sufficiency is established by showing that if $\underset{\sim}{a}$ satisfies the conditions at $h > 0$, then there is an $\underset{\sim}{a}'$, differing from $\underset{\sim}{a}$ by exactly one seat, say at state k, such that $\underset{\sim}{a}'$ satisfies the conditions at h-1 and $k \in L(\underset{\sim}{p},\underset{\sim}{a}') \cap U(\underset{\sim}{p},\underset{\sim}{a}')$. The result then follows from an induction argument using Theorem 7.2.

The characterization in Theorem (7.2) suggests an approach, first proposed by Still (1979) in which various classical methods, like the divisor methods or Hamilton's method, can be modified so as to be both house monotone and stay within the quota. Let d(a) be a divisor criterion and in the recursion of Theorem 7.2 (ii) give the $(h+1)^{st}$ seat to some eligible state that maximizes $p_i/d(a_i)$ over all eligible states. In this way we obtain the Quota-Jefferson method (the same as the Quota method), the Quota-Webster method, the Quota-Hill method, etc. Likewise we could define the Quota-Hamilton method by giving the $(h+1)^{st}$ seat to some eligible state with maximum "remainder" $p_i(h+1)/p - a_i$.

To illustrate how such methods work, we apply the Quota method to the four-state example shown in Table 7.1.

State	Population	Quota	Q-Appt.	Quota	Q-Appt.	Quota	Q-Appt.
A	501	4.59	5*	5.01*	6	5.43	6
B	394	3.61	4	3.94	4	4.27	5
C	156	1.43	1	1.56	1	1.69	1
D	149	1.37	1	1.49	1	1.61	1
Total	1200	11.00	11	12.00	12	13.00	13

Table 7.1. Example Illustrating the Quota Method (Starred State Gets the Next Seat).

The rule for computing the Quota method is simpler than the others because (by Theorem 7.1) it is not necessary to calculate the sets $L(\tilde{p},\tilde{a})$ at each stage of the recursion. It suffices to maximize $p_i/(a_i+1)$ over all $i \in U(\tilde{p},\tilde{a})$, i.e. the eligible states are those that could get one more seat without violating upper quota at the next larger house size. The recursion begins at h = 0 and leads to the solution at h = 11 shown in Table 7.1. The quotas at h = 12 identify the eligible states at this point to be A, C, and D. Of these $p_A/(a_A+1)$ is maximum hence it gets one more seat at house size 12. At this point B, C, and D are eligible and state B deserves the next seat according to the Jefferson divisor criterion.

Consider now a variation of the above example in which state B gains in population and all other states stay the same. Table 7.2 illustrates the Quota method calculation for this case, beginning with h = 11.

State	Population	Quota	Q-Appt.	Quota	Q-Appt.	Quota	Q-Appt.
A	501	4.57	5	4.99	5*	5.40	6
B	400	3.65	4	3.98	4	4.31	4
C	156	1.42	1*	1.55	2	1.68	2
D	149	1.36	1	1.48	1	1.61	1
Total	1206	11.00	11	12.00	12	13.00	13

Table 7.2. Example showing Quota method not Population Monotone

The solutions at h = 1 are the same, but now the only eligible states are C and D, so C gets the next seat. At h = 12 states A, B and D are eligible and A gets the next seat according to the Jefferson divisor criterion. Therefore while state B gained in population relative to all other states, it actually lost a seat at h = 13. Similar examples show that the Quota-Webster method, the Quota-Hill method, and indeed all methods of this type violate population monotonicity and hence are not to be recommended.

Proposition 7.2. The analog of Theorem 7.2 holds in the presence of minimum and maximum requirements by defining $U(\underset{\sim}{p},\underset{\sim}{a}) = \{i : a_i < \bar{q}_i(h+1)\}$ and $L(\underset{\sim}{p},\underset{\sim}{a}) = \{i : \bar{q}_i(h+\alpha) > a_i\}$ for some least α such that $\sum_L (\bar{q}_i(h+\alpha) - a_i) \geqq \alpha$ or, if no such α exists, $L(\underset{\sim}{p},\underset{\sim}{a})$ is the set of all states.

Proposition 7.3. Given requirements $\underset{\sim}{r} \leqq \underset{\sim}{r}^+$ and $\underset{\sim}{a} \geqq \underset{\sim}{0}$, $\sum_i a_i = h$, there is a house monotone solution f that stays within the quota with $\underset{\sim}{a} = f(\underset{\sim}{p},h)$ if and only if

$$\sum_i \max\,(a_i, \lfloor \bar{q}_i(h+\alpha) \rfloor) \leqq h + \alpha,$$

and

$$\sum_i \min\,(a_i, \lceil \bar{q}_i(h-\beta) \rceil) \geqq h - \beta \quad \text{for all } \alpha,\ \beta \geqq 0 \ .$$

8. UNIFORMITY

The problem of fair representation is but one instance of a problem of *fair division*: How should an inheritance be shared among heirs? How should a public good (e.g., airport capacity) be shared among users (e.g., air carriers)? How should taxes be shared among residents? How should seats in Congress be shared among states? An inherent principle of any fair division is that *every part of a fair division should be fair*. For example, one property of a fair division of an inheritance should be that no subset of heirs would want to make trades after the division is made. The principle is very general.

In the context of fair representation this principle says that an apportionment that is acceptable to all states must be acceptable if restricted to any subset of states considered alone. The way in which two states share a given number of seats is independent of the populations of *other* states. This is somewhat reminiscent of Arrow's "independence of irrelevant alternatives" axiom in the theory of social choice, but with an important difference: it can be realized. In the context of fair division and even more particularly in apportionment, it is a central consideration, since inevitably each state will compare its representation with each of its sister states.

Specifically, a method M is *uniform* (Balinski and Young, 1978b) if for every t, $2 \leq t \leq s$, $(a_1,\ldots,a_s) \in M((p_1,\ldots,p_s),h)$ implies $(a_1,\ldots,a_t) \in M((p_1,\ldots,p_t),\sum_1^t a_i)$, and if also $(b_1,\ldots,b_t) \in M((p_1,\ldots,p_t)\sum_1^t a_i)$ then $(b_1,\ldots,b_t, a_{t+1},\ldots,a_s) \in M((p_1,\ldots,p_s),h)$. In other words, each restriction of a fair apportionment is fair; moreover, if a restriction admits a different apportionment of the same number of seats --

that is, if the restricted problem admits a tie -- then there is a corresponding tie in the entire apportionment. The min-max description of divisor methods immediately establishes that all divisor methods are uniform.

Uniform methods include more than the divisor methods, however. In fact, the class of uniform methods can be described by generalizing the recursive procedure for computing divisor methods.

Let a *rank-index* r(p,a) be any real-valued function of rational $p > 0$ and integer $a \geqq 0$ that is monotone decreasing in a: $r(p,a-1) > r(p,a)$. Now define a particular house monotone solution f as follows:

(i) for $h = 0$ let $f(\underset{\sim}{p},0) = \underset{\sim}{0}$.

(ii) if $f(\underset{\sim}{p},h) = \underset{\sim}{a}$, then $f(\underset{\sim}{p},h+1)$ is found by giving a_i+1 seats to some state i such that $r(p_i,a_i) \geqq r(p_j,a_j)$ and a_j seats to each $j \neq i$.

Let $\mathscr{F}$ be the set of all particular solutions defined recursively by (i) and (ii). The *rank-index method based on* r(p,a) is defined by

$$M(\underset{\sim}{p},h) = \{\underset{\sim}{a} : \underset{\sim}{a} = \underset{\sim}{f}(\underset{\sim}{p},h) \text{ for some } f \in \mathscr{F}\} .$$

Divisor methods are just special cases of rank-index methods in which the rank-index has the form $r(p,a) = p/d(a)$. The following theorem shows that the min-max inequality holds also for rank-index methods, hence in particular all rank-index methods are uniform.

Theorem 8.1 $M(\underset{\sim}{p},h) = \{\text{integer } \underset{\sim}{a} \geqq \underset{\sim}{0}: \min_{a_i>0} r(p_i,a_i-1) \geqq \max_{a_i \geqq 0} r(p_i,a_i)\}$ *is the rank-index method based on* $r(p,a)$.

Proof of Theorem 8.1. First we show that if $\underset{\sim}{a}$ satisfies the inequalities for some given $\underset{\sim}{p}$ and h, then there exists $f \in \mathcal{F}$ with $f(\underset{\sim}{p},h) = \underset{\sim}{a}$. If false, let $\bar{h}$ be the least $h \geqq 0$ for which it is false for the given $\underset{\sim}{p}$. Let state i satisfy $r(p_i,a_i-1) \leqq r(p_j,a_j-1)$ for all $j \neq i$ and define $\underset{\sim}{a}'$ as follows: $a_i' = a_i-1$ and $a_j' = a_j$ for all $j \neq i$. Then

$$r(p_j,a_j'-1) \geqq r(p_i,a_i') \geqq r(p_j,a_j') \quad \text{for all } j \neq i$$

and

$$r(p_i,a_i'-1) > r(p_i,a_i') \quad .$$

Hence $\underset{\sim}{a}'$ satisfies the inequalities for $\underset{\sim}{p}$ and h-1 and by the induction hypothesis there is some solution $f \in \mathcal{F}$ with $f(\underset{\sim}{p},h-1) = \underset{\sim}{a}'$. Let f^{h-1} be the restriction of f up to house size h-1. Since $r(p_i,a_i') \geqq r(p_j,a_j')$ for all j, the recursive procedure implies that there is some extension g of f^{h-1} which is in $\mathcal{F}$ and such that $g(\underset{\sim}{p},h) = \underset{\sim}{a}$.

Conversely, suppose that for some $f \in \mathcal{F}$, $\underset{\sim}{a} = f(\underset{\sim}{p},h)$ does not satisfy the min max inequality; say $r(p_i,a_i) > r(p_j,a_j-1)$ for some $i \neq j$. Let $k \leqq h$ be the house size at which (by the recursive procedure constructing f) state j received its a_jth seat. At k-1 state i had $a_i' \leqq a_i$ seats and state j had $a_j' = a_j-1$ seats. But then $r(p_i,a_i') \geqq r(p_i,a_i) > r(p_j,a_j')$ so $r(p_j,a_j')$ was not maximum, contradicting the recursive procedure. □

The same rank-index method can be represented by many different rank-indices. In fact any two rank-indices r and r' that are *order equivalent*, in the sense that $r(p,a) \geqq r(q,b)$ iff

$r'(p,a) \geqq r'(q,b)$, yield the same rank-index method, by Theorem 8.1. The converse is true as well.

<u>*Theorem 8.2*</u> *Two rank-indices* r' *and* r'' *represent the same rank-index method* M *if and only if* r' *is order equivalent with* r''.

<u>*Proof of Theorem 8.2.*</u> Let M be a rank-index method based on r' and also based on r''. We will show that r' is order equivalent to r''. Suppose, by way of contradiction, that $r'(p,a) > r'(a,b)$ while $r''(p,a) \leqq r''(q,b)$ for some pairs (p,a) and (q,b).

Let $(\bar{a},\bar{b})$ be an M-apportionment of $h = a+b+1$ seats for the two-state problem (p,q). If $\bar{a} \leqq a$ then $\bar{b} \geqq b+1$ and by the monotonicity of r',

$$r'(p,\bar{a}) \geqq r'(p,a) > r'(q,b) \geqq r'(q,\bar{b}-1) \quad ,$$

but this contradicts the min max inequality of Theorem 8.1. Hence

(8.1) $\bar{a} \geqq a+1$ and $\bar{b} \leqq b$ for all M-apportionments $(\bar{a},\bar{b})$.

But then

$$r''(p,\bar{a}-1) \leqq r''(p,a) \leqq r''(q,b) \leqq r''(q,\bar{b})$$

again contradicting the min max inequality unless these inequalities are all equalities and $\bar{a} = a+1$, $\bar{b} = b$. Therefore by Theorem 8.1 $(a,b+1)$ is also an M-apportionment, contradicting (8.1). □

Rank-index methods are not only uniform: they are essentially the *only* methods that are uniform. This is true with conditions that are much weaker than those that have been assumed heretofore. A method M is said to be *balanced* if whenever two states have equal populations then their apportionments do not differ by more than one seat. This is a much less demanding concept than weak proportionality and is satisfied by every non-frivolous method known to the authors. (See Proposition 8.1). The only other elementary principle that is necessary is that of symmetry. *Until further notice* we assume only that all methods are balanced and symmetric.

Theorem 8.3 *A method is uniform if and only if it is a rank-index method.*

Corollary. *Every uniform method is house-monotone.*

This surprising corollary is an immediate consequence of the theorem and the definition of a rank-index method. It highlights both the importance and the strength of the uniformity concept.

Proof of Theorem 8.3. We have already noted that a rank-index method is uniform; and, of course, it is balanced and symmetric.

To prove the converse we show that a suitably defined priority relation on pairs (p,a) is an order, and that it can be represented by a real-valued, order-preserving function r(p,a).

Suppose that M is uniform, balanced and symmetric. The following must then hold.

(8.2) If (p,q;h) has M-apportionment (a,b), and (p,q;h') has M-apportionment (a',b'), and $a' < a$, $b' > b$, then $h = h'$, $a' = a-1$ and $b' = b+1$.

Without loss of generality assume that $a'+b' \geq a+b$, say $a'+b' = a+b+k$. To prove (8.2) consider first the two cases $k = 0$ and $k = 1$.

<u>$k = 0$</u>. Let (x,x',y,y') be an M-apportionment for $(p,p,q,q;2h)$. Since M is balanced, $|x-x'| \leq 1$ and $|y-y'| \leq 1$. Either $x+x'$ and $y+y'$ are both even or both are odd. Hence, either $x = x'$ and $y = y'$ or (using symmetry) $x' = x+1$ and $y' = y-1$. In either case $x'+y' = x+y = h = h'$. Uniformity implies that (a,b) may be substituted for (x,y) and (a',b') for (x',y') to obtain the M-apportionment (a,a',b,b'). By balanced, $|a-a'| \leq 1$ and $|b-b'| \leq 1$, so, since $a' < a$ and $b' > b$, we have $a' = a-1$ and $b' = b+1$.

<u>$k = 1$</u>. $a'+b' = a+b+1$, $b' > b$, $a' < a$ implies $b' \geq b+2$. Use balanced to deduce that $(p,p,q,q;2h+1)$ has an M-apportionment either of form $(x,x+1,y,y)$ or $(x,x,y,y+1)$ where $x+y = h$. By uniformity it follows that (a,a',b,b') is an M-apportionment. But $b' \geq b+2$ contradicts balanced.

Therefore, if (a,b) apportions $(p,q;h)$ and (a',b') apportions $(p,q;h+1)$ then we must have $a' \geq a$ *and* $b' \geq b$. That is, *every particular* M-*solution* f *is house-monotone on two states.* It is therefore impossible to have $a' < a$, $b' > b$ and $a'+b' > a+b$. This completes the proof of (8.2).

Consider now the family $\mathscr{P} = \{(p,a) : p > 0,\ a \geq 0,\ p \text{ rational},\ a \text{ integer}\}$. Define a relation $\gtrsim$ on $\mathscr{P}$ as follows:

(8.3) $(p,a-1) \gtrsim (q,b)$ if and only if there is some M-apportionment $(\bar{a},\bar{b})$ for (p,q) and $h = \bar{a} + \bar{b}$ such that $\bar{a} \geq a > 0$, $\bar{b} \leq b$.

Intuitively, the method M gives preference for an extra seat to a state having population p and a-1 seats over that of a state having population q and b seats.

Write $(p,a) > (q,b)$ if $(p,a) \gtrsim (q,b)$ and not $(q,b) \gtrsim (p,a)$. Also write $(p,a) \sim (q,b)$ if both $(p,a) \gtrsim (q,b)$ and $(q,b) \gtrsim (p,a)$. Then from (8.2) we have

(8.4) $(p,a-1) > (q,b)$ if and only if either $a' \geqq a$ or $b' \leqq b$ for every M-apportionment (a',b') for (p,q) and all h,

and

(8.5) $(p,a-1) \sim (q,b)$ if and only if $(a-1,b+1)$ and (a,b) are both apportionments of $(p,q;a+b)$.

Next we show that *the relation* $\gtrsim$ *is a partial order on* $\mathcal{P}$, that is,

(8.6) $(p,a) \gtrsim (q,b)$ and $(q,b) \gtrsim (r,c)$ implies $(p,a) \gtrsim (r,c)$.

Note, first, that (p,a) and (r,c) are comparable since one can consider the problem $(p,r;a+c+1)$. To prove (8.6) we suppose that $(r,c) > (p,a)$ and derive a contradiction.

Construct the problem $(p,q,r;a+b+c+1)$. By (8.4) and uniformity *every* apportionment (x,y,z) of this problem satisfies either $z \geqq c+1$ or $x \leqq a$.

Suppose $z \geqq c+1$. $(q,b) \gtrsim (r,c)$ implies that *either* $y \geqq b+1$ (or $z \leqq c$) for every (y,z), *or* that $(b+1,c)$ and $(b,c+1)$ are both apportionments for (q,r). If $y \geqq b+1$ then $x \leqq a-1$ and so $(q,b) \gtrsim (p,a-1)$, a contradiction. Otherwise, $(b+1,c)$ must be an apportionment, again a contradiction.

Suppose, then, that $x \leqq a$ and $z \leqq c$, which means $y \geqq b+1$. Then $(q,b) \gtrsim (p,a)$, and so $(q,b) \sim (p,a)$ and $(a+1,b)$ and $(a,b+1)$ must both be apportionments, a contradiction. This establishes (8.6).

The fact that $\gtrsim$ is a partial order on the countable set $\mathscr{P}$ implies that there exists an order preserving function r from $\mathscr{P}$ to the real numbers. The function r may be constructed as follows. Let $\pi^0, \pi^1, \pi^2, \ldots$ be an enumeration of all pairs in $\mathscr{P}$, and for each k define an order-preserving function ϕ^k: $\{\pi^0, \ldots, \pi^k\} \to R$ such that ϕ^k agrees with ϕ^{k-1} on $\{\pi^0, \ldots, \pi^{k-1}\}$. Define r to be the union of all the ϕ^k's.

Let $\underset{\sim}{a}$ be any M-apportionment for the problem $(\underset{\sim}{p};h)$. By definition of $\gtrsim$ and uniformity, $(p_i, a_i - 1) \gtrsim (p_j, a_j)$ for all $i \neq j$, $a_i > 0$. Moreover, by balanced $(p_i, p_i; 2a_i)$ has the *unique* apportionment (a_i, a_i), whence $(p_i, a_i - 1) > (p_i, a_i)$. Therefore $\underset{\sim}{a}$ satisfies the min max inequality

$$(8.7) \qquad \min_{a_i > 0} r(p_i, a_i - 1) \geqq \max_{a_j} r(p_j, a_j)$$

and, moreover,

$$(8.8) \qquad r(p, a-1) > r(p, a) \text{ for all } p, \ a > 0.$$

Suppose, conversely, that $\underset{\sim}{a}$ satisfies (8.7), $\sum a_i = h$. We must then show that $\underset{\sim}{a}$ is an M-apportionment for $(\underset{\sim}{p};h)$. Let $\underset{\sim}{b}$ be *some* M-apportionment for $(\underset{\sim}{p};h)$ and suppose $\underset{\sim}{b} \neq \underset{\sim}{a}$. Choose i and j so that $a_i < b_i$ and $a_j > b_j$. Both $\underset{\sim}{a}$ and $\underset{\sim}{b}$ satisfy (8.7) and $r(p,a)$ is monotone decreasing in $\underset{\sim}{a}$, implying

$$r(p_j, b_j) \geqq r(p_j, a_j - 1) \geqq r(p_i, a_i) \geqq r(p_i, b_i - 1) \geqq r(p_j, b_j) \quad ,$$

and so all inequalities are equalities. Moreover, by (8.8), $b_j = a_j - 1$ and $b_i = a_i + 1$. But $r(p_i, b_i - 1) = r(p_j, b_j)$ says that $(p_i, b_i - 1) \sim (p_j, b_j)$, so by (8.5) $(b_i - 1, b_j + 1) = (a_i, a_j)$ is an alternative apportionment of $a_i + a_j$ seats between p_i and p_j. By uniformity (a_i,a_j) may be substituted for (b_i,b_j) in $\underset{\sim}{b}$ to produce an M-apportionment $\underset{\sim}{b}'$ that differs from $\underset{\sim}{a}$ in fewer components. Continuing in this manner we conclude that $\underset{\sim}{a}$ itself must be an M-apportionment. This completes the proof of Theorem 8.3.* □

Although uniformity as a property seems innocuous, since so natural, it has surprisingly strong implications. It can be satisfied only by rank-index methods under very general conditions. *Reinstate now* the elementary principles that any method should be homogeneous, symmetric and weakly proportional (but not necessarily complete).

Theorem 8.4 *A method* M *is uniform and weakly population monotone if and only if it is a divisor method.*

Proof of Theorem 8.4. A uniform method that is weakly proportional is balanced (Proposition 8.1). Therefore, M is a rank-index method with some $r(p,a)$ that is monotone decreasing in a. Furthermore, M homogeneous means that $(xp,a) \geq (xq,b)$ for rational $x > 0$ if and only if $(p,a) \geq (q,b)$ and so $r(xp,a) \geqq r(xq,b)$ if and only if $r(p,a) \geqq r(q,b)$.

Suppose M is weakly population monotone. Then, for any populations $p' > p > 0$ and any integer $a \geqq 0$ the two-state problem $(p',p;2a+1)$ has only apportionments of form (a',a), where $a' \geqq a+1$. Therefore, $r(p',a) > r(p,a)$, that is, $r(p,a)$ is monotone increasing in p.

Fix $a \geqq 2$. Then $r(1,a) < r(1,1) < r(a+1,a)$. Let $\mathscr{P}^a = \{p : r(p,a) \geqq r(1,1)\}$ and define $p^a = \inf \mathscr{P}^a$. By the preceding inequalities $1 \leqq p^a \leqq a+1$, so p^a is finite. It may be irrational; however, in any case, for any increasing sequence of rationals

*A variant of this result characterizing rank-index methods by a slightly weaker formulation of uniformity together with house monotonicity is proved in (Balinski and Young 1977b).

$p^{a(n)} \to p^a$ and any rational $\varepsilon > 0$ there is an $n(\varepsilon)$ so that for $n \geqq n(\varepsilon)$

$$r(p^{a(n)},a) \leqq r(1,1) < r(p^{a(n)}+\varepsilon,a) \quad .$$

Therefore, since M is homogeneous,

(8.9) $$r(p,a) \leqq r(p/p^{a(n)},1) \text{ and } r(p/(p^{a(n)}+\varepsilon),1) < r(p,a)$$

for any rational $\varepsilon > 0$ and n sufficiently large.

Define $d(a) = p^a$ for $a \geqq 2$ and $d(1) = 1$. If there exists a finite $p^0 = \inf \mathscr{P}^0 > 0$, let $d(0) = p^0$; otherwise, $d(0) = 0$. We claim that by this definition $p/d(a)$ is order equivalent to $r(p,a)$.

Consider, first, the case $d(0) > 0$. Suppose $r(p,a) \leqq r(q,b)$. Then by (8.9), with rational increasing $p^{a(n)} \to p^a$, $q^{b(n)} \to q^b$ and any rational $\varepsilon > 0$,

$$r(p/(p^{a(n)}+\varepsilon),1) < r(p,a) \leqq r(q,b) \leqq r(q/q^{b(n)},1)$$

for n sufficiently large, and so, by weak population monotonicity

$$p/(p^{a(n)}+\varepsilon) < q/q^{b(n)}$$

implying

$$p/p^a \leqq q/q^b \qquad \text{or} \qquad p/d(a) \leqq q/d(b) \quad .$$

If $r(p,a) = r(q,b)$ then $r(p,a) \leqq r(q,b)$ and $r(q,b) \leqq r(p,a)$ and so $p/d(a) = q/d(b)$.

Consider, then, the case $d(0) = 0$. The same reasoning as above applies for all $a,b > 0$. $d(0) = 0$ means that $r(p,0) > r(1,1)$ for all p, thus $r(p,0) > r(q,b)$ for all $b > 0$ and all p,q. Similarly, $p/d(0) > q/d(b)$ since $d(b) > 0$. For $a = b = 0$, $r(p,0) \geqq r(q,0)$ if and only if $p \geqq q$ which holds if and only if $p/0 \geqq q/0$ by convention. □

This theorem has strong implications. Notice, first, that without invoking the completeness property one obtains divisor methods, which can trivially be made complete since they are continuous. Most important, the properties that (i) any part of a fair division must be fair (uniformity) and (ii) a greater claim of one player over another guarantees the first at least as great a part of the spoils as the second (weak population monotonicity) are sufficient to characterize divisor methods. This provides yet another way of seeing why divisor methods are really the only ones that are suitable candidates for apportionment.

Proposition 8.1. If M is uniform and weakly proportional, then M is balanced.

(One shows that for every $a \geq 0$ the problem $(p,p,p,p;4a+2)$ has a unique apportionment of form (x,x,y,y) up to permutations. Using this and induction on a it may be shown that in each of the problems $(p,p;2a+1)$, $(p,p,p;3a)$, $(p,p,p;3a+1)$, and $(p,p,p;3a+2)$ each state must receive either a or a+1 seats. Conclude finally that M is balanced.)

In the next three propositions a method M is assumed to be *only* uniform and symmetric (see Balinski and Young 1978a).

Proposition 8.2. M is the method of Jefferson if and only if it satisfies lower quota.

Proposition 8.3. M is the method of Adams if and only if it satisfies upper quota.

Proposition 8.4. M is the method of Webster if and only if it is near to quota.

Proposition 8.5. There exists no uniform and symmetric method that satisfies quota.

Proposition 8.6. The concepts of rank-index methods and of uniformity may be generalized to include minimum requirements, and the various conclusions of this section hold for the general case.

A related problem is that of *proportional representation*; here $\underset{\sim}{p}$ represents the votes of the parties and $\underset{\sim}{a}$ the allocation of seats. Consider a problem $(\underset{\sim}{p},h)$, where one party receives p^* votes and a^* seats, and another gets $\bar{p}$ votes and $\bar{a}$ seats. Suppose that before the seats are distributed these two parties coalesce into one party with vote total $p^* + \bar{p}$, the vote totals of all other parties staying the same. It is of interest to ask how the total number of seats the coalesced party now gets compares with $a^* + \bar{a}$. A symmetric method M is said to be *stable* if after coalescing there exists an allocation for h in which the coalesced party gets b seats and $a^* + \bar{a} - 1 \leqq b \leqq a^* + \bar{a} + 1$. This idea was apparently first proposed by Erlang (1907).

Proposition 8.7. A divisor method with divisor criterion d(a) is stable if and only if $d(a_1+a_2) \leqq d(a_1) + d(a_2) \leqq d(a_1+a_2+1)$.

Proposition 8.8. The five traditional divisor methods are stable.

Proposition 8.9. The method of Hamilton is stable.

A symmetric method M is said to *encourage coalitions* if after coalescing there exists an allocation of the seats satisfying $b \geqq a^* + \bar{a}$, and to *encourage schisms* if $b \leqq a^* + \bar{a}$. (See Erlang, 1907, Balinski and Young 1978b, 1979a).

Proposition 8.11. The method of Jefferson is the unique uniform balanced method that encourages coalitions.

Proposition 8.12. The method of Adams is the unique uniform balanced method that encourages schisms.

REFERENCES

M.L. Balinski and H.P. Young (1974) A New Method for Congressional Apportionment, Proceedings of the National Academy of Sciences, 71, 4602-4606.

———————— (1975) The Quota Method of Apportionment, American Mathematical Monthly, 82, 701-730.

———————— (1977a) Apportionment Schemes and the Quota Method, American Mathematical Monthly, 84, 450-455.

———————— (1977b) On Huntington Methods of Apportionment, SIAM Journal on Applied Mathematics - Part C, 33, No. 4, 607-618.

———————— (1978a) The Jefferson Method of Apportionment, SIAM Review, 20, No. 2, 278-284.

———————— (1978b) Stability, Coalitions and Schisms in Proportional Representation Systems, American Political Science Review, 72, No 3, 848-858.

———————— (1979a) Criteria for Proportional Representation, Operations Research, 27, No. 1, 80-95.

———————— (1979b) Quotatone Apportionment Methods, Mathematics of Operations Research, 4, No. 1, 31-38.

———————— (1980) The Webster Method of Apportionment, Proceedings of the National Academy of Sciences, 77, No. 1, 1-4.

G. Birkhoff (1976) House Monotone Apportionment Schemes. Proceedings of the National Academy of Sciences, U.S.A., 73, 684-686.

O.R. Burt and C.C. Harris, Jr. (1963) Apportionment of the U.S. House of Representatives: a minimum range, integer solution, allocation problem, Operations Research, 11, 648-652.

J.M. Cotteret and C. Emeri (1970) Les Systemes Electoraux. Paris: Presses Universitaires de France.

A.K. Erlang (1907) Flerfoldsvalg efter rene partilister. Nyt Tidskrift for Matematik, Afdeling B, 18, 82-83.

V. d'Hondt (1878) La Representation Proportionelle des Partis par un Electeur. Ghent.

——— (1882) Systeme pratique et raisonne de representation proportionnelle. Brussels: Muquardt.

E.V. Huntington (1921) The Mathematical Theory of the Apportionment of Representatives, Proceedings of the National Academy of Sciences, U.S.A., 7, 123-127.

——— (1928) The Apportionment of Representatives in Congress, Transactions of the American Mathematical Society, 30, 85-110.

A. Hylland (1975) En merknad til en artikkel fra 1907 om forholdsvalg. Nordisk Matematisk Tidskrift, 23, 15-19.

J.P. Mayberry (1978a) Quota methods for Congressional apportionment are still non-unique. Proceedings of the National Academy of Sciences, U.S.A., 75, 3537-3539.

——— (1978b) A Spectrum of Quota Methods for Legislative Apportionment and Manpower allocation. Broch University, St. Catharines, Ontario (mimeographed).

J.W. Still (1979) A class of new methods for Congressional apportionment, SIAM J. Appl. Math., 37, No. 2, 401-418.

Sainte-Lagüe (1910) Le representation et la methode des moindres carres. Comptes Rendus de l'Academie des Sciences, 151, 377-378.