International Game Theory Review
Vol. 15, No. 3 (2013) 1340014 (14 pages)

DOI: 10.1142/S0219198913400148

ON ALLOCATING SEATS TO PARTIES AND DISTRICTS: APPORTIONMENTS

GABRIELLE DEMANGE
PSE-EHESS, 48 Boulevard Jourdan
75014 Paris, France
demange@pse.ens.fr

Received 31 August 2011
Revised 15 December 2011
Accepted 18 March 2013
Published 1 July 2013

The paper presents the problem of choosing the representatives in an assembly when the whole electoral region is subdivided into several electoral districts. Because of the two dimensions, geographical (districts) and political (parties), the problem is called bi-apportionment. The main purpose of the paper is to discuss fairness and proportionality axioms and to describe their implementation.

Keywords: Party-proportional representation; power indices; (bi-)apportionment; (bi-)divisor methods; fair shares.

Subject Classification: D70, D71

1. Introduction

The representation problem is how to assign a fixed number of seats to different categories of a population as a function of some data. Most important representation problems, often called apportionments, arise in the political field under various forms. In one form, the seats of an assembly have to be distributed to the parties after an election according to some pre-specified electoral rule; here the data consists of the number of votes obtained by each party. Another form, though not always so explicit or transparent, pertains to the distribution of the seats among geographical areas such as districts, regions, cantons in Switzerland, or countries in the European Parliament for example; here the relevant data is described by the population sizes. A more complicated but common situation appears when the two dimensions, geographical and political, matter for determining the representation. Here the data consists of the number of votes obtained by each party in each district, and the seats of the assembly have to be distributed to the parties *and* the districts. This is called a bi-apportionment problem. This paper reviews both the uni- and bi-dimensional problems with a focus on proportional representation.

To illustrate the difficulties, first note that in the apportionment debate, no method establishes itself as a reference. In the uni-dimensional setting, the difficulty stems from the fact that seats are indivisible; as a result there are various ways to approximate an ideal proportional representation. In the bi-dimensional setting, the interaction between the constraints at the district level and the parties' results at the global level creates other difficulties. Some current methods have been proven to be 'bad'. For example, the Italian 'bug', as called by Pennisi [2006], refers to a serious flaw since the procedure itself is ill-defined and may lead to contradictory outcomes depending on the way it is implemented. Another example is the election of the Bundestag deputies in Germany where increasing the ballots for a party may induce a loss in its seats. Such a nonmonotone behavior is difficult to detect due to the complexity of the electoral rule (see Pukelsheim, 2006). The Italian and German cases are just two examples. In many countries, the procedures for the election of representatives at an assembly are cumbersome. They start by allocating seats district per district and then adjust the parties' seats in a complex way so as to approximate a proportional party representation. The complexity reflects the more or less explicit goal of satisfying two criteria: on the one hand, a proportional party representation and, on the other hand, decentralization, meaning that local candidates are chosen on the basis of the local votes. However, as we will discuss, proportional party representation and decentralization cannot be both satisfied for all electoral results, and a priority between the two criteria should be determined.

The paper is organized as follows. Section 2 recalls some approaches to apportionments in the uni-dimensional setting, with a focus on fairness and proportionality. Section 3 reviews the few studies on bi-dimensional apportionments and the methods that have been introduced to account for geographical and political dimensions simultaneously. These methods are based on a 'target', that we call 'fair share', which must be transformed into seats, i.e., integers. This can be performed by a 'controlled rounding' method as first illustrated by Gassner [1991] for the Belgian Senate election, or by a bi-divisor method as introduced by Balinski and Demange [1989a]. A bi-divisor method serves as a basis for the design of the New Apportionment Procedure adopted in various cantons in Switzerland starting with Zurich. The section ends up with some discussion and open questions.

2. Uni-Dimensional Apportionment

The uni-dimensional apportionment problem has been thoroughly studied. To fix the idea, let us introduce some notation. Let H be the total number of seats to be assigned in the house. Let us first describe apportionment problems in terms of parties. In a party-apportionment problem, the total number of seats has to be assigned to parties as a function of the number of votes they have received. Let m be the number of parties, v_j be the total votes for party j, and $\mathbf{v} = (v_1, \ldots, v_m)$ be the vector of all votes. A (party-)apportionment assigns the H seats to the parties. Denoting by s_j the number of seats — integer-valued — received by party j, an apportionment is represented by a vector $\mathbf{s} = (s_1, \ldots, s_m)$ where s_j is an integer (possibly null) and $\sum_j s_j = H$. A (district)-apportionment problem is obtained by considering districts instead of parties and by interpreting the v_j as the electoral population: the total number of seats has to be assigned to districts as a function of their population sizes. When unspecified, j is referred as a category.

A method specifies the apportionment as a function of the data: An *apportionment method* assigns to each possible electoral outcome $\mathbf{v}$ an apportionment $\mathbf{s}$.

In the case of parties, an apportionment method is simply an electoral rule. In the case of districts, the method is not always specified, but there is a trend towards greater transparency. In Europe for instance, as explained by Grimmett [2012], 'the Committee of Constitutional Affairs commissioned a Symposium of mathematicians to *identify a mathematical formula for the distribution of the seats which will be durable, transparent and impartial to politics*.'

A main approach to judge methods relies on 'axioms' or 'properties' that the method should reasonably satisfy. I present the main axioms considered in the literature that relate to some form of fairness. Some of these axioms are based on the (Hare)-quota. Let V be the sum of the votes, $V = \sum v_i$. The (Hare)-*quota* of j is defined by $q_j = \frac{v_j}{V} H$, and is interpreted as the number of seats that j should receive provided the seats were divisible and strict proportionality of the seats to category sizes was the goal of the apportionment.

2.1. *Fairness axioms*

- Nonreversal: If a category has a strictly smaller size than another, then it should not get more seats.
- Exactness: If the quotas are all integers, the apportionment is given by the quotas.
- Respect of quota: Each category should receive a number of seats equal to its quota rounded either up or down.
- Consistency (or Uniformity): Dropping some categories and the seats they obtain, the distribution of the seats to the remaining categories is unchanged.
- Population-Monotonicity: If, following an election, category i grows stronger than a category j, others being unchanged, category i does not lose a seat.
- House-Monotonicity: If the total number of seats H increases then no category should lose a seat.

Nonreversal does not need any comment. Exactness and respect of quota both rely on the premise that proportionality as embodied in the quotas is the 'ideal' or target. When this target can be achieved, exactness requires that it is the solution. When it cannot be achieved due to the indivisibility of seats, quotas should be rounded to an adjacent integer to approximate proportionality. Exactness is weaker than the property of respect of quota.

Consistency, sometimes called uniformity, is the key property in many fair division problems that asks that any part of a (fair) allocation to be itself a (fair) allocation.

The monotonicity properties state desirable properties of the method when some parameters, the size of the categories or the number of seats, vary. In a sense they state minimal fairness properties on the variations in the apportionments.

The first four axioms bear on a given apportionment. As such they can be readily checked; a reversal, for example, is observable at the apportionment in place. Nonmonotonicity properties instead are more difficult to check if one does not dwell into the mathematics. A nonmonotonic behavior is only observed after an adequate sequence in the change of the parameters, population or house sizes. Indeed the nonmonotonicity of some common methods was uncovered only when some 'paradoxes' occurred.[a]

The interpretation and relevance of the axioms in both the geographical and political contexts will be discussed after the presentation of some 'classical' methods.

2.2. *Some common methods*

Let us recall briefly two classes of methods that are exact, namely for which quotas are the target.

The first class is based on the natural idea of 'rounding' each quota to an adjacent number in such a way that the number of allocated seats is kept equal to H. By construction these methods respect the quotas. The largest remainder method is a well-known example. The description is simple. First, give every party its lower quota (i.e., round down), and second distribute the remaining seats to the parties that have the largest remainder left.

The second class is that of divisor methods. It was uncovered by Huntington [1921], who showed that five well-known methods were computed in the same way, each one associated with a specific rounding rule. Standard rounding for example, which rounds each number to the nearest integer, yields the so-called Webster or Sainte-Laguë's rule or Major fractions.[b] The method works as follows. Find a scale (or 'divisor') such that rounding and summing all scaled data yield the right total. Formally, a d-rounding method is characterized by thresholds, $d(n)$ in $[n, n+1]$ defined for each natural number n. Each number x in the interval $]d(n), d(n+1)[$ is rounded to n, and we denote $[x]_d = n$. Standard rounding[c] has $d(n) = n + 1/2$. The *divisor method based on* d associates to voting results $\mathbf{v}$ the apportionment(s) (typically unique) that satisfy: $\{\mathbf{s} = (s_i) \mid s_j = [\lambda v_j]_d$ for a scalar λ chosen so that $\sum_j s_j = H\}$.

[a] Paradoxes are often called Alabama, population and new state (or Oklahoma) paradoxes.

[b] A common method has many various names, associated to its proponents in a country (as Webster in the US), or to mathematicians who discovered it (as Sainte-Laguë'), or to its mathematical definition (Major fractions).

[c] The four alternative well-known divisor methods are Jefferson's (or D'Hondt's or Greatest divisors) based on $d(n) = n + 1$ (so that any real number $x \in (n, n+1]$ is rounded down to $n + 1$), Hill's method based on $d(n) = \sqrt{n(n+1)}$, Dean's method based on $d(n) = 2n(n+1)/(2n+1)$, and Adams' method based on $d(n) = n$ (up-rounding).

In a nutshell, the necessary adjustments due to the integer-valued requirement are made in proportion of the numbers in a divisor method and in absolute terms in a method that respects quotas. This explains why the divisor methods are monotone (in whatever sense) but not the methods that respect quotas. There is indeed a fundamental conflict between population-monotonicity and the respect of quota, as shown by Balinski and Young [1982]: no method satisfies both axioms.

In view of the incompatibility of these axioms, Gambarelli [1989] proposed to set priorities on the criteria. The criteria are applied in sequence until a single apportionment is obtained. The method is called Minimax because at each stage only the apportionments that minimize the largest differences among parties according to the specified criterion are kept. One possible drawback of this method is its computability/complexity because all apportionments have to be evaluated (at least at the first step when the first criterion is applied).

To summarize, assuming that exact proportionality as embodied in the quotas is a goal, a variety of methods are reasonable and no one is ideal. This lesson is confirmed by another approach based on optimization rather than axioms. The five main divisor methods and the largest remainders introduced above all minimize the distance to the quotas, each one with a different 'reasonable' distance.

2.3. *Discussion*

Even though the just presented methods differ and the differences in the produced outcomes play a determinant political role under some circumstances, they are all based on the same premise. Quotas constitute the benchmark, which is not feasible only because seats are indivisible. In particular, each method delivers the quotas when they are all integers (the exactness property). There are arguments against this premise, both in the geographical and political contexts.

In the political context, the constitution of a stable majority party is often the priority. Such a priority is executed in various ways, often indirectly, by imposing a minimal threshold for representation for example. Even for a country such as Netherlands, which aims at a proportional representation, both the minimal requirement and the divisor rule that most favors large parties — the Jefferson rule — are used. A relevant approach to assess whether proportionality is appropriate for a representative assembly depends on the whole institutional framework and the role devoted to this assembly. This is out of the scope of the paper and the next section will take as given the goal of the electoral rule, be proportionality or whatever other objective.

As for the geographical context, somewhat implicitly population-proportionality as described by the Hare-quota would achieve an equitable distribution of voting power of the citizens. However not only population-proportionality is most often far from being achieved but also there is still a lively debate about whether the Hare-quota is indeed a good measure of equity. For example, in the current discussion about the apportionment method for the European Parliament, the principle of 'degressive proportionality' is retained (see an account of the debate in the special issue around the Cambridge compromise edited by Laslier [2012]). A theoretical normative approach challenges the fairness property of proportionality. If fairness means that each individual has the same 'power', then Hare-quotas may be unfair. Penrose [1946] already expresses this idea and justifies an apportionment in which the seats are allocated proportionally to the square root of the population. Let us describe the basic argument, so as to make clear the underlying assumptions. Assume the representatives of a district (country) vote as a 'block'. In other words, these representatives do not represent the possible diversity in the opinions *within* their district. Thus each district in the assembly acts as a single player with a voting weight proportional to the number of seats it has been assigned to. 'Power' is defined as the chance of being pivotal, that is, reflects the power to change a decision.[d] The key point is that, under majority rule (and under some assumptions on the preferences on the issues) the decisiveness 'power' is increasing more steeply than the voting weight. As a result, allocating seats in proportion to size yields disproportionally more power to citizens in districts with large size. Under the statistical assumptions made by Penrose [1946], seats should be allocated proportionally to the square root of the population; this provides a fair benchmark that differs from the Hare-quotas.

This line has been followed up recently by promoting the use of power indices to help to design fair rules. A 'fair' rule is defined as one that gives an equal power to each citizen. It should be clear that this approach crucially depends on the chosen index. Power is linked with decisiveness, that is the chances of being pivotal. These chances depend on the representatives' voting behavior, both within a district (block vote or not) and across districts through the extent of correlation in their votes (see Widgren, 2005 for a survey). Power also depends on the electoral rule in the assembly since the chances of being pivotal depend on the set of 'winning' coalitions. To sum up, power indices may incorporate various important features. If these features indeed hold and are stable overtime, then the rule that gives an equal power to each citizen provides the fair benchmark (instead of the quota). Nonetheless, there still remains to convert this benchmark into integers in order to obtain an apportionment. An inspection of the axioms reveals that they make sense unchanged, but for exactness or respect of the quotas where quotas have to be replaced by the fair benchmark. The various rounding methods apply. I do not know if any previous study exists, but the basic conflict between population-monotonicity and respect of the benchmark is likely to be still present.

3. Bi-Apportionment

In many elections for choosing representatives in an assembly, the electoral body is divided into several electoral districts. The results obtained in each district matter

[d]This is similar to the model underlying the well-known Banzhaf index, now called the Penrose–Banzhaf index by some authors.

for electing the representatives of that district in the assembly, but at the same time, the elections in the districts cannot be viewed as separate elections. These representatives, who will all gather in the assembly, are affiliated to parties in most circumstances. So the outcome not only determines the representatives but also the strength of the parties and ultimately, in some countries, the prime minister. This explains why some electoral laws account for both local and global levels. Specifically the numbers of votes are distinguished by district and party. This is called a bi-apportionment problem.

Formally, let H be the total number of seats to be assigned in the house, n be the number of districts, m the number of parties. The result of an electoral outcome is now described by a $n \times m$ matrix $\mathbf{v} = (v_{ij})$ in which row i represents district i, column j party j, and v_{ij} the votes obtained by j in district i. To simplify the presentation, the matrix $\mathbf{v}$ is assumed to have all its elements positive (otherwise additional conditions are needed for the methods to be well-defined).

The district-apportionment is fixed prior to the election, specified by the number of representatives that each district is entitled to, h_i for district i. The district-apportionment is thus described by $\mathbf{h} = (h_1, h_2, \ldots, h_n)$, a n-tuple of positive integers which sum to H, $H = \sum h_i$. Now, given an electoral outcome $\mathbf{v}$, the electoral rule assigns seats to parties in each district so as to satisfy the district constraints. Let b_{ij} represent the number of seats received by j in district i.

Definition 1. Let the district-apportionment $\mathbf{h}$ be given. A *bi-apportionment* is represented by a matrix $\mathbf{b} = (b_{ij})$ where the b_{ij} are non-negative integer and satisfy the district constraints: $\sum_j b_{ij} = h_i$ for each i. A *bi-apportionment method* $\mathbf{A}$ assigns to each voting matrix $\mathbf{v}$ a bi-apportionment.

There has been so far few studies on bi-apportionment methods. Of course, there is the possibility of full 'decentralization', under which an apportionment method is used separately in each district. But there are various arguments in favor of accounting for the overall totals obtained by parties. Accounting simultaneously for the districts' constraints and the parties' results makes the problem truly bi-dimensional.

3.1. *Two-step procedures: priority to aggregate results*

We present first two arguments against full decentralization.

When each district has a single representative (or a small number), separate elections may lead to rather extreme outcomes and drastically diminish the representation of intermediate parties due to the rounding effects. Furthermore, party-proportionality cannot be achieved by allocating seats district per district when there are distortions in the district apportionment. This result is independent of any integer-valued requirement (see Demange, 2012). The difficulty lies at the root of complex additional features in electoral rules that aim at achieving party-proportionality. For example, there are rules that allow for a variable number of seats, as for the Faroese Parliament (Zachariassen and Zachariassen, 2006) or the Bundestag in Germany with the 'overhang' seats (Pukelsheim, 2006). Without entering into details, these rules first allocate seats to the parties on a district basis; second, if the obtained apportionment is too 'unfair' for some parties, then additional seats are distributed so as to mitigate unfairness.

Another argument in favor of accounting for national parties' votes to some extent arises when some districts have very few representatives relative to others. Citizens in small districts in favor of a minor party are almost sure to have no influence on the final outcome in case of separate calculations while it might not be true in larger districts. Equity among citizens then requires an adjustment mechanism that takes into account all the votes in favour of a party, as called for by the Swiss federal court in 2002. As a result, the electoral rule use the votes to the parties at the national/aggregate level.

In what follows, whatever method, the priority is given to the votes to the parties at the national/aggregate level. This is achieved by carrying out the method in two steps. In the first step, the H seats are apportioned to the parties on the basis of their overall vote totals: $\mathbf{s} = (s_j)$ would be the result of a national election if districts did not matter.

The second step computes a bi-apportionment assigned to the votes $\mathbf{v}$ under the district and party constraints given by $\mathbf{h}$ and $\mathbf{s}$: each district i is entitled to h_i seats and each party j to s_j seats.

The chosen (uni-)apportionment method in the first step allows to achieve the desired properties on the parties' representation. If party-proportionality is the priority,[e] a 'proportional' method should be chosen among the ones described in Sec. 2. If party-proportionality is not the priority, another apportionment method may be chosen to determine this first party-apportionment. The second step goes through producing a bi-apportionment that satisfies the required properties on party's representation and meets the pre-defined district requirements.

The second step can be stated as follows: given $(\mathbf{v}, \mathbf{h}, \mathbf{s})$ how to transform $\mathbf{v}$ into integers so as to meet the constraints. This problem differs from the uni-dimensional problem. Even without the integer-valued requirement, it is not obvious what the solution should be. So we first start by considering the (bi)-allocation problem, i.e., the (bi)-apportionment problem where the integer assumption is relaxed.

3.2. *The fair share benchmark*

Given $(\mathbf{h}, \mathbf{s})$ the feasible *bi-allocations* are represented by $\mathbf{a} = (a_{ij})$, non-negative, that satisfy

$$\sum_j a_{ij} = h_i \text{ for each } i, \sum_i a_{ij} = s_j \text{ for each } j. \tag{1}$$

[e]Gassner [1991], motivated by the severe drawbacks of the Belgian electoral law, expresses the idea of apportioning seats first at the global level. This idea is used in the New Apportionment Procedure applied in Zurich. In both case proportionality is the goal in both dimensions.

Feasible bi-allocations exist because $\sum_i h_i = \sum_j s_j = H$. Among them, the fair share matrix is a good candidate to represent the idea of proportionality to $\mathbf{v}$ while accounting for *a priori* constraints. It is obtained by multiplying rows and columns by appropriate multipliers so as to satisfy the constraints:

Definition 2. The fair share (matrix) to problem ($\mathbf{v}$, $\mathbf{h}$, $\mathbf{s}$) is the unique matrix $\mathbf{f}$ of the form ($f_{ij} = \lambda_i v_{ij} \mu_j$) that matches the row and column sums (1).

The fair share method, which assigns its fair share to a problem ($\mathbf{v}$, $\mathbf{h}$, $\mathbf{s}$), is characterized by three axioms, exactness, homogeneity, and consistency, as shown by Balinski and Demange [1989a].

The fair share solves the problem of adjusting a matrix so as to satisfy specified values for the rows and columns totals. The adjustment is used in a wide variety of contexts, for example in statistics for adjusting contingencies tables or in economics for balancing international trade accounts (in the RAS model, see e.g., Bacharach, 1965). It is sometimes called bi-proportional matrix, but bi-proportionality may introduce some confusion as it may suggest that the row and column sums are proportional apportionments respectively for districts and parties. The fair share can be defined in the more general setting in which row and column sums are only constrained to belong to some intervals rather than being assigned specific values. The properties of the fair share method described above, exactness, homogeneity, and consistency, extend (Balinski and Demange, 1989a).

3.3. *Bi-apportionments based on fair shares*

The fair share method provides a natural benchmark or target, which constitutes the basis for bi-apportionment methods. Recall that a bi-apportionment must be integer-valued. Starting from the fair share benchmark, a solution would be to 'round' its elements, say to the nearest integer (standard method). However, the obtained matrix may not meet the row and column totals (the same is true for any specified rounding method). Two main approaches have been followed, which parallel those used in a uni-dimensional setting.

The first approach is based on the idea of 'rounding' each element of the fair share matrix to an adjacent number in some way while still satisfying the desired constraints on the row and column sums. (The rounding may differ across elements, that is, no rounding method is pre-specified.) It should be first noted that there are bi-apportionments that 'respect fair shares'. This is due to the specific structure of the linear system described by (1) which makes all extreme points integer-valued when the $\mathbf{h}$ and the $\mathbf{s}$ are integer-valued.[f] The electoral rule however should assign a unique bi-apportionment (in most circumstances). This can be accomplished, for example, by an optimization procedure known as *controlled rounding* and developed

[f]Such a result is often referred to as Birkhoff theorem. A well-known example is the matching or assignment game where $n = m$ and all the components of $\mathbf{h}$ and $\mathbf{s}$ are equal to 1.

by Cox and Ernst [1982] (see also Gassner, 1991 who applies an alternative rounding method).

The second approach performs simultaneously the rounding and the adjustment through scale factors (divisors). This leads to a variety of *bi-divisor methods*, each one characterized by a distinct rounding method, as introduced and characterized in Balinski and Demange [1989a]. The idea is to use divisors as in the uni-dimensional case, but now there is one for each constraint.[g] The adjustment through the multipliers and the rounding are 'simultaneous' so that the bi-apportionment is not necessarily a rounding of the fair share. Formally, just as in the one-dimensional case, bi-dimensional divisor methods are based on a d-rounding method (recall that such a method assigns to each scalar x a (almost unique) rounded value $[x]_d$). One looks for multipliers λ_i for district i, μ_j for party j such that d-rounding each element of the matrix $(\lambda_i v_{ij} \mu_j)$ is a bi-apportionment, namely the constraints on row- and column-totals are met. The obtained bi-apportionment exists and is typically unique.[h] Formally a *bi-divisor method* (*based on* d) assigns to any positive problem $(\mathbf{v}, \mathbf{h}, \mathbf{s})$ the bi-apportionment(s) $\mathbf{b}$ (typically unique) that satisfy

$$\mathbf{b} = (b_{ij}) \,|\, b_{ij} = [\lambda_i v_{ij} \mu_j]_d, \quad \sum_j b_{ij} = h_i \text{ each } i \text{ and } \sum_i b_{ij} = s_j \text{ each } j. \qquad (2)$$

A bi-divisor method satisfies exactness, monotony, consistency, proportionality. Furthermore, if $\mathbf{s}$ is obtained by applying a proportional apportionment method to the global parties votes, the bi-divisor method satisfies party-proportionality.

The New Apportionment Procedure developed by Pukelsheim [2006] and adopted by several cantons in Switzerland is based on a bi-divisor method. Algorithms have been designed[i] to compute a bi-apportionment, starting with the Tie and Transfer algorithm, which uses a formulation in terms of transportation flows (Balinski and Demange, 1989b). Alternative algorithms are compared in Maier *et al.* [2010]. For a systematic treatment of the problem in terms of network flows, see Pukelsheim *et al.* [2011].

As far as I know, there are no studies comparing the various bi-apportionment methods we have just described (see, however, the analysis of Zachariassen and Zachariassen, 2006 applied to the Faroese Parliament). Intuitively the outcomes should not differ that much. Once the party and district apportionments are fixed, there is less flexibility than in the general uni-dimensional problem (in particular, restricting the problem to a single district, the second step becomes a vacuous

[g]The uni-dimensional setting considered in Sec. 2 has only one overall constraint corresponding to the total number of the seats. When there are additional constraints, say a minimum number of seats per districts, additional divisors are introduced.

[h]More precisely, this is true for divisor methods that assign no seat to small enough numbers, i.e., that satisfy $d(0) > 0$ where d describes the thresholds. This is a reasonable assumption since otherwise, for $d(0) = 0$, each party should receive at least one seat in each district.

[i]'BAZI', A Free Computer Program for Proportional Representation provides useful programs at http://www.math.uni-augsburg.de/stochastik/bazi/welcome.html.

problem). The standard comparison between Jefferson's and Adam's methods for example does not apply here.

Gambarelli and Palestini [2007] recently extend the Minimax method to bi-apportionments (for which the district apportionment is fixed). Recall that the Minimax method is based on a priority ordering of criteria. In particular, criteria bearing on the global votes can be put in top of the list, as is recommended by Gambarelli and Palestini [2007]. This is in the same spirit as fixing the party-apportionment as a function of the global votes.

Finally, when party-proportionality is not the goal, the same methodology applies if the party-apportionment is determined by the total votes for each party through a nonproportional method. Given the values for the district- and party-apportionments, the **h** and the **s**, one computes the associated fair share. The various ways for rounding the cells in the matrix and their properties apply. Hence the methodology carries over whatever district- and party-apportionments.

3.4. *Discussion and open problems*

I discuss first the issue of the complexity of a bi-dimensional method and the rationale for district distortions.

One may worry about the complexity of the bi-dimensional methods for voters. A method that is truly bi-dimensional involves handling matrices. Furthermore, the rounding issue adds another difficulty in computing the outcomes. Voters should be able to check the outcome following an election. As such, both types of methods, although 'transparent' in theory, may appear quite obscure to many voters or politicians, who cannot compute the outcome without an adequate software. There is nevertheless a difference between the controlled rounding and the bi-divisor methods. For the latter, once the multipliers are made public, the outcome can be checked by hand. For the controlled rounding method, which minimizes some distance measure to the 'target', this is not the case. Recently Serafini and Simeone [2011a, 2011b] have proposed to use a different distance measure for rounding the target. The computation of the optimal apportionment relies on flow techniques and the Max Flow-Min Cut Theorem produces 'certificates' that allow voters to check the optimality of the outcome.

In the previous section, the district-apportionment is taken as 'given'. The party-apportionment is then determined, say to obtain a proportional party-representation by applying a (uni-dimensional) proportional method on global parties' results. The interaction between the constraints on district- and party-apportionments, the **h** and the **s**, has not been tackled yet, as far as I know. This issue deserves some discussion. Consider the main theoretical argument against proportionality in district-apportionments, which justifies the Penrose square root law (see Sec. 2.3). It is based on the fact that the citizens' preferences in a district are not well represented because their representatives vote in 'block'. This argument is no longer valid when citizens vote for parties. Since they express their preferences related to the 'general issues' handled by parties, citizens are represented not only by their district's representatives but also by their party's representatives. The extent to which districts with small population sizes should be favored should depend on the type of issues handled by the 'representative' assembly. The question is worth studying.

Representation problems abound outside politics, in business for designing the board of directors, in schools or universities for the constitution of assemblies representing the various bodies — students, parents, teachers, administrative — in international institutions such as IMF, ONU for representing countries. In these frameworks, institutions often use more flexible rules than in politics. In a uni-dimensional setting, a simple way to approximate proportionality (assuming it is a target) is to assign weights to representatives, hence avoiding difficulties due to the indivisibility of seats. Weighted voting is used in international institutions such as the International Monetary Fund (each Member State receives a weighted vote proportional to its contribution to the Fund) or the World Bank (each member receives a fixed part plus a part proportional to its shares in the Bank). In publicly owned companies, most often, shareholders' votes are proportional to their shares. In condominium, owners' votes are functions of the size of their flats. In politics, the European Council also uses weighted voting. All these contexts consider a single category. It is unclear how to extend weighted voting to a bi-dimensional context, especially if one wants to favor some categories.

Studies on bi-dimensional representation have been so far mostly restricted to situations in which each person casts a single vote. In the context of a scientific association, however, Brams [1990] proposed to use approval voting, in which each person casts a list of acceptable candidates. The association wanted to find a way to achieve a more equitable representation of regions and specialties, making the problem bi-dimensional. Using approval voting in a bi-dimensional political setting needs more investigation.

Finally, in some problems, more than two characteristics are relevant for the representation. Let us illustrate with an example. For the Parliament of the Federation of Bosnia and Herzegovina, Hylland [2000] proposed a representation of the population of the 10 Cantons and the three 'Constituent people', Bosniacs, Croats and Others. Such a two-dimensional representation is fixed before the elections, and plays the role of the district-apportionment in our paper. Adding the parties' dimension makes the problem three dimensional. Some of the results obtained with two dimensions do not carry over to more than two dimensions.[j] This opens up challenging questions.

[j] Controlled rounding, for example, relies on Birkhoff theorem, which is valid in two dimensions only.

Acknowledgments

I am very grateful to Margaux Luflade for research assistance. I also would like to thank the referees and the editor G. Gambarelli.

References

Bacharach, M. [1965] Estimating non-negative matrices for marginal data, *Int. Econ. Rev.* **6**, 294–310.

Balinski, M. L. and Demange, G. [1989a] An axiomatic approach to proportionality between matrices, *Math. Oper. Res.* **14**, 700–719.

Balinski, M. L. and Demange, G. [1989b] Algorithm for proportional matrices in reals and integers, *Math. Program.* **45**, 193–210.

Balinski, M. L and Young, H. P. [1982] *Fair Representation: Meeting the Ideal of One Man, One Vote*, 1st edn. (Yale University Press).

Brams, S. [1990] Constrained approval voting: A voting system to elect a governing board, *Interfaces* **20**, 65–79.

Cox, L. H. and Ernst, L. R. [1982] Controlled rounding, *INFOR* **20**, 423–432.

Demange, G. [2012] On party-proportional representation under district distortions, *Math. Soc. Sci.* **63**, 181–191.

Gambarelli, G. [1989] Minimax apportionments, *Group Decis. Negot.* **8**(6), 441–461.

Gambarelli, G. and Palestini, A. [2007] Minimax multi-district apportionments, *Homo Oeconomicus* **24**, 335–356.

Gassner, M. [1991] Biproportional delegations: A solution for two-dimensional proportional representation, *J. Theor. Polit.* **3**, 21–342.

Grimmett, G. [2012] European apportionment via the Cambridge Compromise, *Math. Soc. Sci.* **63**, 68–73.

Huntington, E. V. [1921] The mathematical theory of apportionment of representatives, *Proc. Nat. Acad. Sci. U.S.A.* **7**, 123–127.

Hylland, A. [2000] Proportionality in two dimensions. Applied to the representation of cantons and constituent peoples in the House of Peoples of the Parliament of the Federation of Bosnia and Herzegovina, available at http://www.sv.uio.no/econ/personer/vit/aanundh/upubliserte-artikler-og-notater/Prop-2d%5B1%5D.pdf.

Laslier, J. F. [2012] Around the Cambridge compromise: Apportionment in theory and practice, special issue (ed J. F. Laslier), *Math. Soc. Sci.* **63**, 65–191.

Maier, S., Zachariassen, P. and Zachariassen, M. [2010] Divisor-based biproportional apportionment in electoral systems: A real-life benchmark study, *Manag. Sci.* **56**, 373–387.

Pennisi, A. [2006] The Italian bug: A flawed procedure for bi-proportional seat allocation, in *Mathematics and Democracy: Recent Advances in Voting Systems and Collective Choice*, eds. Simeone, B. & Pukelsheim, F. (Springer), pp. 151–165.

Penrose, L. S. [1946] The elementary statistics of majority voting, *J. R. Stat. Soc.* **109**, 53–57.

Pukelsheim, F. [2006] Current issues of apportionment methods, in *Mathematics and Democracy: Recent Advances in Voting Systems and Collective Choice*, eds. Simeone, B. & Pukelsheim, F. (Springer), pp. 167–176.

Pukelsheim, F., Ricca, F., Scozzari, A., Serafini, P. and Simeone, B. [2011] Network flow methods for electoral systems, *Networks*, in press, doi: 10.1002/net.20480.

Serafini, P. and Simeone, B. [2011a] Parametric maximum flow methods for minimax approximation of target quotas in biproportional apportionment, *Networks*, in press, doi: 10.1002/net.20434.

Serafini, P. and Simeone, B. [2011b] Certificates of optimality: The third way to biproportional apportionment *Soc. Choice Welfare*, in press, doi: 10.1007/s00355-010-0528-8.

Widgren, M. [2005] Power in the design of constitutional rules, in *Group Formation in Economics: Networks, Clubs, and Coalitions*, eds. Demange, G. & Wooders, M. (Cambridge University Press), pp. 312–334.

Zachariassen, P. and Zachariassen, M. [2006] A comparison of electoral formulae for the Faroese Parliament, in *Mathematics and Democracy: Recent Advances in Voting Systems and Collective Choice*, eds. Simeone, B. & Pukelsheim, F. (Springer), pp. 235–251.