



HAL
open science

On allocating seats to parties and districts: apportionments

Gabrielle Demange

► **To cite this version:**

Gabrielle Demange. On allocating seats to parties and districts: apportionments. 2011. halshs-00644439

HAL Id: halshs-00644439

<https://shs.hal.science/halshs-00644439v1>

Preprint submitted on 24 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

WORKING PAPER N° 2011 – 36

On allocating seats to parties and districts: apportionments

Gabriel Demange

JEL Codes: D70, D71

Keywords: Party Proportional Representation, Power Indics, (Bi-) Apportionment, (Bi-) Divisor Methods, Fair Shares



PARIS-JOURDAN SCIENCES ÉCONOMIQUES

48, Bd JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

On allocating seats to parties and districts: apportionments

Gabrielle DEMANGE

PSE-EHESS, address 48 bd Jourdan, 75014 Paris, France

e-mail demange@pse.ens.fr

<http://www.parisschoolofeconomics.eu/fr/demange-gabrielle/>

November 21, 2011

Abstract

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

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

JEL 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 arise in the political field under various forms, and are often called apportionments. For example 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 (districts, regions, cantons in Switzerland, countries in the European Parliament for example); here the relevant data are the population sizes. A more complicated but common situation arises when the two dimensions, geographical and political, matter for determining the representation. This paper deals mainly with this two-dimensional problem, called bi-apportionment.

To make a sense of the difficulties, first note that in the apportionment debate, no method establishes itself as a reference. However some methods prove 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. The causes of such phenomenon are due to the complexity of the procedure, and cannot be explained in a few sentences (see Pukelsheim, [2006]). The Italian and German cases are just two examples; in many countries, the electoral rule for choosing the representatives in an assembly is cumbersome. The complexity often reflects the more or less explicit goal of satisfying distinct criteria which may be conflicting, meaning that they cannot be satisfied simultaneously for all electoral results. For instance, as we will argue, proportional representation of the parties is in conflict with the power of local votes on the local candidates, at least when the allocation of the seats to the districts is biased, which is the most common situation.

The paper is organized as follows. Section 2 recalls some approaches to apportionment 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 both 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 ‘controlled rounding’ as first illustrated by Gassner [1991] for the belgian Senate election, or by bi-divisor methods as introduced by Balinski and Demange [1989-a]. 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 as a function of the number of votes received by the parties, as is performed by an electoral rule. Let m be the number of parties, v_j be the total votes for party j , and $\mathbf{v} = (v_1, \dots, v_m)$ 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 $\mathbf{s} = (s_1, \dots, s_m)$ 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. When unspecified, j is referred to 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 more transparency. In Europe for instance, as explained by Grimmett [2011], ‘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. 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

- Non-reversal : 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.

Non-reversal 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 familiar and 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. Non-monotonicity properties

instead are more difficult to check if one does not dwell into the mathematics. A non-monotonic behavior is only observed after an adequate sequence in the change of the parameters, population or house sizes. Indeed it is only when some 'paradoxes' occurred that the non-monotonicity of some common methods was revealed.¹

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

It turns out that there is a fundamental conflict between population-monotonicity and the respect of quota, as shown by Balinski and Young [1982]. This result shows how strong the integrality requirement is. In view of this result, no method is singled out. Let us recall briefly two classes of methods that are exact, namely for which quotas can be seen as the ideal.

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 . The most well-known method in this class is the largest remainders. 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. By construction these methods respect the quotas.

The second class, divisor methods, 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.² Thresholds serve as a basis to define thresholds for rounding: Standard rounding for example, which rounds each number to the nearest integer, yields the so-called Webster rule. The method works as follows: Find a scale (or 'divisor') such that rounding all scaled data and summing yields the right total. As a result, the necessary adjustments due to the integer requirements are made in proportion of the numbers, and not in absolute terms as in a method that respect the quotas. This explains why the divisor methods are monotone (in whatever sense) and not the rounding-the-quota methods.

In view of the incompatibility of these axioms, Gambarelli [1999] proposes to set priorities on the criteria. The criteria are applied in sequence until a single apportionment is obtained. (For example one may start with the apportionments that present no-reversal). The method is called Minimax because at each stage only the apportionments that minimize the largest differences between parties according to the specified criterion are kept. One possible drawback of this method is its computability/complexity because one needs to evaluate all apportionments (at least when the first criterion is applied). Computability will be addressed more generally in the next section.

¹Paradoxes are often called Alabama, population and new state (or Oklahoma) paradoxes.

²Formally, a divisor method is characterized by thresholds values $d(n) \in [n, n + 1]$ for each natural number n . Each number in the interval $]d(n), d(n + 1)[$ are rounded to n . The *divisor method (based on d)* associates to voting results \mathbf{v} the apportionments that satisfy given by: $\{\mathbf{s} = (s_i) | s_i = [\lambda v_j]_d \text{ for a } \lambda \in \mathbb{R} \text{ chosen so that } \sum_j s_j = H\}$. Generically the apportionment is unique.

To summarize, assuming that exact proportionality as embodied in the quotas is a goal, there is no ideal method but there are many different reasonable methods. This lesson is confirmed by another approach based on optimization rather than axioms. The five main divisor methods and the largest remainders introduced below 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 a benchmark, which is not feasible only because seats are indivisible. In particular, each method delivers the quotas when these are all integers (the exactness property). There is a fundamental critic to this premise, which is independent of the integrality requirement. The critic takes very different forms in the geographical and political contexts. I start with the political one.

In many countries, stability and the constitution of a majority party is 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 a 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 depends on the whole institutional framework and the role of the representative assembly we are considering. 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 criterion.

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 Grimmet [2011]). 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 chances of being pivotal, that is the power to change a decision, hence linked with decisiveness.³ The key point is that, under

³This is similar to the model underlying the well-known Banzhaf index, now called the Penrose-Banzhaf index, by

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 disproportionately 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, which now is the benchmark instead of 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, which in turn depend on the assumptions one makes on representatives' voting behavior, both within a district (block vote or not) and across districts such as the possible correlations of the 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, rules using power indices may incorporate important features that make them more fair, under the important proviso that the underlying modeling assumptions hold and are stable. Nonetheless, now that a benchmark has been defined, there still remains to convert it into integers for obtaining an apportionment. An inspection of the axioms reveal that they still make sense, but for exactness or respect of the quotas where quotas are replaced by the relevant benchmark. The various rounding methods apply. I do not know any study, 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 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 in most circumstances affiliated to parties. So the outcome not only determines the representatives but also the strength of the different parties and ultimately, in some countries, the prime minister. This explains why some electoral laws account for both the local and global levels.

The difference with the analysis in the previous section is that the number 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 defining the methods).

some authors.

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 $\mathbf{h} = (h_1, h_2, \dots, h_n)$, a n -tuple of positive integers that 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 for taking into account also of the global results, which make 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, if party-proportional representation is the goal, a simple fact is that 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 requirement (see Demange [2011]). This difficulty is 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 allocate the seats on a district basis to the parties and then may add extra seats if the parties apportionment is too ‘unfair’.

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 of citizens then would require for some adjustment mechanism, as called for by the Swiss federal court in 2002. As a result, some electoral rules take into account the party-representation at the national/global level and aims at achieving some amount of proportional representation.

In what follows, whatever method, the priority is given to results at the 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. When party-proportionality is a goal,⁴ a 'proportional' method should be chosen among the ones described in Section 2. If party-proportionality is not a goal and another apportionment method is chosen to determine this first party-apportionment, the second step goes through. The only difference is that the bi-apportionment will satisfy some properties that differ from party-proportionality but that will meet other pre-defined 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 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

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

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

The fair share corresponds to a common solution for adjusting a matrix used in various areas such as 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

⁴Gassner [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 method applied in Zurich. In both case proportionality is the goal in both dimensions.

column sums are proportional apportionments respectively for districts and parties. In addition the fair share can be defined in the more general setting in which row and column sums may be only constrained to belong to some intervals rather than being assigned some values. The properties of the fair share method described above extend (Balinski and Demange [1989-a]).

3.3 Bi-apportionments based on fair shares

Fair shares provide 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 to some extent 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 method is not pre-specified and may differ across elements. It should be first noted that there are indeed bi-apportionments that ‘respect fair shares’. This is due to the special 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.⁵ For an electoral rule the outcome should be determinate. One can use for example a well-defined procedure, known as *controlled rounding*, as developed 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 [1989-a]. The idea is to use divisors as in the uni-dimensional case, but now there is one for each constraint.⁶ 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 function. One looks for multipliers λ_i for district i , μ_j for party j such that *d*-rounding each element of the matrix with general element $\lambda_i v_{ij} \mu_j$ is a bi-apportionment, namely the constraints on row- and column-totals are met. The obtained bi-apportionment exists and is typically unique.⁷ Formally a *bi-divisor method (based on d)* assigns

⁵Such 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.

⁶The uni-dimensional setting considered in Section 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.

⁷More precisely, this is true for divisor methods that assign no seat to small enough numbers, i.e., that satisfy $d(0) > 0$ where d is defined as in footnote 2. This is a reasonable assumption since otherwise, for $d(0) = 0$, each party should receive at least one seat in each district.

to any positive problem $(\mathbf{v}, \mathbf{h}, \mathbf{s})$ the bi-apportionment(s) \mathbf{b} (typically unique) that satisfy

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

A bi-divisor method satisfies exactness, monotony, uniformity, 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 to compute the bi-apportionment, starting with the Tie and Transfer algorithm using a formulation in terms of transportation flows (Balinski and Demange [1989-b]). Alternative algorithms are defined and compared in Maier, Zachariassen, and Zachariassen [2010].⁸

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

Gambarelli and Palestini [2007] recently extends the Minimax method to bi-apportionments (for which the district apportionment is fixed). Recall that the minimax method is based on an ordering of priority on criteria. In particular, as the authors recommend, criteria bearing on the global votes can be applied first. This is in the same spirit as fixing the party-apportionment.

Finally, when the 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 non-proportional method. Given the values for the district- and party- apportionments, the \mathbf{h} and the \mathbf{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 and furthermore the rounding issue adds another difficulty in computing the outcomes. Voters should be able to check the outcome following an

⁸For a systematic treatment of the problem in terms of network flows, see Pukelsheim *et al.* [2011]. 'BAZI', A Free Computer Program for Proportional Representation provides useful programs at <http://www.math.uni-augsburg.de/stochastik/bazi/welcome.html>.

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 [2011-a] and [2011-b] have proposed to use a different distance measure for rounding the target. The computation of the optimal apportionment relies on flow techniques and some ‘certificates’, produced by the Max Flow-Min Cut Theorem, allow voters to check the optimality of the outcome.

In the previous section, the district-apportionment is taken as ‘given’. The party-apportionment was then determined, say to achieve proportional party-representation (modulo the variations in the proportional methods) or another goal. The interaction between the constraints on district- and parties- apportionments, the \mathbf{h} and the \mathbf{s} has not been tackled, 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. 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, hence express their preferences related to ‘general issues’ handled by parties. Citizens are represented not only by their district representatives but also by their parties’ representatives. This suggests that the rationale for favoring districts with small population sizes is much weaker. The answer should depend on the type of issues handled by the ‘representative’ assembly, but in any case is worth studying.

The representation problem, prominent in the setting of electoral rules, arises in many other contexts as well. Still in the political domain, Hylland [2000] suggests a representation for the Parliament of the Federation of Bosnia and Herzegovina. There are ten Cantons and three ‘Constituent people’, Bosniacs, Croats and Others. Here the matrix gathers the population of the three Constituent people (instead of parties) in the ten cantons according to the census.

Representation problems also 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 world 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, in which each Member State receives a weighted vote proportional to its contribution to the Fund (which is determined by its economy) or the World Bank (members received a fixed part plus a part proportional to their shares in the Bank). In most publicly owned companies, shareholders have votes proportional to their shares. In condominium, owners votes are functions of the size of their

flats etc. 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.

We have restricted our analysis to a somewhat poor set of data, in which each person casts a single vote. In the context of a scientific association, Brams [1990] proposed to use approval voting. The association wanted to find a way to achieve a more equitable representation of regions and specialties, making the problem bi-dimensional. Extending such a procedure (i.e., using approval voting) to a political setting needs to be investigating.

Finally a further development would be to investigate the more general following formulation of the representation problem. Given a partition of a population into categories, find a set of representatives that meets some constraints or achieves some goal. In the presented bi-dimensional setting, a category is characterized by the pair district-party with constraints on the totals in each dimension. Considering more than two dimensions is worth investigating.⁹

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

References

- Bacharach, M. [1965] ‘Estimating Non-Negative Matrices for Marginal Data’. *International Economic Review* 6(3): 294–310
- Balinski, M.L. and Demange, G. [1989-a] ‘An Axiomatic Approach to Proportionality between Matrices’. *Mathematics of Operations Research* 14: 700–719.
- Balinski, M.L. and Demange, G. [1989-b]. ‘Algorithm for Proportional Matrices in Reals and Integers’. *Mathematical Programming* 45: 193–210.
- Balinski, M.L. and Young, H. P. [1982]. *Fair Representation: Meeting the Ideal of One Man, One Vote* (1st ed.). Yale Univ Press.
- Brams, S. [1990]. ‘Constrained Approval Voting: A Voting System to Elect a Governing Board,’ *Interfaces* 20 (September-October): 65-79.
- Cox, L.H. and Ernst, L.R. [1982]. ‘Controlled Rounding’. *INFOR* 20: 423–432.
- Demange, G. [2011]. ‘On party-proportional representation under district distortions’ *Mathematical Social Sciences* doi:10.1016/j.mathsocsci.2011.10.002.
- Gambarelli G. [1989]. *Minimax Apportionments Group Decision and Negotiation*, 1999 - Springer.
- Gambarelli G. and Palestini A. [2007]. ‘Minimax Multi-District Apportionments’ *Homo Oeconomicus* 24 (3/4): 335–356.

⁹Some of the results however do not carry over. Controlled rounding for example relies on Birkhoff theorem, which is valid in two dimensions only.

- Gassner M. [1991]. ‘Biproportional Delegations: A Solution for Two-Dimensional Proportional Representation’. *Journal of Theoretical Politics* 3(3):321–342.
- Grimmett G. [2011]. ‘European Apportionment via the Cambridge Compromise’. Available at <http://www.statslab.cam.ac.uk/~grg/papers/USep3.pdf>.
- Huntington E.V. [1921]. ‘The Mathematical Theory of Apportionment of Representatives’. *Proceedings of the National Academy of Science of the USA* 7(4): 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/aaunundh/upubliserte-artikler-og-notater/Prop-2d%5B1%5D.pdf>
- Leech, Dennis (2002). ‘Designing the Voting System for the Council of Ministers of the European Union’, *Public Choice* 437-464.
- Maier, S., Zachariassen P. and M. Zachariassen (2010). ‘Divisor-Based Biproportional Apportionment in Electoral Systems: A Real-Life Benchmark Study’. *Management Science* 56(2): 373–387.
- Pennisi, A. [2006]. ‘The Italian Bug: A Flawed Procedure for Bi-Proportional Seat Allocation’. In B. Simeone and F. Pukelsheim (eds), *Mathematics and democracy: recent advances in voting systems and collective choice* 151–165. Springer.
- Penrose, L.S. [1946]. ‘The elementary statistics of majority voting’. *Journal of the Royal Statistical Society* 109: 53–57.
- Pukelsheim, F. [2006]. ‘Current issues of apportionment methods’ In B. Simeone and F. Pukelsheim *op. cit.*
- Pukelsheim F., Ricca F., Scozzari A., Serafini P., Simeone B. [2011]. Network flow methods for electoral systems, *Networks*, in press, DOI 10.1002/net.20480.
- Serafini, P. and Simeone B. [2011-a], 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. [2011-b]. ‘Certificates of optimality: the third way to bi-proportional apportionment’. *Social Choice and Welfare* in press, DOI: 10.1007/s00355-010-0528-8.
- Widgren, M. [2005]. ‘Power in the Design of Constitutional Rules’. In G. Demange and M. Wooders (eds) *Group Formation in Economics: Networks, Clubs, and Coalitions* 312–334. Cambridge University Press.
- Zachariassen, P. and Zachariassen, M. (2006). ‘A comparison of electoral formulae for the Faroese Parliament’ In B. Simeone and F. Pukelsheim, *op. cit.*