Optimal Apportionment

Yukio KORIYAMA*, Jean-François LASLIER†
École Polytechnique‡

April 29, 2011

Abstract

This paper presents an argument in favor of the "degressive proportionality principle" in apportionment problems. The core of the argument is that each individual derives utility from the fact that the collective decision matches her own will with some frequency, with marginal utility being decreasing with respect to this frequency. Then classical utilitarianism at the social level recommends decision rules which exhibit degressive proportionality. Application is done to the case of the 27 states of the European Union.

[Preliminary and incomplete. Comments are welcome.]

1 Introduction

1.1 Background

Consider a situation in which repeated decisions have to be taken under the (possibly qualified) majority rule by representatives of groups (countries) that differ in size. In that case, the principle of equal representation (each representative should represent the same number of individuals) translates into a principle of proportional apportionment (the number of representatives of a country should be proportional to its population). Arguments have been raised against this principle and in favor of a principle of *degressive proportionality* according to which the ratio of the number of representatives to the population size should decrease with the population size rather than be constant.

The degressive proportionality principle is endorsed by most politicians and actually enforced (up to some qualifications) in the European institutions (Duff

*yukio.koriyama@polytechnique.edu
†jean-francois.laslier@polytechnique.edu
‡Département d'Économie, Palaiseau Cedex, 91128 France. For useful remarks, we thank Antonin Macé, Rafael Treibich, Michel Le Breton, Karine Van der Straeten, Eduardo Perez, Pierre Picard, Francisco Ruiz Aliseda, Stéphane Zuber, Annick Laruelle, Ani Guerdjikova, and the participants of the 2011 London workshop on Voting Power.

2010a, 2010b, TEU 2010). It is sometimes termed the Lamassoure-Séverin requirement following the European Parliament Resolution on "Proposal to amend the Treaty provisions concerning the composition of the European Parliament" which was adopted on October 11, 2007 after the report by Lamassoure and Severin (2007). On that occasion it was noted that the treaties and amendments of the European Union were referring to degressive proportionality "without defining this term in any more precise way." Then the October 2007 Resolution stated:

> [The European Parliament] Considers that the principle of degressive proportionality means that the ratio between the population and the number of seats of each Member State must vary in relation to their respective populations in such a way that each Member from a more populous Member State represents more citizens than each Member from a less populous Member State and conversely, but also that no less populous Member State has more seats than a more populous Member State.

It is known that, in the case of a Parliament, in which each member must have one and only one vote, the degressive proportionality requirement is impossible to satisfy exactly, due to unavoidable rounding problems (see for instance Cichocki and Życzkowski, 2010). But if one seeks to respect the principle "up to one", or "before rounding", and (rather obviously) if one allows for fractional weights, then many solutions become available, among which one has to choose (Ramirez Gonzalez, Palomares and Marquez 2006; Martínez Aroza and Ramírez González 2008, Grimmet *et al.* 2011).

The aim of this paper is to justify the principle of decreasing proportionality by an optimality argument and thereby to suggest the computation of optimal weights in specific instances, optimal weights which are degressively proportional.

1.2 Illustration of the argument

The argument in favor of degressively proportional apportionment is based on the maximization of an explicit utilitarian social criterion. Each individual derives utility from the fact that the collective decisions often match her own will. The social objective is simply the sum of such individual utilities. The argument can be explained with a very simple example.

Suppose there are only two countries, of size n_1 and n_2, with $n_1 < n_2$. Then, the majority rule gives full power to the big country. When the two countries agree on which decision to take, they are both satisfied, but when they disagree, country 1, the small one, is never satisfied. Intuition in that case recommends that the power to decide should be sometime given to the small country. To be more specific, suppose that binary decisions have to be taken according to the same decision rule. Among these decisions a fraction α is controversial in the sense that the two countries disagree. Suppose also, for the simplicity of the example, that the citizens within each country always agree on their best choice.

Under the majority rule, a citizen of country 2 is satisfied with probability 1, and a citizen of country 1 is satisfied with probability $1-\alpha$. To evaluate this rule at the collective level, one has to make an assumption as to how a citizen values the fact of seeing her will implemented with some frequency, say p. In this paper, we shall make the assumption that this evaluation is a concave function of p, say $\psi(p)$, the same function for every citizen. This means that the individual may well accept that in a moderate proportion of the cases the collective decision does not follow her will, but she incurs a relatively significant disutility if that proportion becomes too large. The individual would accept more easily to see her p decreasing from 1 to .95 than from .6 to .55. We found this hypothesis psychologically sound, and we will later explain its connection with the literature. Under this hypothesis, the sum of individual utilities under the majority rule is:

$$n_1\psi(1-\alpha)+n_2\psi(1),$$

because the will of the small country's citizens is fulfilled with probability $1-\alpha$, and the will of the big country's citizens is fulfilled with probability 1.

Imagine, for the sake of the illustration, that the decision is delegated at random to one or the other country with respective probabilities q_1 and $q_2 = 1-q_1$. Then the frequency of a decision opposed to country 1's will is αq_2, and the social value is:

$$U(q_1) = n_1\psi(1-\alpha q_2)+n_2\psi(1-\alpha q_1).$$

If ψ is linear then the maximum of utility is achieved for $q_1 = 0$, that is the majority rule, but if ψ is concave the maximum may be achieved at some interior point $0 < q_1 < 1$. More exactly, the condition for an interior optimum is that the marginal social benefit at point $q_1 = 0$,

$$U'(0) = \frac{1}{\alpha}\left(n_1\psi'(1-\alpha)-n_2\psi'(1)\right)$$

be positive, that is:

$$\frac{n_1}{n_2} > \frac{\psi'(1)}{\psi'(1-\alpha)}.$$

Such a condition is satisfied if the two countries are not too different in size, or if the marginal utility ψ' is rapidly decreasing with the probability p. In that case the optimal value of q is some number between 0 and 1 such that:

$$n_1\psi'(1-\alpha q_2) = n_2\psi'(1-\alpha q_1).$$

The optimal voting rule involves randomization, but one should not think of randomization as some dice to be thrown at the moment of the decision. In practice, there are two methods by which randomized-like rules are de facto achieved. One way is to use systems of alternate presidency. Decision is given to each member of the group for a fixed duration, and if questions to be solved arise in a random order through time, each member is decisive on a set of items which can be considered as random. The time slots allocated to the various participants can then be fine-tuned to achieve an optimal randomization. More importantly, with many countries, randomization naturally arises in practice in an even simpler way and without alternate presidency, provided that the coalitions of countries which support the same outcome vary with little or no systematic pattern. This route is followed in the sequel, where we build a stochastic model to render the above ideas and apply it to the 27 countries of the European Union.

1.3 Adjacent literature

Most of the existing literature on the subject deals with the measurement of voting power and the tricky combinatorics arising from the different ways to form a majority winning coalition with integer weighted votes; see the books Felsenthal and Machover (1998) and Laruelle and Valenciano (2008). Our focus is different, as can be seen from the two-country example above. The point made in the present paper rests on the non-linearity of ψ. It should be contrasted with the other contributions which derive an optimal rule from an explicit social criterion.

In Theil (1971) the objective is to minimize the average value of $1/w_{c(i)}$, where $w_{c(i)}$ is the weight of the country to which individual i belongs. This objective is justified as follows by Theil and Schrage (1977): "... let us assume that when such a citizen expresses a desire, the chance is w_i that he meets a willing ear. This implies that, in a long series of such expressed desires, the number of efforts per successful effort is $1/w_i$. Obviously, the larger this number, the worse the Parliament is from this individual's point of view. Our criterion is to minimize its expectation over the combined population." Minimizing this objective yields weights which are proportional to the square root of the country size.

In Felsenthal and Machover (1999), the objective is the mean majority deficit, that is the expected value of the difference between the size of the majority camp among all citizens and number of citizens who agree with the decision. In Le Breton, Montero and Zaporozhets (2010) the objective is to get as close as possible to a situation in which all citizens have the same voting power, as measured by the nucleolus of the voting game, a concept derived from cooperative game theory.

The first, and now classical, argument proposed in favor of degressive proportionality rests on statistical reasonings leading to the so-called "Penrose Law", which stipulates that the weight of a country should be proportional to the square root of the population rather than to the population itself (Penrose 1946), a pattern that exhibits degressive proportionality. The mathematical reason[1] why the square root appears in this literature is linked to the fact that

[1] The realized sum of n independent random variables is approximated by its mathematical expectation up to statistical fluctuations of the order $\sqrt{n}$.

the assumption is made that, within each country, voters' opinions are independent random variables. See Felsenthal and Machover (1998); Ramirez *et al.* (2006); Słomczyński and Życzkowski (2010). The political argument is thus that, in a world where frontiers have no link with the citizens' opinions, the representatives could as well be selected at random with no reference to these countries, but if representants have to be chosen country-wise, then the focus should be on the quality of the representation of the country by its constituents as a function of the size of the country. This argument is different from the one put forth in the present paper.

In Barbera and Jackson (2006), and Beisbart and Bovens (2007) the optimality is with respect to a sum of individual utilities, as in the present paper, but individual utilities are linear in p, so that these models do not capture the phenomenon that we wish to highlight.[2] The basic message of these papers is that country weights should be proportional to the importance of the issue for the country as a whole. In simplest settings this provides weights which are simply proportional to the population size. If we knew in advance the importance for the various countries of the various issues to be voted upon, then we should change accordingly the countries' weights. Of course this is not possible at the constitutional stage, but notice that part of this intuition is endogenized in the setting we propose. Start from weights proportional to the population. Larger countries are more often successful in that game. Therefore the outcome of the system is that a citizen (with concave utility) of a larger country is in a situation of lower marginal utility than a citizen of a smaller country. It may therefore be efficient to distort the weights in favor of the smaller countries if the small loss of the many citizens in the larger countries is more than compensated by the larger benefit for the citizens of the small countries.

1.4 Concavity of ψ

In additional to the psychological effects described above, there are several sources where concavity of ψ may emerge.

First, consider that the individual utility is defined over the sequence of successes. Suppose that the issues come in a sequence $t = 1, 2, \cdots$. Let $z^t \in \{0, 1\}$ be the success ($z^t = 1$) or failure ($z^t = 0$) at period t, and let $z = (z^t)_{t \in \mathbb{N}}$. Now utility u is defined on the equivalence class over the sequences z such that the following limit exists and equal to p:

$$\lim_{T \to \infty} \frac{1}{T} \left(\sum_{t=1}^{T} z^t \right) = p. \tag{1}$$

This is equivalent to say that the individual is indifferent in the order of success/failure in the sequence, and only the frequency matters. Define ψ by $\psi(p) = u(z)$ for any sequence which satisfies (1). Then, we have the following proposition.

[2] Such is also the case of Beisbart and Hartman (2010) who study the influence of inter-country utility dependencies for weights proportional to some power of the population sizes.

Proposition 1 *If u is submodular, then ψ is concave.*

Submodularity of u can be interpreted as the substitutability of different issues. An increase of frequency of successful issues is more favorable when there are less successes among the other issues.

In the same vein, and more concretely, suppose that an individual earns 1 unit of money if the collective decision matches her will, and 0 otherwise, and suppose that a fixed number T of such independent decisions will be taken, each time with the probability p of winning. Then the possible total payoffs are $S = 0, 1, ...,$ or T. A risk-adverse individual evaluates this prospect using a Von Neumann and Morgenstern utility function $v(s)$ concave in s. The expected utility is then a function of p:

$$\psi(p) = \sum_{t=0}^{T} \binom{T}{t} p^t (1-p)^{T-t} v(t).$$

Proposition 2 *If v is concave, then ψ is concave.*

The concavity of ψ can also be regarded as ambiguity aversion. The decision maker may have preferences over the probability measures on the acts. Klibanoff, Marinacci and Mukerji (2005) propose a model which exhibits an explicit separation between ambiguity as a characteristic of the decision maker's subjective beliefs and ambiguity attitude as a characteristic of the decision maker's tastes. Consider the following two scenarios. (i) An individual is certain that her frequency of success is 75%. (ii) There is ambiguity: her frequency of success may be either 50% or 100% with equal probabilities. Concavity of ψ implies that the individual prefers the first scenario. Even though both scenarios give an overall "probability" of success as 75%, the individual prefers the first scenario when she likes to know which lottery she faces. Since what we have in our mind as an application of our theory is the constitutional design, which should not be variant for different issues, it is reasonable to distinguish the risk over the different issues that an individual faces under a certain constitutional rule, and the uncertainty about the .

Lastly, the concavity of ψ can be interpreted as the expression of the aversion to inequality of the social planner (the constitutionalist). If the numbers u_i are money-metric measurements of i's welfare, the social planner may have, as its social objective, the maximization of a Kolm-Atkinson index of the form $W = \sum_i \psi(u_i)$.

Then, as is well-known from the theory of inequality measurement (Dutta 2002), the social objective is egalitarian if the function ψ is concave, for instance $\psi(p) = u_i^\alpha$ for $0 < \alpha < 1$. An extreme, degenerated case is the so-called Rawlsian objective of maximizing the well-being of the worst-off individual. (This case is obtained when α tends to 0, and we will show that it implies identical weights for all countries.) This "Social Welfare" point of view can be philosophically grounded on an intrinsic inequality aversion of the social planner reflected in the formula $W = \sum_i \psi(u_i)$ as well as by a purely utilitarian preference that takes into account decreasing marginal utility, as put forth by Bentham (1822), quoted by Trannoy (2011):

> All inequality is a source of evil – the inferior loses more in the account of happiness than the superior is gained.

2 The Model

2.1 Objectives

There are C countries $c = 1, ..., C$, and country c has a population of n_c individuals. We consider binary decision problems. In such a problem, there is a status quo, labeled 0, and an alternative decision labeled 1. Each individual i has a favorite decision $X_i \in \{0, 1\}$, and the final decision is denoted by $d \in \{0, 1\}$. A voting rule is used to take all such decisions so that, from the opinions of the voters, the final decision is in accordance with i's preference with some frequency:

$$p_i = \Pr[X_i = d].$$

This frequency gives more or less satisfaction to the individual; the utility of i is a function of p_i, say $\psi(p_i)$. We make the following assumptions:

1. ψ is the same for all individuals
2. ψ is increasing
3. ψ is concave

The first assumption can be conceived as methodological since we are dealing with a problem of constitution design. The second is almost without loss of generality (changing the preferred option). The third is psychologically meaningful, as argued in the introduction.

The social goal is defined from the individuals' satisfaction in an additive way:

$$U = \sum_i \psi(p_i).$$

This means that the collective judgment is based only on individual satisfaction with no complementarity at the social level. Notice that, because ψ is concave, the maximization of U tends to produce identical values for the individual probabilities p_i. Here the egalitarian goal is not postulated as a collective principle but follows from the individuals' assumed psychology.[3]

[3] One exception is allowed later in this paper. In Subsection 3.1, we consider the egalitarian case as a benchmark, where U is defined by the Rawlsian criterion.

2.2 Probabilistic model

In order to model the correlations between individual opinions, we assume that opinions are generated as follows: In each country c there is a general opinion $Y_c \in \{0, 1\}$, and each voter i in her country $c(i)$ forms an opinion conditionally on $Y_{c(i)}$. We suppose that the probability for a voter to have the general opinion of her country is the same for every voter in every country, and for both alternatives it is denoted μ.

$$\mu = \Pr\left[X_i = x | Y_{c(i)} = x\right], \ x = 0, 1.$$

We assume that μ is larger than $1/2$, so that Y_c can indeed be interpreted as the general opinion in country c.

The variables $Y_c \in \{0, 1\}$ are assumed to be randomly distributed and independent across countries. This assumption, which is in line with standard assumptions in the literature, captures the idea that the coalitions of countries which share a common view on a question show no systematic pattern. This assumption may be at odds with reality, but it can be defended in two ways. First, the way some countries' interests are aligned is itself variable: on some questions larger countries are opposed to smaller ones, other questions oppose rich countries to poor ones, East against West, North against South, etc. Second, in the spirit of constitutional design, one may wish by principle to be blind to current correlations of interest among some countries and give a strong interpretation to the idea that countries are independent entities.

Denote by γ the probability that any given country approves decision 1. Again γ is supposed to be the same for all countries, meaning that no country is a priori more conservative than the others.

$$\gamma = \Pr\left[Y_c = 1\right].$$

For the applications, the number of countries is moderate (say 27), and the number of voters in each country is large (at least several thousand). Therefore one can neglect intra-country randomness. Then, the proportion of voters who favor a reform in country c is μ with probability γ and $(1 - \mu)$ with probability $1 - \gamma$. The probability that a given voter favors a given reform is $\gamma\mu + (1 - \gamma)(1 - \mu)$.

2.3 Voting Rules

Each country c has a weight w_c. In the **Council** model, the country has in fact a unique representative, who votes according to the country's general opinion Y_c. Then the decision $d = 1$ is taken if, and only if, the total weight of the countries who voted for is larger than a threshold s:

$$d^{\text{council}} = 1 \text{ iff } \sum_c w_c Y_c > s.$$

In the **Parliament** model, the country has w_c representatives, who vote in proportion of the voters' opinions. Then, the number of votes at the parliament in favor of $d = 1$ is $w_c\mu$ for a country such that $Y_c = 1$, and is $w_c(1-\mu)$ for a country such that $Y_c = 0$.

Here, the decision $d = 1$ is taken if and only if the total weight of the representatives who voted for is larger than a threshold s:

$$d^{\text{parliament}} = 1 \text{ iff } \sum_c w_c \left(\mu Y_c + (1-\mu)(1-Y_c)\right) > s.$$

2.4 Questions

The same question can be asked for the Council model and for the Parliament model. The objective is to maximize the expected collective welfare. Given are: the population figures (n_c), the prior probability that a country favors the bill (γ), the intra-country in-homogeneity (μ), and the utility function (ψ). One has to choose the weights w_c and the threshold s; that make $C+1$ variables, but given the form of the two decision rules, we can suppose that $\sum_c w_c = 1$. The expected social welfare is:

$$U = \sum_i \psi(p_i) = \sum_i \psi(\Pr[X_i = d]) = \sum_c n_c \psi\left(\pi_c\right), \tag{2}$$

with

$$\pi_{c(i)} = \Pr[X_i = d]$$

for any citizen i of country c. This probability can be decomposed conditionally on the country's general opinion Y_c:

$$\begin{aligned} \pi_c = {} & \gamma\mu \Pr\left[d = 1 | Y_c = 1\right] + (1-\gamma)(1-\mu)\Pr\left[d = 1 | Y_c = 0\right] \\ & + \gamma(1-\mu)\Pr\left[d = 0 | Y_c = 1\right] + (1-\gamma)\mu \Pr\left[d = 0 | Y_c = 0\right]. \end{aligned} \tag{3}$$

Especially, when the prior is symmetric (i.e. $\gamma = 1/2$),

$$\pi_c = 1 - \mu + \left(\mu - \frac{1}{2}\right)\left\{\Pr\left[d = 1 | Y_c = 1\right] + \Pr\left[d = 0 | Y_c = 0\right]\right\}. \tag{4}$$

One therefore needs to compute the probabilities $\Pr\left[d | Y_c\right]$. We use the following Lemma.

Lemma 1 *Given the weighted voting rule (w, s), we have*

$$\Pr\left[d = 1 | Y_c = 1\right] = \Pr\left[\sum_{k \neq c} w_k Y_k \geq s' - w_c\right]$$

$$\Pr\left[d = 0 | Y_c = 0\right] = \Pr\left[\sum_{k \neq c} w_k Y_k < s'\right]$$

where $s' = s$ for the Council model and $s' = \frac{s-(1-\mu)}{2\mu-1}$ for the Parliament model.

Our first result is about the threshold. When the prior is symmetric, the optimal voting rule is the weighted majority rule.

Proposition 3 *For $\gamma = 1/2$, the optimal threshold is $s = 1/2$ for both the Council model and the Parliament model.*

In the next Section, we report both theoretical and numerical results concerning the optimal weights.

3 Optimal weights

3.1 Two benchmarks

The linear case

Suppose that the function ψ is linear; then without loss of generality we can take $\psi(p) = p$. Then the optimal weights are simply proportional to the population.

Proposition 4 *If $U = \sum_i p_i$ the optimal decision rule is a weighted majority, with weights w_c proportional to the population.*

The Rawlsian case

On the other hand, suppose that the social criterion gives absolute priority to the worse-off individual, what is sometimes called the MaxMin, or Rawls's criterion. Then the optimal weights are independent of country populations.

Proposition 5 *If $U = \min_i p_i$ the optimal decision rule is the simple majority among countries: all countries have equal weight.*

3.2 Normal approximation

The probabilities $\Pr[d|Y_c]$ are derived from the weighted sum of $C-1$ identical and independent Bernoulli variables. Explicit description of these probabilities may require complex computations. However, when C is large enough (e.g. $C > 15$), approximation by normal distribution is sufficiently accurate. Let m_{-c} and σ_{-c} denote the mean and the standard deviation:

$$m_{-c} = \mathbb{E}\left[\sum_{k \neq c} w_k Y_k\right] = \gamma \sum_{k \neq c} w_k = \gamma(1 - w_c),$$

$$\sigma_{-c}^2 = \mathbb{V}\left[\sum_{k \neq c} w_k Y_k\right] = \gamma(1-\gamma) \sum_{k \neq c} w_k^2.$$

Our approximation is:

$$\sum_{k \neq c} w_k Y_k \rightsquigarrow \mathcal{N}\left(m_{-c}, \sigma_{-c}\right).$$

Then, by Lemma 1,

$$\Pr[d=1|Y_c=1] = 1-\Phi\left(\frac{s'-w_c-m_{-c}}{\sigma_{-c}}\right),$$
$$\Pr[d=0|Y_c=0] = \Phi\left(\frac{s'-m_{-c}}{\sigma_{-c}}\right),$$

where Φ is the cumulative distribution function of the standard normal distribution. One can also check (with no surprise) that the same result as Proposition 3 is true for the normal approximation.

Proposition 6 *For $\gamma = 1/2$, in the normal approximation the optimum threshold is $s = 1/2$.*

When $\gamma = 1/2$ and $s = 1/2$,

$$\Pr[d=1|Y_c=1] = \Pr[d=0|Y_c=0] = \Phi\left(\frac{w_c/2}{\sigma_{-c}}\right).$$

Let $\widetilde{U}$ denote the approximated collective welfare. By (2), (3) and Lemma 1, we have:

$$\widetilde{U} = \sum_c n_c g\left(\frac{w_c/2}{\sigma_{-c}}\right)$$

where $f(x) = 1-\mu+(2\mu-1)x$ and $g = \psi \circ f \circ \Phi$. Note that f is a linear function.

Theorem 1 *Suppose that the prior is symmetric ($\gamma = 1/2$). If ψ is sufficiently concave, then the optimal weights should exhibit degressive proportionality. More precisely, if $-\frac{g''(x)}{g'(x)} > \frac{3x}{1+x^2}$ for $x > 0$, then $n_c < n_{c'}$ implies $\frac{w_c}{n_c} > \frac{w_{c'}}{n_{c'}}$.*

Without assuming any sufficient condition on the degree of concavity, degressive proportionality is also obtained when no country is large.

Theorem 2 *Suppose that the prior is symmetric ($\gamma = 1/2$). If all countries are sufficiently small, then the optimal weights should exhibit degressive proportionality.*

To see how strong the sufficient condition in Theorem 1 is, consider a family of functions, $\psi_a(p) = \log(p-a)$ for $a \in (0, 1/2)$. Then, $-\psi_a''(p)/\psi_a'(p) = (p-a)^{-1}$, which is increasing in a. It is straightforward to see numerically that the sufficient condition is satisfied if $a > 0.367$. Therefore, optimal weights are proportionally degressive for such functions ψ_a. Now, suppose $a = 0.3$. Then, the sufficient condition is not satisfied. However, condition (6) in the Proof of Theorem 2 is satisfied for $x < 0.66$. For a given value of w_c, maximum x is attained if the weights are the same for all $k \neq c$: $w_k = (1-w_c)/(n-1)$. Then, $x^{\max} = \sqrt{n-1}w_c/(1-w_c)$. For example, for $n = 27$, $x < 0.66$ is guaranteed if $w_c < (0.66)/\left(0.66+\sqrt{26}\right) \simeq 0.115$. Therefore, if no country has a weight bigger than 0.115, it is guaranteed that the optimal weights are proportionally degressive.

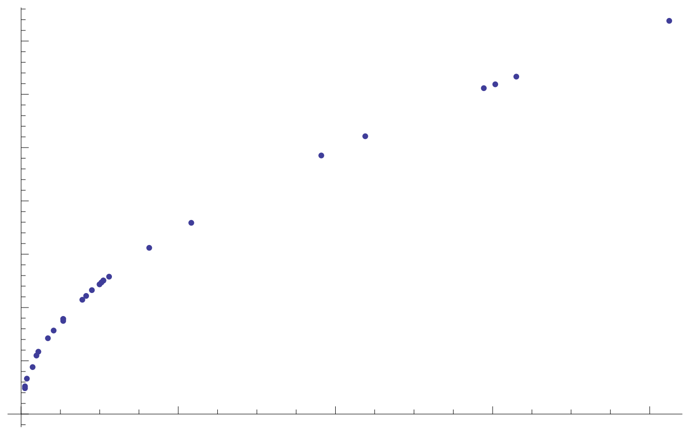

Figure 1: Optimal weights function of the populations.

3.3 Numerical results

Let us consider the 27 European countries, with 751 seats to be allocated. We take the following parameters:

$$\gamma = 1/2$$
$$\mu = 1$$
$$\psi(p) = \log(p - 1/2).$$

Table 1 provides for each country the population, the number of representatives proposed by Pukelsheim (2010), and the optimal number of representative for our model. The apportionment method used by Pukelsheim (2010), often called the Fix+Prop method, uses a base of 6 seats per citizenry, with the remaining 589 seats for proportional apportionment using standard rounding methods. As to the optimal weights for our model, the optimal threshold is $s = .5$, as proved in Proposition 3. Table 1 indicates the non-rounded optimal number of seats, that is $751 \times w_c$, to be compared with Fix-Prop. The figures have been obtained using the FindMinimum program in *Mathematica*. They increase from 4.86 for Malta to 72.83 for Germany in a concave way, as it can be seen on Figure 1. The last column of Table 1 indicates the value of the probability $p_i = \Pr[X_i = d]$ that the collective decision matches individual i's will. Naturally, the optimal utilitarian weights are such that this probability depends on the country and is larger in larger countries.

3.4 Discussion

In the symmetric model ($\gamma = 1/2$), the values of the probabilities π_c are always larger than 1/2 at the optimum. The above computation was done for the

	population	Fix+Prop	optimal	probability
Germany	82 438 000	96	73.8340	0.678508
France	62 886 200	83	63.4343	0.650804
United Kingdom	60 421 900	80	62.0563	0.647231
Italy	58 751 700	77	61.1115	0.644794
Spain	43 758 300	59	52.1370	0.622098
Poland	38 157 100	52	48.4857	0.613075
Romania	21 610 200	32	36.0653	0.583127
Netherlands	16 334 200	26	31.2435	0.571756
Greece	11 125 200	20	25.6958	0.55881
Portugal	10 569 600	19	25.0368	0.557281
Belgium	10 511 400	19	24.9668	0.557119
Czech Republic	10 251 100	18	24.6516	0.556388
Hungary	10 076 600	18	24.4380	0.555893
Sweden	9 047 800	17	23.1413	0.552892
Austria	8 265 900	16	22.1076	0.550503
Bulgaria	7 718 800	15	21.3558	0.548768
Denmark	5 427 500	13	17.8812	0.540773
Slovak Republic	5 389 200	13	17.8176	0.540627
Finland	5 255 600	12	17.5938	0.540113
Ireland	4 209 000	11	15.7343	0.535849
Lithuania	3 403 300	10	14.1411	0.532203
Latvia	2 294 600	9	11.6032	0.526405
Slovenia	2 003 400	8	10.8400	0.524663
Estonia	1 344 700	8	8.87718	0.520189
Cyprus	766 400	7	6.69932	0.51523
Luxembourg	459 500	7	5.18634	0.511788
Malta	404 300	6	4.86469	0.511057

Table 1: Population, Fix+Prop rounded weights, optimal weights and individual probabilities

objective function $\log(p - 1/2)$. Using the family of objectives

$$\psi_a(p) = \log(p - a)$$

for different values of a with $0 < a < 1/2$, one finds that the optimal weights exhibit less and less non-linearity when a is smaller. This is due to the fact that the concavity of the function ψ_a is increasing with a. This confirms the intuition on which this paper is based: the optimal weights exhibit degressive proportionality because the objective function is concave.

A Appendix

A.1 Proofs

Proof of Proposition 1. We first show that for any $p_1, p_2 \in \mathbb{Q}$ with $p_1 < p_2$, we have $u(p_1) + u(p_2) \leq 2u((p_1+p_2)/2)$. To see that, let z be a sequence which repeats $\left(\underbrace{1, \cdots, 1}_{(p_1+p_2)/2}, 0, \cdots, 0\right)$, and let z' be a sequence which repeats

$$\left(\underbrace{0, \cdots, 0}_{(p_2-p_1)/2}, \underbrace{1, \cdots, 1}_{(p_1+p_2)/2}, 0, \cdots, 0\right).$$

Then, $u(z) = u(z') = (p_1+p_2)/2$. Now, obviously $z \vee z'$ is a sequence which repeats

$$\left(\underbrace{1, \cdots, 1}_{p_2}, 0, \cdots, 0\right),$$

and $z \wedge z'$ is a a sequence which repeats

$$\left(\underbrace{0, \cdots, 0}_{(p_2-p_1)/2}, \underbrace{1, \cdots, 1}_{p_1}, 0, \cdots, 0\right).$$

By assumption, u is submodular: $u(z) + u(z') \geq u(z \vee z') + u(z \wedge z')$. Hence, $\psi(p_1) + \psi(p_2) \leq 2\psi((p_1+p_2)/2)$. By continuity and monotonicity, ψ is concave. ■

Proof of Proposition 2. It is straightforward to show that:

$$\psi'(p) = T\sum_{t=0}^{T-1}\binom{T-1}{t}p^t(1-p)^{T-1-t}\{v(t+1)-v(t)\},$$

$$\psi''(p) = T(T-1)\sum_{t=0}^{T-2}\binom{T-2}{t}p^t(1-p)^{T-2-t}\begin{bmatrix}\{v(t+2)-v(t+1)\} \\ -\{v(t+1)-v(t)\}\end{bmatrix}.$$

Since v is increasing and concave,

$$v(t+1)-v(t) > 0,$$
$$\{v(t+2)-v(t+1)\} - \{v(t+1)-v(t)\} < 0.$$

Hence, $\psi'(p) > 0$ and $\psi''(p) < 0$ for $p \in (0,1)$. ■

Proof of Lemma 1. We first give a proof for the Parliament model. By definition,

$$\begin{aligned}\Pr[d=1|Y_c=1] &= \Pr\left[\sum_k w_k\left(\mu Y_k+(1-\mu)(1-Y_k)\right)>s \,\middle|\, Y_c=1\right]\\ &= \Pr\left[\sum_{k\neq c} w_k\left(\mu Y_k+(1-\mu)(1-Y_k)\right)>s-w_c\mu\right]\\ &= \Pr\left[\sum_{k\neq c} w_k Y_k>\frac{s-(1-\mu)}{2\mu-1}-w_c\right].\end{aligned}$$

Similarly,

$$\begin{aligned}\Pr[d=0|Y_c=0] &= \Pr\left[\sum_k w_k\left(\mu Y_k+(1-\mu)(1-Y_k)\right)<s \,\middle|\, Y_c=0\right]\\ &= \Pr\left[\sum_{k\neq c} w_k Y_k<\frac{s-(1-\mu)}{2\mu-1}\right].\end{aligned}$$

By setting $\mu = 1$, we obtain the result for the Council model. ■

Proof of Proposition 3. Recall that $\pi_c = \Pr[X_i = d]$ for any citizen i in country c. Denote

$$\widetilde{Y}_{-c} = \sum_{k\neq c} w_k Y_k.$$

Then, by Lemma 1,

$$\begin{aligned}\Pr[d=0|Y_c=0] &= \Pr\left[\widetilde{Y}_{-c}\leq s'\right],\\ \Pr[d=1|Y_c=1] &= \Pr\left[\widetilde{Y}_{-c}>s'-w_c\right]=1-\Pr\left[\widetilde{Y}_{-c}\leq s'-w_c\right].\end{aligned}$$

Hence, by (4)

$$\pi_c = 1-\mu+\left(\mu-\frac{1}{2}\right)\Pr\left[s'-w_c<\widetilde{Y}_{-c}\leq s'\right].$$

Notice that this implies that π_c is increasing with w_c in the sense that if one compares two countries c, c' with $w_c < w_{c'}$ then $\pi_c \leq \pi_{c'}$. The random variable $\widetilde{Y}_{-c}$ is a weighted sum of Bernoulli variables, each of whom take value 0 and 1 with probability $\gamma = 1/2$. The density of $\widetilde{Y}_{-c}$ is a step function which is symmetric around the average $\overline{w}_{-c} := (1/2)\sum_{k\neq c} w_k$, non-decreasing before $\overline{w}_{-c}$, and non-increasing after $\overline{w}_{-c}$. Therefore the integral of $\widetilde{Y}_{-c}$ on an interval of fixed length attains its maximum when the interval is centered on $\overline{w}_{-c}$. In that case, the mid-point of the interval $\left[s' - w_c < \widetilde{Y}_{-c} \leq s'\right]$, $s' - w_c/2$, is equal to $\overline{w}_{-c} = (1 - w_c)/2$. It follows that, for any c, π_c is maximum for $s' = 1/2$. For both the Council model and the Parliament model, it implies $s = 1/2$. Since ψ is an increasing function, the maximum of $U = \sum_c n_c \psi(\pi_c)$ is obtained at $s = 1/2$. ■

One should remark that this result holds for all values of the weights w_c, even non-optimal ones.

Proof of Proposition 4. The objective is $U = \sum_i \Pr[X_i = d]$. Conditionally on a realization of the vector of variables $(Y_c)_{c \in C} \in \{0,1\}^C$, the social utility of taking decision $d = 0$ or 1 is

$$\begin{aligned} U(d=0) &= \sum_{c:Y_c=0} \mu n_c + \sum_{c:Y_c=1} (1-\mu) n_c \\ U(d=1) &= \sum_{c:Y_c=1} \mu n_c + \sum_{c:Y_c=0} (1-\mu) n_c, \end{aligned}$$

so that $d = 1$ is strictly better if and only if $(2\mu - 1)\sum_{c:Y_c=1} n_c > (2\mu - 1)\sum_{c:Y_c=0} n_c$. Since $\mu > 1/2$, we know which decision d maximizes the criterion, that is majority rule: $d = 1$ if $\sum_{c:Y_c=1} n_c > \sum_{c:Y_c=0} n_c$ and $d = 0$ otherwise. This optimal rule is indeed a weighted majority rule with weight $w_c = n_c / \sum_{c'} n_{c'}$ and threshold 1/2. ■

Proof of Proposition 5. The objective is $U = \min_c \pi_c$. By Proposition 4, the optimal decision rule is the simple majority rule with the equal weight, if $n_c = 1$ for $\forall c$. That is, for any rule, $\sum_c \pi_c \leq C p^{eq}$, where p^{eq} is the probability of winning under the equal weight. Now, suppose that $p^{eq} < \min_c \pi_c$. Then, $p^{eq} < \pi_c$ for $\forall c$, implying $C p^{eq} < \sum_c \pi_c$, a contradiction. Therefore, $\min_c \pi_c \leq p^{eq}$ for any rule. Hence, optimal U is p^{eq}, which is attained by the equal weight. ■

Proof of Proposition 6. By (4) and Lemma 1, for $\gamma = 1/2$,

$$\pi_c = 1 - \mu + \left(\mu - \frac{1}{2}\right)\left\{1 - \Phi\left(\frac{s' - 1/2 - w_c/2}{\sigma_{-c}}\right) + \Phi\left(\frac{s' - 1/2 + w_c/2}{\sigma_{-c}}\right)\right\}.$$

Observe that

$$\begin{aligned} \frac{\partial \pi_c}{\partial s'} = 0 &\Leftrightarrow \Phi'\left(\frac{s' - 1/2 - w_c/2}{\sigma_{-c}}\right) = \Phi'\left(\frac{s' - 1/2 + w_c/2}{\sigma_{-c}}\right) \\ &\Leftrightarrow \exp\left(-\frac{1}{2\sigma_{-c}^2}\left(s' - \frac{1}{2} - \frac{w_c}{2}\right)^2\right) = \exp\left(-\frac{1}{2\sigma_{-c}^2}\left(s' - \frac{1}{2} + \frac{w_c}{2}\right)^2\right) \end{aligned}$$

which implies $s' = 1/2$. On the other hand, it is straightforward to see that $\left.\frac{\partial^2 \pi_c}{\partial s'^2}\right|_{s'=1/2} = (2\mu - 1)\,\Phi''(w_c/2) < 0$. Therefore, π_c is maximized at $s' = 1/2$, (i.e. $s = 1/2$ both in the Council and in the Parliament model) for each country c. Since $\widetilde{U} = \sum_c n_c \psi(\pi_c)$ and ψ is increasing, $\widetilde{U}$ is maximized at $s = 1/2$. ■

Proof of Theorem 1. Suppose, to the contrary, that w is an optimal weight vector and there exists a pair (c_1, c_2) such that $n_{c_1} < n_{c_2}$ and $\frac{w_{c_1}}{n_{c_1}} \leq \frac{w_{c_2}}{n_{c_2}}$. We show that there exists a pair $\left(w'_{c_1}, w'_{c_2}\right)$ such that $w_{c_1}^2 + w_{c_2}^2 = \left(w'_{c_1}\right)^2 + \left(w'_{c_2}\right)^2$ and $n_{c_1} g_{c_1}(w) + n_{c_2} g_{c_2}(w) < n_{c_1} g_{c_1}(w') + n_{c_2} g_{c_2}(w')$, where w' is the weight vector of which w_{c_1} and w_{c_2} are replaced by w'_{c_1} and w'_{c_2}. Then w' is an improvement of w, contradicting the optimality of w.[4] Let

$$\eta(x) = n_{c_1} g\left(\sqrt{\frac{w_{c_1}^2 + x}{W - \left(w_{c_1}^2 + x\right)}}\right) + n_{c_2} g\left(\sqrt{\frac{w_{c_2}^2 - x}{W - \left(w_{c_2}^2 - x\right)}}\right).$$

We want to show $\eta'(0) > 0$. Let $h(x) = g\left(\sqrt{\frac{w_{c_1}^2 + x}{W - \left(w_{c_1}^2 + x\right)}}\right)$. Then, $\eta(x) = n_{c_1} h(x) + n_{c_2} h\left(w_{c_2}^2 - w_{c_1}^2 - x\right)$. Hence,

$$\eta'(x) = n_{c_1} h'(x) - n_{c_2} h'\left(w_{c_2}^2 - w_{c_1}^2 - x\right),$$
$$\eta'(0) = n_{c_1} h'(0) - n_{c_2} h'\left(w_{c_2}^2 - w_{c_1}^2\right).$$

We want to show $n_{c_1} h'(0) > n_{c_2} h'\left(w_{c_2}^2 - w_{c_1}^2\right)$. By definition,

$$h'(x) = g'\left(\sqrt{\frac{w_{c_1}^2 + x}{W - \left(w_{c_1}^2 + x\right)}}\right) \frac{1}{2\sqrt{\frac{w_{c_1}^2 + x}{W - \left(w_{c_1}^2 + x\right)}}} \frac{W}{\left(W - \left(w_{c_1}^2 + x\right)\right)^2}.$$

$$h'(0) = g'\left(\sqrt{\frac{w_{c_1}^2}{W - w_{c_1}^2}}\right) \frac{1}{2\sqrt{\frac{w_{c_1}^2}{W - w_{c_1}^2}}} \frac{W}{\left(W - w_{c_1}^2\right)^2},$$

$$h'\left(w_{c_2}^2 - w_{c_1}^2\right) = g'\left(\sqrt{\frac{w_{c_2}^2}{W - w_{c_2}^2}}\right) \frac{1}{2\sqrt{\frac{w_{c_2}^2}{W - w_{c_2}^2}}} \frac{W}{\left(W - w_{c_2}^2\right)^2}.$$

We want to show:

$$\frac{n_{c_1}}{w_{c_1}} g'\left(\sqrt{\frac{w_{c_1}^2}{W - w_{c_1}^2}}\right) \frac{1}{\left(W - w_{c_1}^2\right)^{\frac{3}{2}}} > \frac{n_{c_2}}{w_{c_2}} g'\left(\sqrt{\frac{w_{c_2}^2}{W - w_{c_2}^2}}\right) \frac{1}{\left(W - w_{c_2}^2\right)^{\frac{3}{2}}}.$$

Since we assumed $\frac{w_{c_1}}{n_{c_1}} \leq \frac{w_{c_2}}{n_{c_2}}$, it is sufficient to show that

$$g'\left(\sqrt{\frac{w_{c_1}^2}{W - w_{c_1}^2}}\right) \frac{1}{\left(W - w_{c_1}^2\right)^{\frac{3}{2}}}$$

[4] Note that w' does not sum up to one in general. But once such a vector $w'$is found, we can obtain exactly the same probability of winning by normalizing w'. Hence, it suffices to find an unnormalized vector w'.

is strictly decreasing in w_{c_1}. Let $x = \sqrt{w_{c_1}^2 / \left(W - w_{c_1}^2\right)}$. Then, $\frac{1}{W - w_{c_1}^2} = \frac{1+x^2}{W}$. Hence, what we want to show is equivalent to that $g'(x)\left(1+x^2\right)^{\frac{3}{2}}$ is decreasing in $x\,(> 0)$. This is equivalent to:

$$-\frac{g''(x)}{g'(x)} > \frac{3x}{1+x^2} \text{ for } x > 0. \tag{5}$$

■

Proof of Theorem 2. Let $\Psi = f \circ \Phi$. Then,

$$g'(x) = \psi'(\Psi(x))\,\Psi'(x),$$
$$g''(x) = \psi''(\Psi(x))\,(\Psi'(x))^2 + \psi'(\Psi(x))\,\Psi''(x).$$

Hence,

$$-\frac{g''(x)}{g'(x)} = -\frac{\psi''(\Psi(x))}{\psi'(\Psi(x))}\Psi'(x) - \frac{\Psi''(x)}{\Psi'(x)}.$$

Since Φ is the cdf of the standard normal distribution, $-\frac{\Phi''(x)}{\Phi'(x)} = x$. Since f is linear, $-\frac{\Psi''(x)}{\Psi'(x)} = x$. Therefore, condition (5) is equivalent to:

$$-\frac{\psi''(\Psi(x))}{\psi'(\Psi(x))} > \frac{1}{\Psi'(x)}\left(\frac{3x}{1+x^2} - x\right) \text{ for } x > 0. \tag{6}$$

The right hand side is zero for $x = 0$ (note that $\Psi'(0) = (2\mu - 1)\,\Phi'(0) > 0$). By assumption, $-\frac{\psi''(\Psi(0))}{\psi'(\Psi(0))} > 0$. Therefore, $\exists x^0$ such that for all $x \in \left(0, x^0\right)$, condition (6) is satisfied. Remember $x = \sqrt{w_c^2 / \left(W - w_c^2\right)}$. Hence, $\exists w^0$ such that for all $w_c \in \left(0, w^0\right)$, x is small enough so that (5) is satisfied. ■

References

[1] Salvador Barberà and Matthew O. Jackson (2006) "On the weights of nations: assigning voting weights in a heterogeneous union" *The Journal of Political Economy* 114: 317—339.

[2] Claus Beisbart and Luc Bovens (2007) "Welfarist evaluations of decision rules for boards of representatives" *Social Choice and Welfare* 29: 581—608.

[3] Claus Beisbart and Stephan Hartmann (2010) "Welfarist evaluations of decision rules under interstate utility dependencies" *Social Choice and Welfare* 34: 315—344.

[4] Jeremy Bentham (1822) *First Principle Preparatory to the Constitutional Code*. reprinted in: *The Works of Jeremy Bentham, Volume IX*. Elibron Classics, Adamant Media Corporation (2005).

[5] Cichocki, M. and Życzkowski, K., eds.(2010) *Institutional Design and Voting Power in the European Union*. Ashgate, London.

[6] Andrew Duff (2010a) "Explanatory Statement. Proposal for a modification of the Act concerning the election of the members of the European Parliament by direct universal suffrage of 20 September 1976" *European Parliament, Committee on Constitutional Affairs*, 2010.

[7] Andrew Duff (2010b) "Draft Report on a proposal for a modification of the Act concerning the election of the members of the European Parliament by direct universal suffrage of 20 September 1976 (2009/2134(INI))" *European Parliament, Committee on Constitutional Affairs*, 2010.

[8] Bhaskar Dutta (2002) "Inequality, poverty and welfare" pp. 597—633 in: *Handbook of Social Choice and Welfare, Volume 1*, Edited by K.J. Arrow, A.K. Sen and K. Suzumura. Amsterdam: Elsevier.

[9] Dan Felsenthal and Moshé Machover (1998) *The Measurement of Voting Power*, Cheltenham: Edward Elgar.

[10] Dan Felsenthal and Moshé Machover (1999) "Minimizing the mean majority deficit: the second square root rule" *Mathematical Social Sciences* 37: 25—37.

[11] Geoffrey Grimmet, Friedrich Pukelsheim, Jean-François Laslier, Victoriano Ramírez González, Wojciech Słomczyński, Martin Zachariasen, and Karol Życzkowski (2011) "The allocation between the EU Member States of the seats in the European Parliament: The Cambridge Compromise" European Parliament Policy department, Constitutional affairs. Brussels: European Parliament

[12] Peter Klibanoff, Massimo Marinacci and Sujoy Mukerji (2005) "A Smooth model of Decision Making under Ambiguity" *Econometrica* 73(6): 1849—1892.

[13] Alain Lamassoure and Adrian Severin (2007) European Parliament Resolution on "Proposal to amend the Treaty provisions concerning the composition of the European Parliament" adopted on 11 October 2007 (INI/2007/2169).

[14] Michel Le Breton, Maria Montero, and Vera Zaporozhets (2010) "Voting power in the EU Council of Ministers and fair decision making." *mimeo.*

[15] J. Martínez and Victoriano Ramírez (2008) "Several methods for degressively proportional allocations. A case study" *Mathematical and Computer Modelling*, 48: 1439—1445.

[16] Guillermo Owen (1995) *Game Theory*, (3rd edition) Academic Press Inc.

[17] Lionel S. Penrose (1946) "The elementary statistics of majority voting" *Journal of the Royal Statistical Society* 109: 53—57.

[18] Friedrich Pukelsheim (2010) "Putting citizens first: Representation and power in the European Union." Pages 235—253 in: *Institutional Design and Voting Power in the European Union* (Edited by M. Cichocki and K. Życzkowski), Ashgate: London.

[19] Victoriano Ramírez, Antonio Palomares, Maria Luisa Márquez (2006) "Degressively proportional methods for the allotment of the European Parliament seats amongst the EU member states." Pages 205-220 in: *Mathematics and Democracy – Recent Advances in Voting Systems and Collective Choice*. (Edited by Bruno Simeone and Friedrich Pukelsheim), Springer: Berlin.

[20] Wojciech Słomczyński and Karol Życzkowski (2010) "Jagiellonian Compromise – An alternative voting system for the Council of the European Union." Pages 43—57 in: Institutional Design and Voting Power in the European Union (Edited by M. Cichocki and K. Życzkowski), Ashgate: London.

[21] Henri Theil (1971) "The allocation of power that minimizes tension" *Operations Research* 19: 977—982.

[22] Henri Theil and Linus Schrage (1977) "The apportionment problem and the European parliament" *European Economic Review* 9: 247—263.

[23] Alain Trannoy (2010) "Mesure des inégalités et dominance sociale" *Mathématiques et Sciences Humaines* 192: 29—40.

[24] Treaty on European Union: «Articles 1-19.» *Official Journal of the European Union* C 83 (30.2.2010) 13—27.