# Optimal Apportionment

Yukio KORIYAMA*, Jean-François LASLIER,
Antonin MACÉ, and Rafael TREIBICH
Ecole Polytechnique†

December 17, 2012

**Abstract**

This paper provides a theoretical foundation which supports the *degressive proportionality principle* in apportionment problems, such as the allocation of seats in a Parliament. The core of the argument is that the utility assigned to a constitutional rule is a non-linear function of the frequency with which each collective decision matches the individual's own will. If the function is concave, then classical utilitarianism at the social level recommends decision rules which exhibit degressive proportionality with respect to the population size.

# 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 (e.g. countries) that differ in size. In this case, the principle of equal representation translates into a principle of proportional apportionment. In other words, if we require each representative to represent the same number of individuals,

*Corresponding author: yukio.koriyama@polytechnique.edu

†Département d'Économie, Palaiseau Cedex, 91128 France. For useful remarks, we thank Ani Guerdjikova, Annick Laruelle, Michel Le Breton, Eduardo Perez, Pierre Picard, Francisco Ruiz-Aliseda, Karine Van der Straeten, Jörgen Weibull, Stéphane Zuber, and the participants of the 2011 workshop on Voting Power and Procedures at LSE, Advances in the Theory of Individual and Collective Decision-Making at Istanbul Bilgi University, SAET 2011, D-TEA Workshop 2011, seminar participants at Ecole Polytechnique, Paris School of Economics, University of Edinburgh, BETA (Université de Strasbourg) and CREM (Université de Caen Basse Normandie & Rennes 1).

the number of representatives of a group 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 2010a, 2010b, TEU 2010). It is sometimes termed the Lamassoure-Severin 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 has been referring to degressive proportionality "without defining this term in any more precise way." 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", then many solutions become available, among which one has to choose (Ramírez-González, Palomares and Marquez 2006; Martínez-Aroza and Ramírez-González 2008; Grimmet *et al.* 2011). Such is also the case (rather obviously) if one allows for fractional weights.

The same principles formally apply to the case where a country is represented by a number of representatives, each of whom is given one vote, and to the case where a country is represented by a single delegate who is given a weight in relation to the country size. We shall refer to the two cases as a *Parliament* and a *Council*.

This paper applies Normative Economics to Politics. Its aim is to justify the principle of degressive proportionality by an optimality argument.

## 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. To evaluate a constitutional rule at the collective level, one has to describe how the society evaluates the fact that the will of each citizen is reflected in the social decision under the rule. Let $\psi(p)$ be the expected utility derived from the decision rule as a function of the frequency with which her preferred alternative matches the social decision determined by the rule. We assume that $\psi$ is increasing and concave.[1] The social objective is simply the sum of such individual utilities. The argument can be explained with a very simple example.

Suppose there are $C$ countries. There is one large country with population $n_1$, while the others are equally small, with population $n_2$. Let us assume that country 1 is so large that it contains more than the half of the entire population in the society: $n_1 > (C-1)\, n_2$. Proportional apportionment $w = (n_1, n_2, \cdots, n_2)$ entitles the full decisional power to the large country: 100 % of the decisions will be made by the large country. Intuition, in that case, may recommend that the decisional power should be occasionally given to small countries. A decrease in decision frequency from 100 % to slightly less for the citizens of the large country may be more than compensated by some small increase for the citizens of the small countries. To be more precise, suppose that a series of binary decisions (to pass the bill or not) is taken and that each country's preferences over the binary decision are symmetrically and independently distributed. Citizens in the big country are always satisfied, while those in small countries are satisfied with probability 1/2, as their preferred choice may happen to agree with that of the big country by chance. The utilitarian social welfare is thus:

$$U = n_1 \psi(1) + (C-1)\, n_2 \psi\left(\frac{1}{2}\right).$$

Now, suppose instead that a slightly smaller weight is apportioned to the big country: $w' = (w_1, w_2, \cdots, w_2)$ where $w_1/w_2 \in (C-3, C-1)$. Then, the big country loses when all small countries disagree. Such an event occurs with probability $1/2^{C-1}$, implying that the frequency of success decreases

[1] **We do not need any behavioral assumption to obtain concavity of the utility as a function of the frequency. Especially, there is no violation of the von Neumann-Morgenstern axioms. Detailed discussion, including the microfoundation, is given in Section 4.**

(resp. increases) by $1/2^{C-1}$ for the big country (resp. for the small countries). The social welfare is now:

$$U' = n_1\psi\left(1 - \frac{1}{2^{C-1}}\right) + (C-1)\, n_2\psi\left(\frac{1}{2} + \frac{1}{2^{C-1}}\right).$$

If $C$ is sufficiently large, the first order approximation yields:

$$U - U' \simeq \frac{1}{2^{C-1}}\left(n_1\psi'(1) - (C-1)\, n_2\psi'\left(\frac{1}{2}\right)\right).$$

Hence, $U < U'$ if

$$\frac{\psi'(1)}{\psi'(1/2)} < \frac{(C-1)\, n_2}{n_1}.$$

Proportional rule is suboptimal if $C$ is sufficiently large and/or $\psi$ is sufficiently concave.

This example illustrates that the optimal organization entails giving relatively more weights to small countries than proportionality would suggest. The assumption of decreasing marginal utility plays a key role. A stochastic model is introduced below to render the above ideas.

### 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 winning coalition with integer-weighted votes; see the books by Felsenthal and Machover (1998) and Laruelle and Valenciano (2008). Our focus is different, as can be seen from the example above. The point made in the present paper rests on the non-linearity of $\psi$. It should be contrasted with the other contributions which also derive an optimal rule from an explicit social criterion.

The first, and now classical, argument proposed in favor of degressive proportionality rests on statistical reasonings leading to the 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, a pattern that exhibits degressive proportionality (Penrose 1946). The mathematical reason why the square root appears in this literature is linked to the assumption made that, within each country, voters' opinions are independent random variables[2] (see Felsenthal and Machover (1998); Ramirez *et*

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

*al.* (2006); Słomczyński and Życzkowski (2010); Maaser and Napel (2011)). The political argument is that, in a world where frontiers have no link with the citizens' opinions, the representatives may as well be selected at random with no reference to these countries, but if representatives have to be chosen country-wise, then the focus should be on the statistical 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 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 the 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. Feix et. al. (2011) focuses on the majority efficiency, which is known as Condorcet efficiency in Social Choice Theory.

In Barberà and Jackson (2006), and Beisbart and Bovens (2007) the optimality is with respect to a sum of individual utilities, as in the present paper. 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 simple settings this provides weights which are simply proportional to the population size. In these contributions, the individual utilities to be summed at the collective level are, by assumption, linear in $p$. 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. This argument in favor of proportionality, called *pure majoritarian* in Laslier (2012), is different from what we wish to highlight here. If we could know in advance the importance for the various countries of the various issues to be voted upon, then we should change the countries' weights accordingly. Of course this is not possible at the constitutional stage, but notice that part of this intuition is endogenized in the setting we propose, along the following reasoning.

Start from weights strictly 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. The optimal weights should thus exhibit degressive proportionality.

### 1.4 This paper

Many existing apportionment rules show degressive proportionality, in principle and in fact.[3] The main contribution of this paper is to provide a theoretical foundation for the principle of degressive proportionality, which is not sensitive to knife-edge assumptions such as independence or linearity.

Penrose's square-root law is not robust in the following two aspects. First, it hinges on the assumption that the voters' preferences are independent random variables. Common sense suggests that this assumption is far from plausible. Even if not perfectly correlated, citizens of a country tend to have common interest because of geography, culture, economy, etc. The independence assumption is also empirically rejected by Gelman, Katz and Bafumi (2004). To see that the independence assumption is crucial to obtain Penrose Law, consider a group with population $n$. If there is a slight correlation in the preferences, we can show by an elementary computation[4] that the standard deviation of total utility in the group grows by the order of $n$. As will be shown later, the optimal weights which maximize the utilitarian social welfare are not proportional to the square-root of the population in such a case[5]. Only in the situation where voters' opinions are perfectly

[3] Leading examples are the US Electoral College, the European Union Council of Ministers, and the European Parliament. In countries with bicameral legislature, the upper house often uses equal representation while the lower house uses proportional representation. In combination, legislative power can be considered to be distributed with degressive proportionality.

[4] Suppose that $var\left(u_i\right) = \sigma^2$ for $\forall i$ and $cov\left(u_i, u_j\right) = \sigma_\varepsilon^2$ for $\forall i, \forall j \neq i$. Then $var\left(\sum_i u_i\right) = n\sigma^2 + n\left(n-1\right)\sigma_\varepsilon^2$ increases by the order of $n^2$ iff $\sigma_\varepsilon \neq 0$.

[5] This is in accordance with the findings of Beisbart and Hartmann (2010), who show by simulation that the interest group model (perfect correlation) of Beisbart and Bovens (2007) is stable, while the aggregate (independent) model is not.

independent, the standard deviation grows by the order of $\sqrt{n}$.

Second, and maybe more critically, it is commonly assumed in the apportionment literature that each individual's utility is additively separable over the issues, that is, the total utility assigned to an apportionment rule is the simple sum of the payoffs obtained in each issue. However, in general the marginal utility obtained from an additional success may well depend on the utility level attained by the rest of issues. In many economic situations, it is reasonable to assume that the marginal utility decreases.

**Our model brings the decreasing marginal utility assumption commonly used in Economics into Political Science. When the marginal utility is decreasing, the marginal importance of an additional issue for big countries is relatively smaller, since they have higher chances of winning in other issues. Degressive proportionality is obtained as the result of equalizing the marginal utility of the individuals across the countries with heterogeneous sizes so that the utilitarian social welfare is maximized. In Barberà and Jackson (2006), the optimal weights are shown to be proportional to the sum of the expected utilities in the country, which depends solely on the utility distribution, exogenously given, independent of the decision rule. In our model, the importance of the issues for each country is determined endogenously since we do not assume separability of the utility function defined over the sequences of the decisions taken under the constitutional rule. The importance of a certain issue for a country depends on the frequency with which this country is together with the majority of the society. Indeed, we show that the optimal weight is proportional to the endogenously determined importance. In this sense, our result is consistent with Barberà and Jackson (2006) *at the optimum*. However, the reasoning which leads to the support of degressive proportionality is precisely this endogeneity, not the importance exogenously defined by the distribution of the preferences. As a consequence, we provide a theoretical foundation for the principle of degressive proportionality, which is not sensitive to either the linearity of the utility or the independence assumption of the preference distribution across individuals.**

In section 2, we describe the model in which the uncertainty over citizens' opinions is described by a probabilistic distribution. Our main theorem is given in Section 3, where we first compute the optimal weights for two extreme, benchmark cases, and then show that the general cases fall between them. Detailed discussion as to the assumption on the shape of citizens' utility function is given in Section 4. Section 5 concludes. All proofs are provided in the Appendix.

# 2 The Model

## 2.1 Objective

There are $C$ countries, and country $c \in \mathcal{C} = \{1, \cdots, C\}$ has a population of $n_c$ individuals. Let $n = \sum_c n_c$ be the total population. We consider binary decision problems. The status quo is labeled as 0, and the alternative decision is labeled as 1. Each individual $i$ announces her favorite decision $X_i \in \{0, 1\}$, and the final decision is denoted by $d \in \{0, 1\}$.[6] A voting rule is used to take all such decisions so that, from the opinions stated by the voters, the final decision is in accordance with $i$'s preference with some frequency:

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

The main departure of this paper from the literature is that we treat this frequency as the object of preference for individual $i$, and we denote by $\psi(p_i)$ the utility attached to the constitution. We suppose that all individuals share the same utility function $\psi$ and make the usual assumption of decreasing marginal utility. Detailed discussions on these assumptions are provided in Section 4.

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 $\psi$ 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 assumption on individuals' utility.[7]

[6] Since there are only two alternatives, voting for the favorite decision is a dominant strategy. The voting game is dominance-solvable and truthful voting is the unique admissible strategy.

[7] 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, although it can be seen as the limit case where the concavity of $\psi$ goes to infinity.

## 2.2 Probabilistic Opinion Model

In order to model the correlations between individual opinions, we use a probabilistic opinion model. More precisely, we assume that individual preferences $(X_i)_{i=1}^n \in \{0,1\}^n$ are drawn from a joint distribution $f(X_1, \cdots, X_n)$. We focus our attention on a class of distributions with the following two properties: (i) all individuals are ex ante unbiased with respect to the two alternatives,[8] and (ii) the preferences are positively correlated within countries, but independent across countries. We use the parameter $\mu$ to describe the intra-country correlation.

Suppose that voters in country $c$ receive a country-specific signal $Y_c \in \{0,1\}$, and each voter $i$ in her country $c(i)$ forms an opinion conditionally on $Y_{c(i)}$. The conditional probability $\mu$ for a voter to follow her country-specific signal is the same for every voter in every country, and for both alternatives:

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

We assume that $\mu$ is larger than $1/2$, so that $Y_c$ can indeed be interpreted as the general opinion in country $c$. We could have started the other way round and, instead of taking the country's general opinion as a primitive, we could have specified a probability distribution for the correlated opinions of the citizens of country $c$. Then $Y_c$ would be defined as the majority value of the variables $X_i$ for $i \in c$. But since we are dealing with large numbers of individuals (between .4 and 100 millions per country), it is much simpler to take $Y_c$ as the primitive.

The variables $Y_c \in \{0,1\}$ are assumed to follow the Bernoulli distribution with parameter $1/2$, and to be 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 point can be defended in two ways. First, the way some countries' interests are aligned is itself variable: on some issues larger countries are opposed to smaller ones, other issues divide rich countries against poor ones, East against West, 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. (See Laruelle and Valenciano, 2005, and Barr and Passarelli, 2009.) We will discuss

[8] If there is a known bias to one of the two alternatives, welfare-maximizing decision is rather obvious, since the right choice is the preferred alternative and thus the society has little interest to take a vote. Most interesting are the cases in which the voters are unbiased ex ante so that voting works as a device to aggregate preferences.

in the conclusion the consequences of relaxing this assumption.

### 2.3 Weighted Voting Rules

Each country $c$ has a weight $w_c$. Without loss of generality we can normalize the weights so that

$$\sum_{c=1}^{C} w_c = 1.$$

We introduce two weighted decision models. 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 the total weight of the countries who voted for the proposition is strictly larger than a threshold $t$, and the decision $d = 0$ is taken if the total weight of the countries who voted against the proposition is strictly larger than $1-t$. For $Y = (Y_c)_{c\in\mathcal{C}}$,

$$d^{\text{council}}\left(Y|w,t\right) = \begin{cases} 1 \text{ if } \sum_c w_c Y_c > t \\ 0 \text{ if } \sum_c w_c Y_c < t \end{cases}.$$

When the threshold is exactly met, $d = 1$ is taken with a pre-specified probability that depends on the realization of $Y$.

In the **Parliament** model, the country $c$ 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 the total weight of the representatives who voted for is larger than the threshold $t$:

$$d^{\text{parliament}}\left(Y|w,t\right) = \begin{cases} 1 \text{ if } \sum_c w_c\left(\mu Y_c + (1-\mu)(1-Y_c)\right) > t \\ 0 \text{ if } \sum_c w_c\left(\mu Y_c + (1-\mu)(1-Y_c)\right) < t \end{cases}.$$

Indeed, these two models are equivalent up to the threshold.

**Proposition 1** $d^{parliament}\left(Y|w,t\right) = d^{council}\left(Y \middle| w, \frac{t-(1-\mu)}{2\mu-1}\right).$

Proof is immediate.[9] If $t < 1-\mu$ or $t > \mu$ in the Parliament model, the decision is either $d = 0$ or $d = 1$ regardless of the realized values of $Y$. It is as if $t < 0$ or $t > 1$ in the Council model.

Note that if the threshold is $1/2$, the two models are identical. When a weighted voting rule has the threshold $t = 1/2$, we call it a *weighted*

[9] $\sum_c w_c\left(\mu Y_c + (1-\mu)(1-Y_c)\right) > t \Leftrightarrow \sum_c w_c Y_c > \frac{t-(1-\mu)}{2\mu-1}$.

*majority rule*. Weighted majority rules keep the symmetry between the two alternative decisions, up to in the limit case where votes are exactly split. Notice that some voting rules are not even weighted. However, it will be proven that the optimal voting rules are indeed weighted majority rules, i.e. weighted, with threshold 1/2.

The central idea of this paper is the degressive proportionality.

**Definition 1** *Weights are said to exhibit degressive proportionality to the population if*

$$n_c < n_{c'} \Rightarrow w_c \leq w_{c'} \text{ and } \frac{w_c}{n_c} \geq \frac{w_{c'}}{n_{c'}}.$$

### 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 ($n = (n_c)_{c \in \mathcal{C}}$), the intra-country homogeneity ($\mu$), and the utility function ($\psi$). For each model $M \in \{\text{Council, Parliament}\}$, the expected social welfare is:

$$U(w,t) = \sum_i \psi(p_i) = \sum_c n_c \psi\left(\pi_c^M(w,t)\right), \quad (1)$$

with

$$\pi_c^M(w,t) = \Pr\left[X_i = d^M(w,t)\right] \quad (2)$$

for any citizen $i$ in country $c$. Therefore, our problem is to choose optimal weights $w$ and the threshold $t$:[10]

$$\max_{(w,t)} U(w,t). \quad (3)$$

## 3 Optimal weights in theory

In this Section, we first characterize the optimal weights for two extreme cases: linear utility and the Rawlsian social welfare in Section 3.1. Our main result, obtained in Section 3.2, is stated in the general framework of probabilistic simple games. This class of games precisely describes how ties are broken, and also contains non-weighted games. We prove that the optimal games in that class are weighted, with weights which exhibit degressive proportionality.

[10] For the description to be complete, the tie-breaking rule should be specified, although our main focuses are the weight vector and the threshold.

### 3.1 Two benchmarks

**The linear case**

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

**Proposition 2** *If $U = \sum_i p_i$ the optimal decision rule is a weighted majority, with weights $w_c$ proportional to the population.*

This result is compatible with the existing models, such as Barberà and Jackson (2006) or Fleurbaey (2008). Notice that the result applies to any $\mu$ strictly larger than 1/2. If we allow $\mu = 1/2$, then the model is equivalent to the aggregate (independence) model of Beisbart and Bovens (2007), in which the optimal weights are proportional to the square-root of the population. But even a slight degree of correlation in the distribution of preferences implies that the optimal weights are proportional to the population. Proposition 2 gives evidence which indicates that Penrose's square-root law hinges on the independence assumption when the utility function is assumed to be a linear function of the number of successes.

**The Rawlsian case**

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

**Proposition 3** *For any $\mu > 1/2$, if $U = \min_i p_i$ the optimal decision rule is the simple majority among countries: all countries have equal weight.*

The Rawlsian case corresponds to the limit where the concavity of $\psi$ goes to infinity. Obviously, equal weight is an extreme example of degressive proportionality, where $w_i/n_i$ decreases most rapidly among all degressively proportional rules.

### 3.2 Optimal Apportionment and Simple Games

We introduce the concept of weighted probabilistic simple games corresponding to the weighted voting rules. For each realization of $C$ Bernoulli variables $(Y_1, Y_2, ..., Y_C)$, we can naturally associate the subset of countries (or coalition) for which the Bernoulli variable takes the value 1: $\{c | Y_c = 1\}$. For any of the $2^C$ possible coalitions, the social decision can be either to accept or to reject the proposal. The problem can thus be viewed as the selection of a subset $\Gamma \subset \mathcal{P}(\mathcal{C})$ of winning coalitions, the coalitions for which the proposal is accepted. For any $\Gamma$, the pair $(\mathcal{C}, \Gamma)$ is called a *simple game.*[11] In the corresponding voting rule $d^\Gamma$, the decision $d = 1$ is taken if and only if the coalition of countries which vote in favor of the proposal belongs to $\Gamma$:

$$d^\Gamma = 1 \Leftrightarrow \{c | Y_c = 1\} \in \Gamma.$$

Here we generalize the concept of simple games to allow probabilistic decision rules. For any coalition $S$, we define the probability $q(S)$ that the society accepts the proposition ($d = 1$) when the countries in coalition $S$ vote for it. We call the corresponding function $q : \mathcal{P}(\mathcal{C}) \rightarrow [0, 1]$ a *probabilistic simple game.*

Denote by $\mathcal{PSG}$ the set of probabilistic simple games. Notice that any $q$ in $\mathcal{PSG}$ can be uniquely assimilated to a vector in $[0, 1]^{2^C}$ (and vice versa). A simple game $\Gamma$ is a probabilistic simple game $q$ such that $q(S) = 1$ for any $S \in \Gamma$ and $q(S) = 0$ for any $S \notin \Gamma$. When the deterministic decisions in a probabilistic simple game can be represented by a system of weights, we say that it is a *weighted probabilistic simple game*:

**Definition 2** *A probabilistic simple game $q$ is weighted if there exists a vector of weights $w \in \mathbb{R}^C$ and a threshold $t \in [0, 1]$ such that for any $S \subset \mathcal{C}$,*

$$\sum_{c \in S} w_c > t \quad \Rightarrow \quad q(S) = 1,$$
$$\sum_{c \in S} w_c < t \quad \Rightarrow \quad q(S) = 0.$$

*The subset of coalitions for which the total weight equals the threshold is called the tie set: $T(w, t) = \{S \subset \mathcal{C} \mid \sum_{i \in S} w_i = t\}$. The restriction of $q$ on $T(w, t)$ is called the tie-breaking rule.*

The benefit of considering a probabilistic simple game is straightforward. If we consider only the (deterministic) simple games, we face a maximization problem in which we choose the set of winning coalitions $\Gamma$. Providing an analytical solution to such discrete problems is quite demanding, and computation for large values of $C$ is practically impossible in general. Instead, by considering a larger set of games over the continuous space $[0, 1]^{2^C}$, we can provide an analytical solution.

[11]In what follows, we will omit $\mathcal{C}$ and simply write $\Gamma$.

Of course, the advantage is obtained at the expense of certain costs. A potential problem may be that, by considering the entire set of probabilistic simple games, the optimal game may lie outside of the set of all weighted games. Our original motivation is to find the optimal weights, and indeed there exist many probabilistic simple games which are not weighted. However, in the following we show that the optimal games chosen over the entire set of probabilistic simple games are indeed weighted, and the weights exhibit degressive proportionality, provided that $\psi$ is concave.

For any vector of weights $w \in \mathbb{R}^C$ and any threshold $t \in [0,1]$, we denote by $\mathcal{PSG}(w,t)$ the corresponding set of weighted probabilistic simple games. As is clear by definition, any weighted decision rule can be described as a weighted probabilistic simple game and vice versa. Especially, any weighted voting rule in the Council model can be described as a weighted probabilistic simple game $q \in \mathcal{PSG}(w,t)$. For the Parliament model, any weighted voting rule with weight vector $w$ and threshold $t$ can be described as a weighted probabilistic simple game $q \in \mathcal{PSG}\left(w, \frac{t-(1-\mu)}{2\mu-1}\right)$.[12]

For any vector of population $n = (n_c)_{c\in\mathcal{C}}$, intra-country homogeneity $\mu$, concave utility function $\psi$, and the model $M$, we denote by $(n,\psi,\mu,M)$ the corresponding utilitarian problem over the set of all probabilistic simple games:

$$\max_{q\in\mathcal{PSG}} \sum_{c\in\mathcal{C}} n_c \psi\left(\pi_c(q)\right) \tag{4}$$

where $\pi(q) = (\pi_1(q), ..., \pi_C(q))$ is a function defined over $\mathcal{PSG}$, exactly in the same way as (2).

**Proposition 4** *Let $q^*$ be any solution of the problem $(n,\psi,\mu,M)$.*

*(i) The associated vector of the frequency of success, $\pi^* = \pi(q^*)$ is the same for any $q^*$.*

*(ii) For any two countries $c$ and $c'$ in $\mathcal{C}$, $n_c < n_{c'} \Rightarrow \pi^*_c \leq \pi^*_{c'}$.*

We now state the main result of the paper. We show that any solution to the utilitarian apportionment problem is a weighted majority rule (i.e. threshold is 1/2), with a vector of weights that exhibits degressive proportionality.

**Theorem 1** *Assume that the expected utility attached to a constitutional rule is a concave function of the frequency of success. Define the weight*

[12] See Proposition 1.

*vector $w^*$ so that $w_c^*$ is proportional to $n_c\psi'(\pi_c^*)$, where $\pi^*$ is uniquely determined in Proposition 4. Any solution $q^*$ of $(n, \psi, \mu, M)$ is a weighted probabilistic simple game with the weights $w^*$ and the threshold $t^* = 1/2$. Moreover, $q^*$ is unique up to the tie-breaking rule.*

Since $\psi$ is concave, an immediate corollary is the following:

**Corollary 2** *The optimal weight $w^*$ exhibits degressive proportionality.*

Moreover, since the optimal threshold is $1/2$, the Council model and the Parliament model are equivalent (Proposition 1). We thus obtain another Corollary.

**Corollary 3** *Given $n$, $\psi$ and $\mu$, the optimal weights are the same in both the Council model and the Parliament model.*

As we mentioned above, some probabilistic simple games cannot be described by any weighted voting rule, but any weighted voting rule can be described as a probabilistic simple game. Therefore, we have

$$\max_{(w,t)} U(w,t) \leq \max_{q \in \mathcal{PSG}} \sum_{c \in \mathcal{C}} n_c \psi(\pi_c(q)).$$

Theorem 1 implies that $(w^*, 1/2)$ is a solution of our original problem (3). It also provides a formula that characterizes the optimal weights $(n_c\psi'(\pi_c^*))_{c \in \mathcal{C}}$ and threshold $(1/2)$, although it is silent about the tie-breaking rule. Unfortunately, even with this formula, obtaining the exact values of the optimal weights is challenging, because the probabilities of success $\pi_c^*$ depend, themselves, on the weights. Computation of the optimal weights for an concrete example may require some methodological technique. See, for example, Mac and Treibich (2012) for a detailed discussion and examples.

# 4 Discussions on the shape of the utility function

The main departure of this paper from the literature is the assumption that the utility assigned to a constitutional rule by each individual is a concave function of the frequency of success: $\psi(p)$.

## 4.1 Frequency of success

First, we argue that the frequency of success is the key variable which describes the utility level attached to the constitutional rule. This frequency, called Rae Index (Rae 1969), is one of the most commonly used indices in the literature of voting power measurement. Literature agrees that success and decisiveness are the two major factors which measure voting power, with the Banzhaf-Penrose Index ($BZ_i$) as one of the most widely accepted indices of the latter. Dubey and Shapley (1979) show that these two indices have a simple affine relationship, $p_i = (1 + BZ_i)/2$, when all vote configurations are equally likely. Since our focus in this paper is the constitutional design, under which the voters are treated symmetrically, independently and without any ex ante inclination to yes or no, the frequency is the right measure of both success and decisiveness.

Some authors have argued that individuals attach an intrinsic value to the fact of being decisive (e.g. Sen (1985), Hausman and McPherson (1996), Kolm (1998)). Our approach of course does not contradict this idea, but in fact we do not need such a departure from standard consequentialism. In what follows the frequency of success appears as the pertinent parameter to use in order to describe the usual outcome-utility.

## 4.2 Independence and separability

A series of decisions are made under a constitutional rule. Certainly the overall utility attached to a series of decisions may be different from the simple sum of the payoffs obtained from each decision. In general, the total utility assigned to the constitutional rule may be a non-linear function of the frequency of success.

To see how, let us consider a simple example which makes clear the main idea of this paper. Suppose that a constitutional rule is used for $K$ independent decisions, each of which is worth one dollar or zero. The possible payoffs range from 0 to $K$ dollars. If $p_i$ is the probability that the collective decision matches the individual $i$'s will, the probability of earning $k$ dollars is $\binom{K}{k} p_i^k (1 - p_i)^{K-k}$. Let $u(k)$ denote the von Neumann and Morgenstern utility attached to payoff $k$. The expected utility is then a function of $p_i$, the frequency with which each social decision matches the individual's will:

$$\psi(p_i) = \sum_{k=0}^{K} \binom{K}{k} p_i^k (1 - p_i)^{K-k} u(k).$$

Elementary calculus shows that the function $\psi$ is concave if the marginal utility $u(k+1) - u(k)$ is decreasing. More generally, the same intuition continues to hold even when the importance of the decisions are heterogeneous. If the marginal gain from an additional success depends on the welfare level attained by the rest of the decisions, and if the assumption of decreasing marginal utility holds, then the preferences are represented by a submodular von Neumann and Morgenstern utility function defined over the sequences of successes. For example, if an individual expects to have a large number of successes under a constitutional rule, then the marginal gain from an additional potential success is smaller than in the situation in which she expects to have a small number of successes.

To be more precise, define the success variable of individual $i$ for decision $k$ as $z_i^k = 1$ for a success $\left(X_i^k = d^k\right)$ and $z_i^k = 0$ for a failure $\left(X_i^k \neq d^k\right)$. Payoff $u_i : \{0,1\}^K \to \mathbb{R}$ is defined over the sequences of successes and failures $z_i = (z_i^1, \cdots, z_i^K)$. The following proposition guarantees that the assumption of decreasing marginal utility $u_i$ boils down to the concavity of the utility function $\psi_i$.

**Proposition 5** *Suppose that $u_i$ is increasing and submodular. Then $\psi_i$ is increasing and concave.*

Whereas decreasing marginal utility is one of the most commonly used assumptions in Economics, somehow departure from separable utility (i.e. linearity with respect to $p_i$) in the literature of apportionment problems has not been much discussed. One of our main contributions in this paper is to take it into account explicitly.

**Notice that, in practice, decisions taken to vote may be linked so that the independence assumption across decisions is violated. For example, decisions about repeated military sanctions may well be such that they will be valid only if a certain number of actions are taken repeatedly.[13] Then, for individuals who favors the sanctions, the overall payoff from these decisions exhibits nonconvexity as a function of the number of successes. Increasing returns in the efficiency of repeated decisions translates into correlations in preferences across decisions. Such questions raise the important issue of the consistency of decisions taken (by voting) in collective decision bodies. Remark that if several decisions are so technically linked in a same issue that preferences are perfectly correlated, then the same individuals will vote for, or against, in**

[13]We thank one of the anonymous referees for suggesting the example.

**a consistent way, so that the decision will be taken by the same group of individuals.[14] Then the overall decision will be consistent, unless the decision rule involves probabilistic tie-breaking (and as far as the individuals are themselves consistent). The question of probabilistic tie-breaking is not a practical concern since, even if the optimal rules that we propose do require probabilistic tie-breaking in theory, the same theory indicates that the probability of a tie is minuscule. .....DONNER LES CHIFFRES.... Therefore it is reasonable to consider that linked decisions are put together into a bundle, which we call a *theme*. Suppose that the preferences are perfectly correlated within a theme and less across themes. Each individual in the society has consistent preferences over, say, economic issues, and these preferences are more or less independent of those over, say, religious issues. Proposition 5 shows that concavity of $\psi$ is obtained if the primitive utility function $u$ is submodular.**

### 4.3 Egalitarianism

The concavity of $\psi$ can as well 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 her social objective, the maximization of a Kolm-Atkinson index of the form:

$$W = \sum_i \psi(u_i).$$

The social objective $W$ is egalitarian if any Pigou-Dalton transfer increases its value. We recall without proof the following result, well-known from the theory of inequality measurement (see Dutta 2002). The social objective is egalitarian if and only if the function $\psi$ is concave, for instance $\psi(u_i) = u_i^\alpha$ for $0 < \alpha < 1$.

**Proposition 6** *$W$ is egalitarian if and only if $\psi$ is concave.*

As put forth by Bentham (1822)[15]:

[14] Logical aggregation problems through majority voting only occur when the majority winning coalitions vary with the issue (Condorcet 1785, Mongin 2012).

[15] 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.

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 on a purely utilitarian preference that takes into account decreasing marginal utility. These two concepts deserve a unified name, and it is called the *utilitarian-egalitarian argument* in Laslier (2012).

An extreme, degenerated case is the Rawlsian objective of maximizing the well-being of the worst-off individual. This case is obtained when $\alpha$ tends to 0, and we show that it implies identical weights for all countries in Proposition 3 above.

# 5 Conclusion

This paper gives a theoretical foundation for the principle of degressive proportionality in the optimal apportionment problem. We consider a model in which the individual utility is a function of the frequency of success in binary decisions, and assume that marginal utility is decreasing. By doing so, we provide a proof which does not hinge on the independence assumption on the distribution of the individual preferences. We believe that our paper provides fundamental support for the degressive proportionality which is currently practiced in many apportionment problems.

Our result includes two important benchmark cases in the literature: in the limit where the concavity diminishes (linear utility), the optimal weights are proportional to the population (except the knife-edge case of zero interdependence); e.g. Barberà and Jackson (2006), Fleurbaey (2008), and the interest group model in Beisbart and Bovens (2007). To the contrary, in the limit where the concavity goes to infinity (MaxMin utility), the optimal weights are equal for all countries. Obviously these two weight profiles are the extreme examples of degressive proportionality, and all the utility functions between the two examples above induce degressive proportionality in between.

These results have been obtained under the assumption that opinions are independent between countries. It should be clear that allowing for any kind of correlation between countries would destroy the result. For instance, suppose that the independence assumption holds except for a given subset of countries, which are, on the contrary, perfectly correlated. Then the above model applies if we treat this set of countries as one large country, summing the populations. Then the optimal weights per country have no reason to be degressively proportional. Nevertheless, it is true that if the optimal values of the probabilities $p_i$ are increasing with the populations, then the optimal weights are degressively proportional. This point is proven in the appendix, as a remark in the proof of the main theorem. Such a paradoxical situation, where a larger country is satisfied less often than a smaller one, cannot happen under independence or if correlations between countries are small.

The next step is to investigate more general conditions which would support the degressive proportionality principle. For example, double correlation within the countries and within the political parties across the countries is a substantial issue in European politics. Integrating these aspects would be in the future research agenda.

# A Appendix

## A.1 Proofs

**Proof of Proposition 2.** 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

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

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 3.** By Proposition 2, if $n_c = 1$ for all $c$, the simple majority rule with the equal weight maximizes the sum of the frequencies. 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 all $c$, implying $C p^{eq} < \sum_c \pi_c$, a contradiction. Therefore, $\min_c \pi_c \leq p^{eq}$ for any rule. Hence, $\max \min_c \pi_c \leq p^{eq}$. The maximum is attained by the equal weight. ■

**Proof of Proposition 4.** Let us remind that $\pi : \mathcal{PSG} \to [0,1]^C$ is the function defined in (4). For any individual $i$ in country $c$,

$$\begin{aligned}\pi_c(q) &= \Pr[X_i = Y_c]\Pr[Y_c = d(q)] + \Pr[X_i \neq Y_c]\Pr[Y_c \neq d(q)] \\ &= 1 - \mu + (2\mu - 1)\Pr[Y_c = d(q)]. \end{aligned} \tag{5}$$

Given a probabilistic simple game $q$, $q(S)$ is the probability that $d = 1$ is chosen. Therefore,

$$\begin{aligned}\Pr[Y_c = d(q)] &= \sum_S \Pr(S)\left(q(S)\mathbf{1}_{\{c\in S\}} + (1-q(S))\mathbf{1}_{\{c\notin S\}}\right) \\ &= \sum_{\{S|c\in S\}} \Pr(S)\, q(S) + \sum_{\{S|c\notin S\}} \Pr(S)\,(1-q(S)) \end{aligned} \tag{6}$$

where $\Pr(S)$ denotes the probability that the set $\{c|Y_c = 1\}$ coincides with $S \subset \mathcal{C}$. Notice that $\pi_c$ is affine in $q$. Hence, the image $\pi(\mathcal{PSG})$ is convex in $[0,1]^C$.

Since $\psi$ is strictly concave, the maximization problem $\sum_c n_c \psi(\pi_c)$ subject to $\pi \in \pi(\mathcal{PSG})$ has a unique solution $\pi^*$. Any solution $q^*$ of the problem $(n, \psi, \mu, M)$ satisfies $\pi^* = \pi(q^*)$.

Suppose now that there exists $c, c' \in \mathcal{C}$ with $n_c < n_{c'}$ and $\pi^*_c > \pi^*_{c'}$. Consider then $\widehat{q}$ defined by $\widehat{q}(\sigma_{cc'}(S))$, where $\sigma_{cc'}$ is the permutation of $\mathcal{C}$ that exchanges $c$ and $c'$. We get $\pi_c(\widehat{q}) = \pi^*_{c'}$, $\pi_{c'}(\widehat{q}) = \pi^*_c$ and $\pi_k(\widehat{q}) = \pi^*_k$, $\forall k \neq c, c'$. Then, $\sum_{c\in\mathcal{C}} n_c \psi(\pi_c(\widehat{q})) > \sum_{c\in\mathcal{C}} n_c \psi(\pi^*)$, which contradicts the optimality of $\pi^*$. ■

**Proof of Theorem 1.** Let $q^*$ be a solution, and $\pi^* = \pi(q^*)$ be the corresponding vector of frequency of success. We can write the first order conditions maximizing $U(q) = \sum_{c\in\mathcal{C}} n_c \psi(\pi_c(q))$ over $2^C$ variables $(q(S))_{S\subseteq\mathcal{C}}$. At the optimum, we have:

$$\begin{aligned}\frac{\partial U}{\partial q(S)}(q^*) > 0 &\;\Rightarrow\; q^*(S) = 1, \\ \frac{\partial U}{\partial q(S)}(q^*) < 0 &\;\Rightarrow\; q^*(S) = 0.\end{aligned}$$

By (5) and (6), we can explicitly compute the partial derivatives of $U$:

$$\frac{\partial U}{\partial q(S)}(q) = (2\mu - 1)\Pr(S)\left(\sum_{c\in S} n_c \psi'(\pi_c(q)) - \sum_{c\notin S} n_c \psi'(\pi_c(q))\right).$$

Hence, $\forall S \subset \mathcal{C}$,

$$\sum_{c \in S} n_c \psi' \left(\pi_c \left(q^*\right)\right) > \sum_{c \notin S} n_c \psi' \left(\pi_c \left(q^*\right)\right) \quad \Rightarrow \quad q^*(S) = 1,$$
$$\sum_{c \in S} n_c \psi' \left(\pi_c \left(q^*\right)\right) < \sum_{c \notin S} n_c \psi' \left(\pi_c \left(q^*\right)\right) \quad \Rightarrow \quad q^*(S) = 0,$$

which is equivalent to:

$$\sum_{c \in S} n_c \psi'(\pi_c \left(q^*\right)) > \frac{1}{2} \sum_{c \in \mathcal{C}} n_i \psi'(\pi_c \left(q\right)) \quad \Rightarrow \quad q^*(S) = 1,$$
$$\sum_{c \in S} n_c \psi'(\pi_c \left(q^*\right)) < \frac{1}{2} \sum_{c \in \mathcal{C}} n_i \psi'(\pi_c \left(q\right)) \quad \Rightarrow \quad q^*(S) = 0.$$

Defining the vector of weights $w^*$ by $w_c^* = \dfrac{n_c \psi'(\pi_c^*)}{\sum_{c' \in \mathcal{C}} n_{c'} \psi'(\pi_{c'}^*)}$ $\forall c \in \mathcal{C}$, we conclude that:

$$\sum_{c \in S} w_c^* > \frac{1}{2} \quad \Rightarrow \quad q^*(S) = 1,$$
$$\sum_{c \in S} w_c^* < \frac{1}{2} \quad \Rightarrow \quad q^*(S) = 0,$$

meaning that the probabilistic simple game $q^*$ is weighted and can be represented by the vector $w^*$ and the threshold 1/2: $q^* \in \mathcal{PSG}\left(w^*, 1/2\right)$. Furthermore, by Proposition 4, we know that for any $c, c' \in C$ with $n_c < n_{c'}$, $\pi_c^* \leq \pi_{c'}^*$, which implies in turn that $w_c^*/n_c = \psi'(\pi_c^*) \leq \psi'(\pi_{c'}^*) = w_{c'}^*/n_{c'}$ because of the concavity of $\psi$.

The last thing we need to show is that the vector $w^*$ is increasing. Let $c$ and $c'$ be two countries such that $n_c \leq n_{c'}$, and assume that $w_c^* = n_c \psi'(\pi_c^*) > n_{c'} \psi'(\pi_{c'}^*) = w_{c'}^*$.

As a first step, let us show that there always exists a coalition $S$ such that $c \in S$, $c' \notin S$ and $q^*\left(S\right) < q^*\left(\sigma_{cc'}\left(S\right)\right)$. By contradiction, assume that for any $S$ which contains $c$ but not $c'$, $q^*\left(S\right) \geq q^*\left(\sigma_{cc'}\left(S\right)\right)$. By (6),

$$\begin{aligned}
&\Pr\left[Y_c = d\left(q^*\right)\right] \\
&= \sum_{\{S | c, c' \in S\}} \Pr\left(S\right) q^*(S) + \sum_{\{S | c, c' \notin S\}} \Pr\left(S\right) \left(1 - q^*(S)\right) \\
&+ \sum_{\{S | c \in S, c' \notin S\}} \Pr\left(S\right) q^*(S) + \sum_{\{S | c \notin S, c' \in S\}} \Pr\left(S\right) \left(1 - q^*(S)\right).
\end{aligned}$$

Then,

$$\begin{aligned}&\Pr\left[Y_c = d\left(q^*\right)\right] - \Pr\left[Y_{c'} = d\left(q^*\right)\right]\\&= \sum_{\{S|c\in S, c'\notin S\}} \left\{\Pr\left(S\right)\left(2q^*(S)-1\right) + \Pr\left(\sigma_{cc'}\left(S\right)\right)\left(1-2q^*\left(\sigma_{cc'}\left(S\right)\right)\right)\right\} \geq 0.\end{aligned}$$

Note that $\Pr\left(S\right) = \frac{1}{2^C} = \Pr\left(\sigma_{cc'}\left(S\right)\right)$. Using (5), this implies $\pi_c^* \geq \pi_{c'}^*$. By Proposition 4, we know that $\pi_c^* \leq \pi_{c'}^*$. Therefore, $\pi_c^* = \pi_{c'}^*$, which implies $w_c^* \leq w_{c'}^*$, a contradiction.

Now, pick a coalition $S$ containing $c$ but not $c'$ with $q^*\left(S\right) < q^*\left(\sigma_{cc'}\left(S\right)\right)$. Because of this last inequality, it is always possible to define another game $q'$ by:

$$\begin{aligned}q'\left(S\right) &= q^*\left(S\right) + \varepsilon,\\q'\left(\sigma_{cc'}\left(S\right)\right) &= q^*\left(\sigma_{cc'}\left(S\right)\right) - \varepsilon,\\q'(T) &= q^*(T),\ \forall T \neq S, \sigma_{cc'}\left(S\right).\end{aligned}$$

Then, we have:

$$\begin{aligned}\pi_c(q') &= 1-\mu+(2\mu-1)\Pr\left[Y_c = d\left(q'\right)\right]\\&= 1-\mu+(2\mu-1)\Big(\Pr\left[Y_c = d\left(q^*\right)\right] + \Pr(S)\varepsilon - \Pr(\sigma_{cc'}(S))(-\varepsilon)\Big)\\&= \pi_c(q^*) + 2(2\mu-1)\Pr(S)\varepsilon,\\\pi_{c'}(q') &= \pi_{c'}(q^*) - 2(2\mu-1)\Pr(S)\varepsilon,\end{aligned}$$

and $\pi_k(q') = \pi_k(q^*)\ \ \forall k \neq c, c'$. Denoting $\kappa = 2(2\mu-1)\Pr(S)\varepsilon$ and $\Delta U = U(q') - U(q^*)$ we get:

$$\begin{aligned}\Delta U &= n_c\left[\psi\left(\pi_c(q^*)+\kappa\right) - \psi\left(\pi_c(q^*)\right)\right] - n_{c'}\left[\psi\left(\pi_{c'}(q^*)\right) - \psi\left(\pi_{c'}(q^*)-\kappa\right)\right]\\&= \kappa\left[n_c\psi'\left(\pi_c\left(q^*\right)\right) - n_{c'}\psi'\left(\pi_{c'}\left(q^*\right)\right)\right] + o_{\kappa\to 0}(\kappa).\end{aligned}$$

By assumption, $n_c\psi'(\pi_c(q^*)) > n_{c'}\psi'(\pi_{c'}(q^*))$. Hence, choosing $\varepsilon$ small enough, we can find $q'$ such that: $U(q') > U(q^*)$. This contradicts the optimality of $q^*$. ■

**Proof of Proposition 5.** Since the index $i$ is obvious and redundant in this Proposition, we remove it in the proof. Let us denote by $|z|$ the number of successes in the sequence $z$: $|z| = \sharp\left\{\ell | z_\ell = 1\right\}$, and by $Z_k$ the set of sequences with $k$ successes:

$$Z_k = \left\{ z \in \{0,1\}^K \middle| |z| = k \right\}.$$

By definition, the expected utility is:

$$\begin{aligned}\psi(p) = \mathbb{E}\left[u\right](p) &= \sum_{z\in\{0,1\}^K} p^{|z|}(1-p)^{K-|z|}u(z) \\ &= \sum_{k=0}^{K} \binom{K}{k} p^k(1-p)^{K-k}U_k\end{aligned}$$

where $U_k = \sum_{z\in Z_k} u(z)/\binom{K}{k}$ is the average of the payoffs obtained from the sequences $z \in Z_k$.

It is straightforward to show that

$$\begin{aligned}\psi'(p) &= K\sum_{k=0}^{K-1}\binom{K-1}{k}p^k(1-p)^{K-1-k}(U_{k+1}-U_k), \\ \psi''(p) &= K(K-1)\sum_{k=0}^{K-2}\binom{K-2}{k}p^k(1-p)^{K-2-k}\left\{(U_{k+2}-U_{k+1})-(U_{k+1}-U_k)\right\}.\end{aligned}$$

Hence, to show that $\psi' > 0$ and $\psi'' < 0$, it suffices to show that $U_k$ is increasing and convex in $k$. Let $k \leq K-1$ be fixed, and we have:

$$\forall z \in Z_k, \text{ if } z_j = 0 \text{ then } u(z) < u(z+\mathbb{1}_j)$$

where $\mathbb{1}_j$ is a vector with an entry 1 at $j$-th component and 0 otherwise. Let us write these inequalities for all sequences $z \in Z_k$ and, given $z$, for all $j$ such that $z_j = 0$. On the left side we obtain the utilities of all $z \in Z_k$, each one appears $(K-k)$ times (for a given $z$, there are $(K-k)$ corresponding $j$). On the right side, we obtain the utilities of all $z \in Z_{k+1}$, each one appearing $(k+1)$ times in this column. Finally, if we sum up these two columns, we have:

$$\begin{aligned}& (K-k)\sum_{z\in Z_k}u(z) < (k+1)\sum_{z\in Z_{k+1}}u(z) \\ \Leftrightarrow\ & (K-k)\binom{K}{k}U_k < (k+1)\binom{K}{k+1}U_{k+1} \\ \Leftrightarrow\ & U_k < U_{k+1}.\end{aligned}$$

Thus $\psi$ is increasing. The intuition is the following; the left (resp. the right) term should be some constant times the average $U_k$ (resp. $U_{k+1}$). Because the two columns have the same length, these two constants should be equal.

Now, fix $k \leq K-2$, we have from the submodularity of $u$:

$$\forall z \in Z_k, \text{ if } z_l = z_j = 0 \text{ and } l \neq j, \text{ then } u(z+\mathbb{1}_l+\mathbb{1}_j)-u(z+\mathbb{1}_l) < u(z+\mathbb{1}_j)-u(z).$$

For the same reasons as before, we obtain $U_{k+2} - U_{k+1} < U_{k+1} - U_k$, this implies that $\psi$ is strictly concave. ■

## References

[1] Barberà, Salvador 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] Barr, Jason and Francesco Passarelli (2009) "Who has the power in EU?" *Mathematical Social Sciences* 57: 339-366.

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

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

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

[6] Cichocki, Marek A. and Karol Życzkowski eds.(2010) *Institutional Design and Voting Power in the European Union*. Ashgate, London.

[7] Condorcet. (1785) "Essai sur l'application de l'analyse la probabilit des dcisions rendues la pluralit des voix, "Paris: Imprimerie Royale.

[8] Dubey, Pradeep and Lloyd S. Shapley (1979) "Mathematical Properties of the Banzhaf Power Index " *Mathematics of Operations Research* 4(2), 99-131.

[9] Duff, Andrew (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.

[10] Duff, Andrew (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.

[11] Dutta, Bhaskar (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.

[12] Engel, Konrad (1997) “Sperner theory ” *Encyclopedia of Mathematics and its Applications, Volume 65,* Cambridge University Press.

[13] Feix, M.R., D. Lepelley, V. Merlin, J.L. Rouet and L. Vidu (2011) “Majority efficient representation of the citizens in a federal union” *mimeo.*

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

[15] Felsenthal, Dan and Moshé Machover (1999) “Minimizing the mean majority deficit: the second square root rule” *Mathematical Social Sciences* 37: 25-37.

[16] Fleurbaey, Marc (2008) “One stake one vote” *mimeo.*

[17] Gelman, Andrew, Jonathan N. Katz and Joseph Bafumi (2004) “Standard Voting Power Indexes Do Not Work: An Empirical Analysis” *British Journal of Political Science* 34: 657-674.

[18] Grimmet, Geoffrey, 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

[19] Hausman, Daniel and Michael McPherson (1996) “Economic Analysis and Moral Philosophy ” Cambridge: Cambridge University Press.

[20] Kolm, Serge-Christophe (1998) “The value of freedom ” in : J.-F. Laslier, M. Fleurbaey, N. Gravel and A. Trannoy (eds.) Freedom in Economics: New perspectives in normative analysis. London: Routledge. 17-44.

[21] Lamassoure, Alain 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).

[22] Laruelle A., Valenciano F., (2005) "Assessing success and decisiveness in voting situations", *Social Choice and Welfare* vol. 24(1), pages 171-197.

[23] Laruelle A., Valenciano F., (2008) "Voting and Collective Decision-Making: Bargaining and Power", *Cambridge University Press, Cambridge, New York.*

[24] Laslier, Jean-François (2012) "Why not proportional?" forthcoming in *Mathematical Social Sciences.*

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

[26] Maaser, Nicola and Stefan Napel (2011). "A Note on the Direct Democracy Deficit in Two-tier Voting", *Mathematical Social Sciences*, forthcoming.

[27] Macé, Antonin and Rafael Treibich (2012) "Computing the optimal weights in a utilitarian model of apportionment" *Mathematical Social Sciences*, 63(2): 141-151.

[28] Martínez-Aroza, José and Victoriano Ramírez-González (2008) "Several methods for degressively proportional allotments. A case study" *Mathematical and Computer Modelling*, 48: 1439-1445.

[29] Mongin, Philippe (2012) "The doctrinal paradox, the discursive dilemma, and logical aggregation theory. " *Theory and Decision* 73: 315-355.

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

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

[32] Pukelsheim, Friedrich (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.

[33] Rae, D., (1969) "Desicion Rules and Individual Values in Constitutional Choice" *American Political Science Review* 63, 40-56.

[34] Ramírez-González, Victoriano, 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.

[35] Sen, Amartya (1985) “Well-being, Agency and Freedom,” Journal of Philosophy, 82: 169–221.

[36] Słomczyński, Wojciech 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.

[37] Theil, Henri (1971) “The allocation of power that minimizes tension” *Operations Research* 19: 977-982.

[38] Theil, Henri and Linus Schrage (1977) “The apportionment problem and the European parliament” *European Economic Review* 9: 247-263.

[39] Trannoy, Alain (2010) “Mesure des inégalités et dominance sociale” *Mathématiques et Sciences Humaines* 192: 29-40.

[40] Treaty on European Union: “Articles 1-19” *Official Journal of the European Union* C 83 (30.2.2010) 13-27.