More on Equal Representation in Two-Tier Voting Systems‡

– Very Preliminary Draft –

Nicola Maaser*
and
Stefan Napel*,†

June 2009

Abstract

For the egalitarian reason that each bottom-tier voter should, in principle, have the same indirect influence on top-tier decisions, delegates have voting weights which increase in the size of their constituency in many assemblies. An earlier Monte-Carlo study (Maaser and Napel, *Social Choice & Welfare* 28: 401–420, 2007) demonstrated that weights proportional to the square root of population sizes come close to ensuring equal representation in a unidimensional spatial voting framework given a 50% decision quota. This paper provides an analytic explanation for this finding. It investigates sophisticated weight allocation rules, which use conventional power indices, and shows that even these fail to extend to quotas $q > 50\%$. More critically, if voters are subject to constituency-specific shocks then, for arbitrary $q \geq 50\%$, a linear rule based on the Shapley-Shubik index outperforms square root rules. This raises the important normative question: which kind of inter-constituency heterogeneity shall be acknowledged behind a constitutional 'veil of ignorance'?

Keywords: equal representation, one person one vote, voting systems, voting power, power indices

1 Introduction

The principle of "one person, one vote" is generally considered to be at the heart of modern democratic constitutions. It guarantees that the collective decision only depends on how many votes an alternative gets, not on whose votes these are. A second basic characteristic of today's idea of democracy is the use of political representatives who make decisions on

‡This paper is to be presented at the Voting Power in Practice Workshop at the University of Warwick, 14-16 July 2009, sponsored by The Leverhulme Trust (Grant F/07-004/AJ).

*Dept. of Economics, University of Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany; phone: +49-921552880; e-mail: nicola.maaser@uni-bayreuth.de, stefan.napel.uni-bayreuth.de

†Corresponding author.

behalf of the citizens. The participation of the latter is largely indirect as they only elect their representatives. We concentrate on systems where representatives who are elected in separate districts participate in a governing body at the union level and cast a block vote for their district. Most often, voting districts are – for geographical, ethnic, or historical reasons – not equally sized so that representatives' voting weights have to somehow reflect their constituency's population size. Prominent examples for such two-tier systems are the EU Council of Ministers and the US Electoral College. The question we wish to answer is: Which rule needs to be applied to define districts' weight at the union level in order to ensure fair representation?

The most intuitive solution to the equal representation problem seems to be to allocate weights proportional to population sizes. Yet, the traditional normative claim from the voting power literature is that indirect powers of citizens are equal iff the powers of the representatives as measured by the Penrose-Banzhaf index are proportional to the *square root* of the respective population (see e. g. Felsenthal and Machover 1998). This solution to the problem of ensuring equal representation in two-tier systems was first offered by Penrose (1946) on the basis of a model with binary (yes or no) decisions. His *square root rule* played a prominent role in the public discussion on the EU constitution (see, for example, the open letter by Bilbao et al. 2004). The controversy regained its momentum at the Council of the European Union (the EU Summit) in June 2007 due to Poland's lobbying for a square-root allocation of weights in the EU Council of Ministers. In various studies on this decision-making body the square root rule has been used as a benchmark (e. g. Felsenthal and Machover 2001, 2004; Leech 2002).

Square root rules have also been demonstrated to be optimal under criteria other than the equality of influence. If weights of the delegates, rather than their powers, are made proportional to the square root of their constituency's population size, we get the *second square root rule* (Felsenthal and Machover 1999) which minimizes the mean deviation of the indirect two-tier decision-making rule from a direct democracy simple majority rule. Beisbart and Bovens (2005) and Beisbart and Hartmann (2006) arrive at the square root rule in a welfarist framework starting from the norm that expected utility should be equalized for all countries. Basically the same result is reached by Barberà and Jackson (2006) who study the design of voting rules that maximize the expected utility of voters for their 'fixed-size-of-blocks model'.

Penrose's square root rule critically depends on equiprobable 'yes' and 'no'-decisions by all voters (or at least a 'yes'-probability which is random and distributed independently across voters with mean exactly 0.5). In this *binomial voting model*, the probability of a tie in a constituency with n_j voters is approximately inversely proportional to $\sqrt{n_j}$.[3] Good and Mayer (1975) and Chamberlain and Rothschild (1981) demonstrate that if the 'yes'-probability is slightly lower or higher, or if it exhibits even minor dependence across voters – say, they are influenced by the same newspapers – the tie probability is substantially smaller than for equiprobable yes/no-decisions. Related empirical studies in two-party

[3] As an individual voter is decisive only if the election is otherwise tied, this means that voting power is in approximation inversely proportional to $\sqrt{n_j}$.

elections have in fact failed to confirm the prediction for the average closeness of ballots (see Gelman, Katz, and Tuerlinckx 2002 and Gelman, Katz, and Bafumi 2004). Larger elections are slightly closer – in proportional terms – than small elections, but by very little, perhaps making a 0.9 power rule appropriate. These findings certainly cast doubts on the accuracy of the binomial distribution model which is the source of the square root rule. It is therefore rather disturbing that implementation of the square root rule may result in highly unequal representation (see Felsenthal and Machover 1998, p. 71f) when its assumptions, namely the binomial voting model, are not fulfilled.

Maaser and Napel (2007) investigate equal representation when policy alternatives are non-binary and decisions are made by simple majority rule, and conclude that weight proportional to the square root of population size is close to optimal in that setting. This finding may be interpreted as extending the scope of Penrose's square root rule beyond the narrow limits of binary decision-making. While that result has so far been based on extensive Monte-Carlo simulations, this paper provides an analytic explanation.

Penrose's square root rule is derived in a voting model where voters only differ in the constituency they live in. This implies that the voting behavior of citizens from the same constituency is not more highly correlated than the voting behavior of citizens from different constituencies. Though the model introduced in Maaser and Napel (2007) is quite different from Penrose's, it shares with the latter the premise that all voters are a priori identical. The equiprobability assumption (as implied by the principle of insufficient reason) is usually maintained on the normative grounds that constitutional design should not rest on volatile correlation patterns of preferences.

We conduct a *sensitivity analysis* on the findings in Maaser and Napel (2007), i.e., we vary the assumptions made there, and observe the effect on the result. Specifically, the aim of this paper is to address the following questions:[4]

1. How does a 'simple' voting rule that derives directly from constituency sizes perform compared with more sophisticated rules that use standard power indices as reference points?

2. What is the fair voting rule under supermajority rules at the top tier?

3. How does the fair voting rule react to heterogeneity across constituencies?

2 Model and Results in Maaser and Napel (2007)

Consider a large population N of *voters*. Let $\mathfrak{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_m\}$ be a partition of the population into m *constituencies* $\mathcal{C}_j$ with $n_j = |\mathcal{C}_j| > 0$ members each. Citizens' preferences are single-peaked with *ideal point* λ_i^j (for $i = 1, \ldots, n_j$ and $j = 1, \ldots, m$) in a bounded convex one-dimensional *policy space* $X \subset \mathbb{R}$. Assume for simplicity that all n_j are odd numbers.

[4] An additional benefit of the present study is its greater precision.

For any random policy issue, let $\cdot : n_j$ denote the permutation of voter numbers in constituency $\mathcal{C}_j$ such that

$$\lambda_j^{1:n_j} \leq \ldots \leq \lambda_j^{n_j:n_j}$$

holds. In other words, $k : n_j$ denotes the k-th leftmost voter in $\mathcal{C}_j$ and $\lambda_j^{k:n_j}$ denotes the k-th leftmost ideal point (i. e., $\lambda_j^{k:n_j}$ is the *k-th order statistic* of $\lambda_j^1, \ldots, \lambda_j^{n_j}$).

A policy $x \in X$ is decided on by a *committee of representatives* $\mathcal{R}$ consisting of one representative from each constituency. Without going into details, we assume that the representative of $\mathcal{C}_j$, denoted by j, adopts the ideal point of his constituency's *median voter*, denoted by

$$\lambda_j \equiv \lambda^j_{(n_j+1)/2:n_j}.$$

In theory, elected representatives are fully responsive to their constituency's median voter. Practically, at least two problems arise: First, systematic abstention of certain social groups can drive a substantial wedge between the median voter's and the median citizen's preference, and non-voters go unrepresented. Second, empirical evidence suggests that a representative may take positions that differ significantly from his district's median when voter preferences within that district are sufficiently heterogeneous (Gerber and Lewis 2004).

Let $\lambda_{k:m}$ denote the k-th leftmost ideal point amongst all the representatives (i. e., the k-th order statistic of $\lambda_1, \ldots, \lambda_m$).

In the top-tier assembly or committee of representatives $\mathcal{R}$, each constituency $\mathcal{C}_j$ has *voting weight* $w_j \geq 0$. Any subset $S \subseteq \{1, \ldots, m\}$ of representatives which achieves a combined weight $\sum_{j \in S} w_j$ above $q \equiv 0.5 \sum_{j=1}^m w_j$, i.e., a *simple majority* of total weight, can implement a policy $x \in X$.

Consider the random variable P defined by

$$P \equiv \min\Big\{l \in \{1, \ldots, m\} \colon \sum_{k=1}^{l} w_{k:m} > q\Big\}.$$

Player $P\!:\!m$'s ideal point, $\lambda_{P:m}$, is the unique policy that beats any alternative $x \in X$ in a pairwise majority vote, i. e., constitutes the *core* of the voting game defined by weights and quota. Without detailed equilibrium analysis of any decision procedure that may be applied in $\mathcal{R}$ (see Banks and Duggan 2000 for sophisticated non-cooperative support of policy outcomes inside or close to the core), we assume that the policy agreed by $\mathcal{R}$ is in the core, i. e., it equals the ideal point of the *pivotal representative* $P\!:\!m$.

By a fair representation of voters in a two-tier system we mean that *Each voter in any constituency should have an equal chance to determine the policy implemented by the electoral college.* Or, more formally, there should exist a constant $c > 0$ such that

$$\forall j \in \{1, \ldots, m\} \colon \forall i \in \mathcal{C}_j \colon \; \Pr\big(j = P\!:\!m \,\wedge\, i = (n_j + 1)/2\!:\!n_j\big) \equiv c. \tag{1}$$

Whereas in Maaser and Napel (2007) we assumed that the ideal points of all voters throughout the union are *independently and identically distributed* (i. i. d.), we now allow for different distributions of the ideal points in different constituencies. Yet, we retain the assumption that voters' ideal points *within* each constituency $\mathcal{C}_j$ are i. i. d., which gives each voter $i \in \mathcal{C}_j$ the same probability to be the constituency's median. Hence,

$$\forall j \in \{1, \ldots, m\}\colon \forall i \in \mathcal{C}_j\colon \Pr\big(i = (n_j+1)/2 \colon n_j\big) = \frac{1}{n_j}.$$

Using that the events $\{i = (n_j+1)/2 \colon n_j\}$ and $\{j = P \colon m\}$ are independent, one can thus rewrite the fairness condition (1) as

$$\forall j \in \{1, \ldots, m\}\colon \frac{\Pr(j = P \colon m)}{n_j} \equiv c. \tag{2}$$

So if constituency $\mathcal{C}_j$ is twice as large as constituency $\mathcal{C}_k$, representative j must have twice the chances to be pivotal than representative k in order to equalize individual voters' chances to be pivotal.

If representatives' ideal points $\lambda_1, \ldots, \lambda_m$ were i. i. d., $\Pr(j = P \colon m)$ would simply be the Shapley-Shubik index (SSI) value, $\phi_j(w, q)$, of representative j in voting body $\mathcal{R}$ defined by weight vector $w = (w^1, \ldots, w^m)$ and quota q (see Shapley and Shubik 1954). Equation (2) then implies that a *linear rule* based on the SSI would guarantee equal representation. In other words, w would have to be chosen such that $\phi_j(w, q)$ is directly proportional to population size n_j for all constituencies $j = 1, \ldots, m$. Solving this *inverse problem* sufficiently accurately is a relatively simple task – at least as long as the number of constituencies is 'large' and no 'pathologies' in the population configuration occur.[5]

But under the assumption in Maaser and Napel (2007) that voters' ideal points are i. i. d., representatives' ideal points $\lambda_1, \ldots, \lambda_m$ are independently but (except in the trivial case $n_1 = \ldots = n_m$) *not* identically distributed. Given F_j with density f_j, the median position λ_j in constituency $\mathcal{C}_j$ is asymptotically normally distributed (see e. g. Arnold et al. 1992, p. 223) with mean

$$\mu_j = F_j^{-1}(0.5)$$

and standard deviation

$$\sigma_j = \frac{1}{2\, f_j(F_j^{-1}(0.5))\sqrt{n_j}}. \tag{3}$$

So, the larger a constituency $\mathcal{C}_j$ is, the more concentrated is the distribution of its median voter's ideal point, λ_j, on the median of the underlying ideal point distribution in this constituency.

The measure used to evaluate the performance of different rules for the allocation of voting weights considers cumulative quadratic deviations between the realized and the ideal chances of an individual. Any voter in any constituency $\mathcal{C}_j$ would ideally determine the outcome with the same probability $1/\sum_{k=1}^m n_k$, but vector $\hat{\pi}$ actually gives him or her the probability $\hat{\pi}_j/n_j$ of doing so. Treating all n_j voters in any constituency $\mathcal{C}_j$ equally then amounts to looking at

$$\sum_{j=1}^m n_j \cdot \left(\frac{1}{\sum_{k=1}^m n_k} - \frac{\hat{\pi}_j}{n_j}\right)^2. \tag{4}$$

[5] An example for such a pathology is provided by Lindner and Machover (2004).

Under this criterion, the increase in deviation is higher if the pivot probability of a large constituency is off the mark than if a small constituency is misrepresented to the same degree.

Focusing on the investigation of *power laws*

$$w_j = {n_j}^{\alpha} \tag{5}$$

with $\alpha \in [0, 1]$,[6] Maaser and Napel (2007) find that $\alpha = 0.5$, i.e., *voting weight proportional to the square root of population*, emerges as close to optimal for both real world examples as well as a vide variety of artificial population configurations.

3 Analytic Arguments

For the reasons stated above, it seems unrealistic to aim for a general analytical solution to the equal representation problem (1), or equivalently, (2) for arbitrary finite configurations $(n_1, \ldots, n_m)$. But is there a way of making progress for a particularly clear layout? The following heuristic arguments suggest that this might indeed be possible.

Assume that representatives' ideal points λ_j are normally distributed with mean $\mu_j = 0$ and standard deviation

$$\sigma_j = \frac{\vartheta\sqrt{2\pi}}{2\sqrt{n_j}} > 0,$$

where $\vartheta > 0$ is a constant. Denote the cumulative density function of λ_j by F_{λ_j}, and the density of λ_j by f_{λ_j}. The latter is given by

$$f_{\lambda_j}(x) = \frac{1}{\sigma_j\sqrt{2\pi}}\, e^{-\frac{x^2}{2{\sigma_j}^2}} = \frac{\sqrt{n_j}}{\vartheta\pi}\, e^{-\frac{x^2\, n_j}{\vartheta^2\pi}}\,. \tag{6}$$

Let Ω denote the set of vectors of median ideal points. Finally, to facilitate notation, let $\pi_j \equiv \Pr\,(j = P\!:\!m)$. A representative k with ideal point λ_k is pivotal in the committee of representatives if, for a given realization $\omega \in \Omega$ of median voters' ideal points, the total weight of the representatives who have ideal points to the left of λ_k is greater than or equal to $q - w_k$, but less than q.

Given weights $w_1, \ldots, w_m$, let $\tilde{w}_j(x)$, $x \in X$, be the random variable defined by

$$\tilde{w}_j(x)(\omega) = \begin{cases} w_j & \text{if } \lambda_j(\omega) \leq x \\ 0 & \text{if } \lambda_j(\omega) > x. \end{cases}$$

where $\omega \in \Omega$ refers to a particular ideal point realization. The random variable $\tilde{w}_j(x)$ is the contribution of constituency $\mathcal{C}_j$ to the total weight of constituencies which have ideal

[6] For big m this approximately includes Penrose's square root rule as the special case $\alpha = 0.5$ (see Lindner and Machover 2004 and Chang, Chua, and Machover 2006).

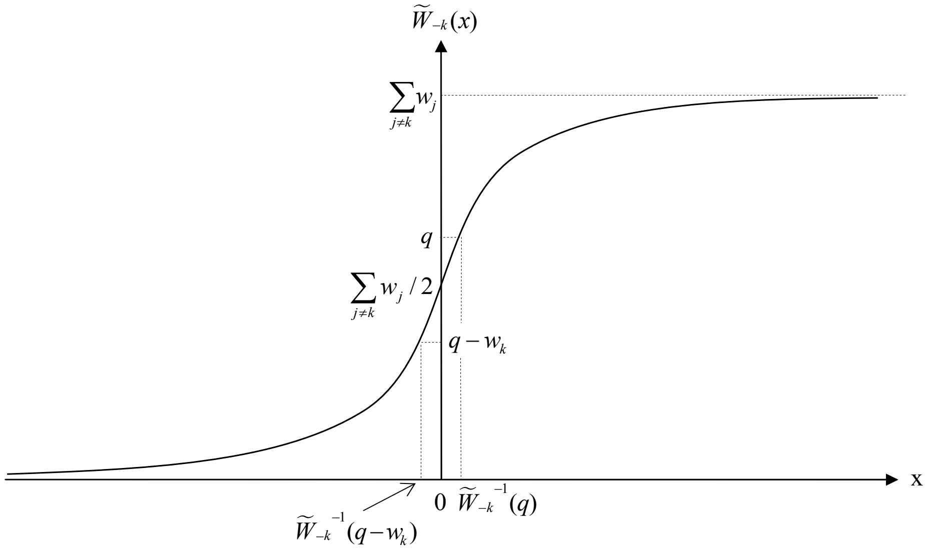

Figure 1: Accumulated weight of constituencies other than k and determination of $\tilde{W}_{-k}^{-1}(q - w_k)$ and $\tilde{W}_{-k}^{-1}(q)$.

points weakly to the left of x. Denote the weight accumulated up to x by constituencies other than any fixed constituency $\mathcal{C}_k$ by

$$\tilde{W}_{-k}(x)(\omega) = \sum_{j \neq k} \tilde{w}_j(x)(\omega).$$

Consider any ideal point realization ω such that $\lambda_k(\omega) = x$. Constituency $\mathcal{C}_k$ is pivotal in the committee of representatives iff

$$\tilde{W}_{-k}(x)(\omega) \leq q < \tilde{W}_{-k}(x)(\omega) + w_k$$

or

$$q - w_k < \tilde{W}_{-k}(x)(\omega) \leq q.$$

The expected value of the probability of this event with respect to the probability density function $f_{\lambda_k}(x)$ yields k's overall power $\Pr(k = P\!:\!m)$,

$$\pi_k = \int_{-\infty}^{\infty} \Pr(q - w_k < \tilde{W}_{-k}(x) \leq q)\, f_{\lambda_k}(x)\, dx. \tag{7}$$

So, π_k is the probability that representative k's median is located between the positions $\tilde{W}_{-k}^{-1}(q - w_k)$ and $\tilde{W}_{-k}^{-1}(q)$ in X at which constituencies $j \neq k$ have accumulated weight $q - w_k$ and q, respectively.[7] This is illustrated in Figures 1 and 2.

[7]Note that $\tilde{W}_{-k}(x)$ is a step function. Thus, $\tilde{W}_{-k}^{-1}(\cdot)$ is the *quasi-inverse* of $\tilde{W}_{-k}(x)$, i.e., $\tilde{W}_{-k}^{-1}(y) = \inf\{x \in X \mid y \leq \tilde{W}_{-k}(x)\}$.

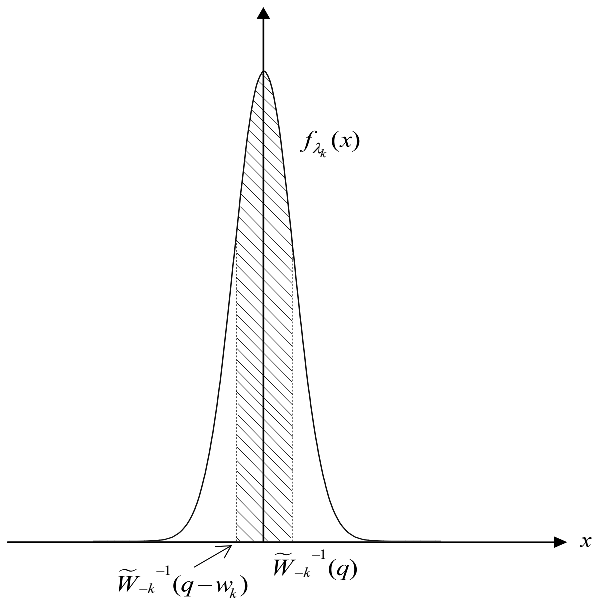

Figure 2: Density of λ_k. The shaded area corresponds to π_k, which is the expectation of the event that λ_k is situated between $\tilde{W}_{-k}^{-1}(q-w_k)$ and $\tilde{W}_{-k}^{-1}(q)$.

As mentioned already, the explicit computation of $\tilde{W}_{-k}(x)$'s distribution, and hence that of π_k, is very involved: for any $\overline{W} \in (q-w_k; q]$, one needs to account for all combinatorial possibilities to reach the aggregate weight $\overline{W}$ without $\mathcal{C}_k$. This would amount to the enumeration of all coalitions S not containing representative k with weight $\sum_{j\in S} w_j = \overline{W}$, followed by the summation of the respective formation probabilities $\prod_{j\in S} F_{\lambda_j}(x) \prod_{j\notin S}(1-F_{\lambda_j}(x))$.

An approximation which ignores these combinatorial complications is of little use for estimating power for any particular weight distribution, but helps in identifying the general behavior of power, as weight and population size is varied.

The key observation is that, for a large number m of constituencies, $\tilde{W}_{-k}(x)$ – as the sum of $m-1$ independent random variables – is approximately normally distributed,[8] with mean

$$\mathbf{E}\tilde{W}_{-k}(x) = \sum_{j\neq k} \mathbf{E}\tilde{w}_j(x) = \sum_{j\neq k} w_j F_{\lambda_j}(x). \tag{8}$$

As m goes to infinity, the variance of the random variable $\tilde{W}_{-k}(x)$ approaches zero. Hence, the error of replacing $\tilde{W}_{-k}(x)$ in (7) by its expected value $\mathbf{E}\tilde{W}_{-k}(x)$ is small when we look

[8] As the random variables $\tilde{w}_j(x)$ are independently, but not identically distributed with finite variance, *Lyapunov's central limit theorem* applies.

at a large number of constituencies. In particular, we can then approximate (7) by

$$\begin{aligned}\pi_k \approx \hat{\pi}_k &= \int_{-\infty}^{\infty} \mathbf{1}_{\{x:\, q-w_k<\mathbf{E}\tilde{W}_{-k}(x)\le q\}}(x)\, f_{\lambda_k}(x)\, dx \\ &= \int_{-\infty}^{\infty} \mathbf{1}_{\{x:\, \mathbf{E}\tilde{W}_{-k}^{-1}(q-w_k)<x\le \mathbf{E}\tilde{W}_{-k}^{-1}(q)\}}(x)\, f_{\lambda_k}(x)\, dx \\ &= \int_{\mathbf{E}\tilde{W}_{-k}^{-1}(q-w_k)}^{\mathbf{E}\tilde{W}_{-k}^{-1}(q)} f_{\lambda_k}(x)\, dx \end{aligned} \tag{9}$$

where $\mathbf{1}_X$ denotes the indicator function of set X.

From the point symmetry of the normal cumulative density function, $F_{\lambda_j}(x) = 1 - F_{\lambda_j}(-x)$, together with (8), it follows that $\mathbf{E}\tilde{W}_{-k}(x)$ is point symmetric in relation to the point $(0; \sum_{j\neq k} w_j/2)$ (cf. Figure 1). Consider the case of simple majority rule, $q = \sum_j w_j/2$. The quota can be rewritten as $q = \sum_{j\neq k} w_j/2 + w_k/2$, whilst $q - w_k = \sum_{j\neq k} w_j/2 - w_k/2$. Thus, if $\mathbf{E}\tilde{W}_{-k}(z) = q$ for some $z \in \mathbb{R}$, then it holds that $\mathbf{E}\tilde{W}_{-k}(-z) = q - w_k$.

Using (6), approximation (9) becomes

$$\hat{\pi}_k = \int_{-z}^{z} f_{\lambda_k}(x)\, dx = \frac{\sqrt{n_k}}{\vartheta\pi} \int_{-z}^{z} e^{-\frac{x^2 n_k}{\vartheta^2\pi}}\, dx = 2\frac{\sqrt{n_k}}{\vartheta\pi} \int_{0}^{z} e^{-\frac{x^2 n_k}{\vartheta^2\pi}}\, dx \tag{10}$$

where z is implicitly defined by $\mathbf{E}\tilde{W}_{-k}(z) = q$. This integral can be written as the Taylor series

$$\begin{aligned}\hat{\pi}_k &= 2\frac{\sqrt{n_k}}{\vartheta\pi} \int_{0}^{z} \left(1 - \frac{\eta x^2}{1!} + \frac{\eta^2 x^4}{2!} - \frac{\eta^3 x^6}{3!} + \dots\right) dx \\ &= 2\frac{\sqrt{n_k}}{\vartheta\pi} \left(z - \frac{\eta z^3}{3\cdot 1!} + \frac{\eta^2 z^5}{5\cdot 2!} - \frac{\eta^3 z^7}{7\cdot 3!} + \dots\right)\end{aligned} \tag{11}$$

with $\eta \equiv \frac{n_k}{\vartheta^2\pi}$.

If constituency k were included in the aggregation, the quota $q = \sum_j w_j/2$ would in expectation be accumulated exactly at $x = 0$. If constituency $\mathcal{C}_k$'s weight w_k is 'small' relative to $\sum_j w_j$, then q will in expectation be accumulated slightly to the right of zero, i.e., z is close to zero. For this case, terms in (11) with degree greater than one have only a second-order effect. So $\hat{\pi}_k$ can rather well be approximated by

$$\hat{\hat{\pi}}_k = \frac{2z\sqrt{n_k}}{\vartheta\pi}. \tag{12}$$

In the neighborhood of $x = 0$, $F_{\lambda_j}(x)$ can be approximated by its Taylor polynomial of degree 1, i. e.

$$\hat{F}_{\lambda_j}(x) = F_{\lambda_j}(0) + x f_{\lambda_j}(0) = \frac{1}{2} + x f_{\lambda_j}(0).$$

Then, solving

$$\sum_{j\neq k} w_j/2 + w_k/2 = q = \mathbf{E}\tilde{W}_{-k}(z) \approx \sum_{j\neq k} w_j \hat{F}_{\lambda_j}(z) = \sum_{j\neq k} w_j/2 + \sum_{j\neq k} w_j z f_{\lambda_j}(0),$$

the location z is obtained approximately as

$$z \approx \frac{w_k}{2\sum_{j\neq k} w_j\, f_{\lambda_j}(0)}.$$

This, together with (12), leads to the conclusion that

$$\pi_k \approx \frac{w_k\sqrt{n_k}}{\vartheta\pi\sum_{j\neq i} w_j\, f_{\lambda_j}(0)}. \tag{13}$$

According to (13), the probability of constituency $\mathcal{C}_k$ to be pivotal at the top tier is approximately proportional to its weight w_k and to the square root of its population n_k. The square root of population, which first showed up in the density (6), describing the distribution of representative j's position, reappears in expression (13). Returning to the equal representation condition (2), it follows from (13) that (2) can be approximately satisfied by choosing weights w_j^* such that

$$w_j^* \propto \sqrt{n_j},$$

where the notation $\propto$ refers to (direct) proportionality between w_j^* and the square root of n_j for all $j = 1, \ldots, m$.

In order to obtain this heuristic result, two major approximations were made: first, the effect of the combinatorial features of a particular weight distribution on power is ignored (leading to formula (9)). Second, the 'lumpiness' of player k's weight which implies that z is actually larger than 0 is not taken into account (leading to (12) and (13)). Nevertheless, expression (13) allows a prediction about the equitable weight allocation for large representative committees and exposes the reason why one would expect a square root rule to eventually emerge from the 'double pivot' model introduced in Maaser and Napel (2007).[9]

While the approximative weight allocation rule $w_j^* \propto \sqrt{n_j}$ may be expected to work well under 'limit conditions', it is of limited use when the number of constituencies is 'small'. The following section for this reason uses Monte-Carlo simulation in order to approximate the probability of any constituency $\mathcal{C}_j$ being pivotal for a given partition of an electorate, or *configuration* $\{\mathcal{C}_1, \ldots, \mathcal{C}_m\}$, and a fixed weight vector $(w_1, \ldots, w_m)$. Based on this, the simulation tries to find weights $(w_1^*, \ldots, w_m^*)$ which approximately satisfy the two equivalent equal representation conditions (1) and (2).

4 Simple and sophisticated square root rules

In Maaser and Napel (2007), the simple square root rule, $w_j \propto \sqrt{n_j}$, has been found to ensure equal representation to an almost optimal extent when the number of constituencies is large. With few constituencies (and representatives), however, it becomes more important

[9]Note that Penrose's square root rule also includes an approximation: in its derivation, *Stirling's formula* is used to approximate the probability that an individual voter is decisive at the lower-tier referendum (or a general two-candidate election).

to have an index that, at least approximately, captures the power distribution generated by voting weights at the top tier. Standard power indices can be ruled out as candidates for the 'theoretically correct' index because they are based on *identical* stochastic behavior of top-tier voters, which is generally inconsistent with identical stochastic behavior of bottom-tier voters. Still, as a second-best solution, sophisticated rules that are based upon the Shapley-Shubik index or the Penrose-Banzhaf index might be expected to do better than the simple rule: the latter ignores all combinatorial aspects of weighted voting, while the former capture them at least for identical top-tier behavior. The latter is not too far off when constituencies have similar sizes.

Implementing such rules requires a solution to the *inverse problem* of finding weights which induce a desired power distribution (see, e. g., Leech 2003, and Leech and Machover 2003). For a finite number n of committee members, the number of different voting rules is also finite, albeit increasing very quickly in n. Therefore, the set of reachable power vectors is discrete, as illustrated in Figure 3. The problem of enumerating all simple games with n players could be solved by determining all *antichains* on 2^N.[10] This (unsolved) problem is known as *Dedekind's problem*, and the corresponding numbers are called Dedekind numbers. The number of games in the important subclass of non-dictatorial, weighted majority games with a quota of half the total weight is considerably smaller: for example, counting permutations, there are four such games with three players,[11] whereas the Dedekind number (excluding the empty antichain which contains no subsets and the antichain consisting of only the empty set) is 18.

It is worthwhile to compare the performance of the simple square root rule to that of sophisticated rules in the double median setting introduced in Maaser and Napel (2007). For the comparison, we use 30 randomly generated configurations of 15 constituencies each.[12] Experience suggests that at this value the distribution of power is not entirely governed any more by the combinatorial particularities of the configuration at hand, but asymptotic properties only begin to operate (see Chang, Chua, and Machover 2006). For larger numbers of constituencies, it becomes increasingly difficult to make meaningful comparisons between weight-based and index-based rules, as the power ratio (measured by the Penrose-Banzhaf or the Shapley-Shubik index) between any two representatives typically approaches the ratio of their voting weights. This convergence is asserted by Penrose's (1952) Limit Theorem, which has been proved to hold under certain conditions (Lindner and Machover 2004), and seems to apply whenever the weight distribution is not too skewed.[13]

[10]A subset of a partially ordered set (or poset) $(P, <_P)$ – where P is a set, and $<_P$ is a partial order relation – is an antichain if any two elements of the subset are incomparable under $<_P$. Applied to simple voting games, the power set 2^N is partially ordered in respect to the inclusion $\subseteq$, and each set of minimum winning coalitions, characterizing a game, corresponds to an antichain.

[11]The minimum integer weight representations of these four games are $(3; 2, 1, 1)$, $(3; 1, 2, 1)$, $(3; 1, 1, 2)$, and $(2; 1, 1, 1)$. In Figure 3, these games correspond to the four points in the interior of the simplex.

[12]In 15 of the configurations, population sizes were drawn from a uniform distribution, and in the other 15 from a Pareto distribution with $\kappa = 1.0$.

[13]For special classes of weighted voting games, Lindner and Machover (2004) prove Penrose's (1952) Limit Theorem with respect to the Penrose-Banzhaf index for $q = 0.5$ and with respect to the Shapley-

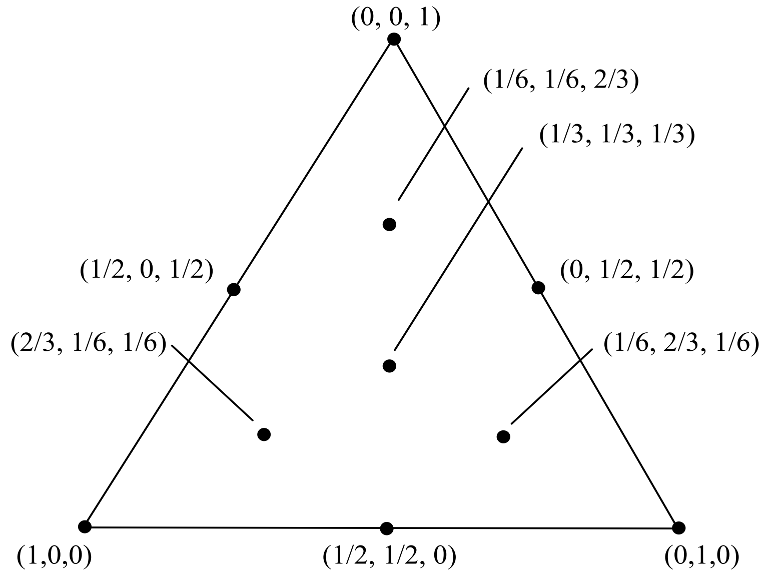

Figure 3: Illustration of the nature of the inverse problem. Numbers are Shapley-Shubik indices for all proper three-player weighted majority games.

Let w_β and w_ϕ denote the weight vectors that are solutions to the inverse problems *"choose weights such that*

(I) $\beta_j(w, q) \propto \sqrt{n_j}$, *and*

(II) $\phi_j(w, q) \propto \sqrt{n_j}$, *for each constituency j"*,

where $\beta(\cdot)$ and $\phi(\cdot)$ refer to the Penrose-Banzhaf measure and the Shapley-Shubik index, respectively. The comparison involves three different weight allocations: simple square root weights, w_β, and w_ϕ. A reference point to evaluate the capacity of these rules to achieve equal representation is provided by the 'best egalitarian weights', as resulting from an unconstrained search for the minimizer of the objective function (4). These, as well as the inverse weights w_β and w_ϕ, are obtained numerically by the *Nelder-Mead simplex method* (see, for example, Avriel 1976, Ch. 9).[14] The deviation from ideal probabilities that is associated with the best unconstrained weights can be considered as inevitable. Owing to the discrete nature of the set of possible power allocations, the discrepancy can, in general, not be eliminated completely.

Shubik index for $q \in (0, 1)$. Their conjecture that the Theorem holds 'almost always' under rather general conditions is corroborated in a simulation study by Chang, Chua, and Machover (2006).

[14]The Nelder-Mead algorithm does not rely on numerical or analytic gradients, which makes it particularly suitable to non-linear optimization problems like the present. In each step of the search, the probabilities $\pi_j \equiv \Pr(j = P\!:\!m)$ of representative j being pivotal in the top-tier committee are approximated by their empirical average over 10 million iterations. A *MATLAB* computer program is used for the computations. The source code is available upon request.

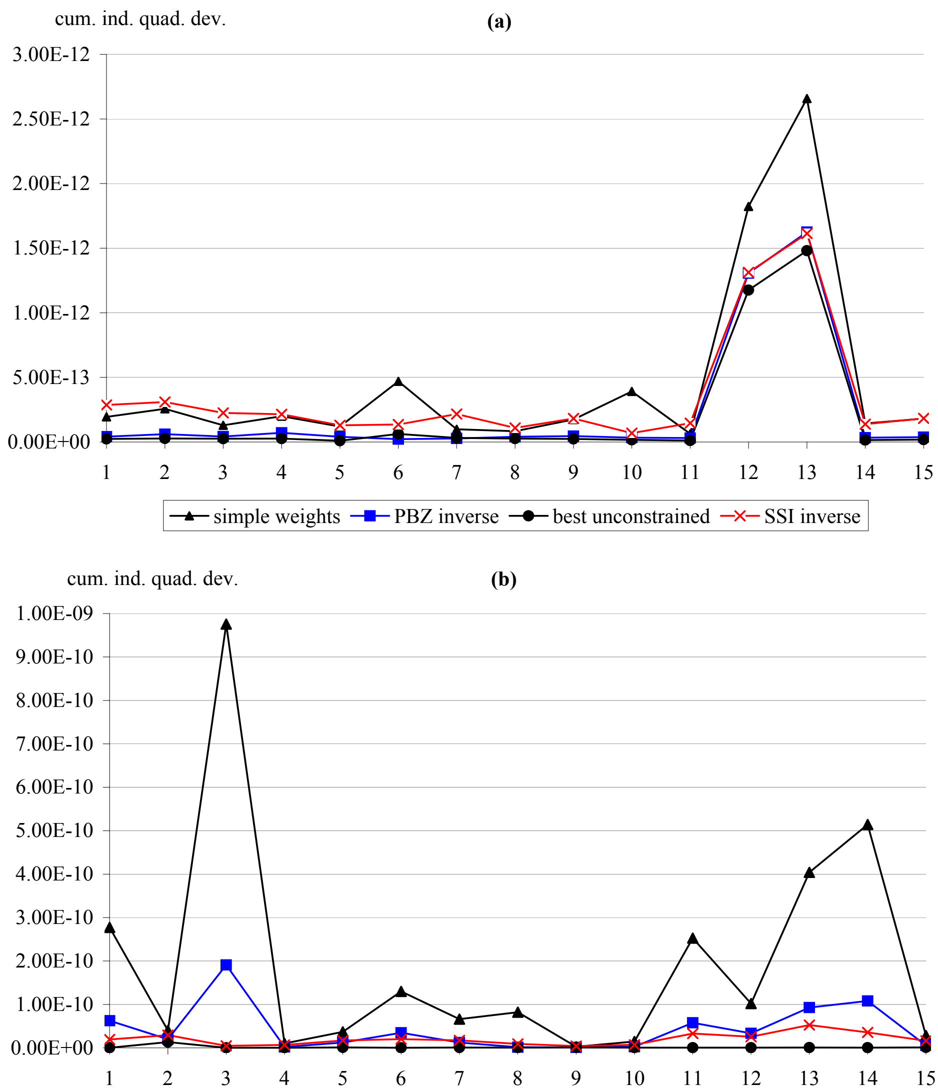

Figure 4: Cumulative individual quadratic deviation under simple weights, w_β, w_ϕ, and best unconstrained weights. Panel (a): 15 configurations with uniformly distributed constituency sizes; Panel (b): 15 configurations with Pareto distributed constituency sizes

Figure 4 already suggests systematic differences in the performance of cumulative deviation (4) under the four sets of weights. The graphic impression is corroborated by a comparison (including all 30 configurations) of cumulative individual quadratic deviations for (1a) w_β, and (1b) w_ϕ, versus simple square root weights, (2a) w_β, and (2b) w_ϕ, versus best unconstrained weights, and (3) w_β versus w_ϕ, using the *Wilcoxon signed rank test* (see, e. g., Hollander and Wolfe 1999, Ch. 3). In tests (1a) and (1b), the null hypothesis that the median difference between pairs of observations is zero could be rejected at the 99% significance level, indicating that both the inverse weights w_β and w_ϕ perform significantly better than simple square root weights.[15] Similarly, the null hypothesis in tests (2a) and (2b) was rejected at the 99% significance level, which suggests that both w_β and w_ϕ are none the less not the correct or first-best weights in the double median setting. In test (3), the null hypothesis could not be rejected, that is, no significant difference between w_β and w_ϕ was detected.

In order to find out whether these differences fade away for larger numbers of constituencies, an additional test including 12 configurations with 30 constituencies each is conducted.[16] For these, cumulative deviations (4) under weights w_ϕ are first compared with those under simple square root weights. The null hypothesis that the median difference between pairs of observations is zero could *not* be rejected at the 95% significance level (it can be rejected at the 90% level). Second, w_ϕ is checked against the best power law weights $w_j = n_j^{\alpha^*}$ with the result that no significant difference in deviations for these two weight allocation rules could be established.

5 Supermajorities and Representation

In view of the important role that supermajority rules play in theory, and their widespread use in real-world decision-making, it is worthwhile to investigate the 'double pivot' model (Maaser and Napel 2007) with respect to the effect of using a quota $q \gg 0.5$ in the top-tier assembly. Of course, even approximately equal representation is impossible under unanimity rule (keeping the bottom-tier role of the median). For $0.5 \ll q \ll 1$, optimal assignments can be expected to give large constituencies greater weight than implied by $\alpha = 0.5$.

In contrast to simple majority rule, the voting game under supermajority rule is not decisive. This means that possibly no policy $x \in X$ exists which defeats all alternatives $x' \neq x$ in a pairwise comparison. The probability that the outcome of collective decision-making is merely a confirmation of the status quo is a measure of the *institutional inertia* created by the decision threshold. The following analysis, however, concentrates on creative power rather than representatives' abilities to preserve the status quo.

[15] Generally, $\alpha = 0.5$ is not exactly the best exponent among all power laws. Obviously, the best power law weights $w_j = n_j^{\alpha^*}$ for a given configuration result in a lower deviation from egalitarian representation than simple square root weights, but they turn out to perform still worse than w_β and w_ϕ.

[16] The 12 configurations consist of 3×4 configurations with population sizes drawn from a uniform, a normal, and a Pareto ($\kappa = 1.0$) distribution, respectively.

To this end, the status-quo, Q, is fixed to a point equal to or left of the leftmost representative's ideal point, which implies that it will always be displaced in favor of some policy to its right by a winning coalition. This assumes that committee members agree about the direction of policy change. Suppose, for illustrative purposes, that $X = [0, 1]$ and $Q = 0$. Moreover, let a continuum of representatives have equal weights and their policy positions be distributed uniformly on X. Then, for a given value of $q \in (0.5, 1)$, all policies $x \in (0,\, 2(1-q))$ are preferred to the status quo by a majority of at least q. However, any policy $x < 1 - q$ could still be improved upon by a share of representatives greater than q. A continuous process of 'displacement' of the status quo in the top-tier committee can be expected to come to a halt at $x = 1-q$. A further movement to the right will be blocked by at least the representative whose ideal point is equal to $1 - q$, and who is strictly necessary to form a winning coalition. For example, the policy outcome under a quota of 0.75 would be the first quartile point of the distribution of representatives' ideal points. Under weighted voting, and for discrete representatives' ideal points, the above reasoning suggests that the policy adopted in the committee of representatives coincides with the ideal point $\lambda_{P:m}$ of representative $P : m$ who is pivotal 'from the right'. The random variable P is defined by

$$P \equiv \min\Big\{ l \in \{1, \ldots, m\} \colon \sum_{k=1}^{l+1} w_{k:m} > (1-q) \sum_{j=1}^{m} w_j \Big\}.$$

In principle, it seems feasible, for 'limit' situations, to extend the analytical arguments put forward in Section 3 concerning the probability of top-tier pivotality under simple majority rule to the case of supermajorities. This being beyond the scope of the present work, we resort to Monte-Carlo simulation to evaluate rules of the form $w_j = n_j^{\alpha}$ and search for the optimal α given alternative values of the quota. Again, the extent to which the considered rule falls short of the egalitarian norm (1) or (2) will be measured by cumulative quadratic deviation at the individual level as given by (4). First, randomly generated configurations will be investigated, then, we briefly look at the EU Council of Ministers.

Randomly generated configurations

Under different assumptions about the distribution of constituency sizes, each of the Tables 1, 2, and 3 reports optimal values of α for four configurations with $m = 30$ constituencies. As mentioned above the difference between simple and sophisticated rules becomes insignificant for this number of players. We therefore concentrate on findings regarding optimal rules of the type $w_j = n_j^{\alpha}$. The values of α run from 0 to 1 in 0.01-intervals, and probabilities π_j were estimated by simulations with 10 mio. iterations. Values in parentheses are the deviations (4) associated with the optimal α.

Three observations apply irrespectively of the distributional assumption. First, α^* increases in the quota. This is due to the fact that the median voter of large constituencies is more central, which lowers the chances of the constituencies to be pivotal when a considerable supermajoritarian rule is used. To compensate this effect, the weight of populous countries has to rise. Second, the deviation from ideal egalitarian probabilities generally also increases. From $q = 55\%$ to $q = 80\%$, the quality of representation deteriorates by up to a factor of 1000. This decline indicates that any power law either gives not enough or too much pivot probability to large constituencies. Third, while one might have expected cumulative individual quadratic deviations to be lowest under simple majority, they reach their minimum at a quota of 55% (among all quotas considered here).

Table 1 relates results for uniformly distributed constituency sizes $n_1, \ldots, n_{30}$. Populations in configurations (I) and (II) come from a uniform distribution over $[0, 10^8]$, and those in (III) and (IV) come from a uniform distribution over $[3 \cdot 10^6, 10^7]$. It will be readily noticed that the deviations in columns (I) and (II) are smaller than deviations in (III) and (IV) except for the highest quotas where no systematic difference is apparent. Moreover, the optimal α exhibits greater stability from configuration (I) to (II) than between (III) and (IV). As the variance of $\mathbf{U}(0, 10^8)$ is by a factor of 200 greater than that of $\mathbf{U}(3 \cdot 10^6, 10^7)$, the data suggest a positive relationship between the variance of the population numbers and the accuracy of the equal representation rule. These findings are corroborated by the data for normally distributed populations contained in Table 2, where again the two left columns pertain to more variable population configurations than the two right columns. The likely explanation for these patterns is that a great amount of variance in population numbers translates *ceteris paribus* into a great variety of weights in the power law allocation $w_j = n_j^{\alpha}$. This implies, for a given quota, that more distinct winning coalitions exist, enabling a closer match between achievable and ideal probability vectors. It is worth noting that the variances of $\mathbf{U}(3 \cdot 10^6, 10^7)$ and the normal distribution $\mathbf{N}(10^7, 2 \times 10^6)$ are roughly the same, but a comparison of columns (III) and (IV) in Table 1 with corresponding columns in Table 2 reveals that estimated probabilities are, in most cases, closer to their ideal values with uniformly distributed constituency sizes. Thus, if one considers configurations of the same distribution type, the more variable distribution can be expected to allow more egalitarian representation, but across different types, variance is a less reliable indicator. Under a normal distribution, many constituencies are of similar size, and the minor differences between them cannot easily be reflected adequately by pivot probabilities. Then, the precise value of α^* and the quality of representation then may depend heavily on the particular constituency configuration at hand, as is the case with configuration (III) in Table 2.

Table 3 shows results for population sizes drawn from a Pareto distribution $\mathbf{P}(\kappa, \underline{x})$. The parameter $\kappa > 0$ determines the shape or skewness of the distribution, and $\underline{x} > 0$ is the minimum possible value.[17] Here, only a single or very few large constituencies exist, which are particularly disadvantaged by their central position when supermajority rules apply. A high value of α would give them a power monopoly, but a moderate α gives them insufficient pivot probabilities. This logic drives the rather low values of α under simple majority rule as well as the comparatively high values for the most demanding quotas.

[17]It is not possible to compare the columns in Table 3 with respect to the variance of constituency sizes because the variance of $\mathbf{P}(\kappa, \underline{x})$ is infinity for $\kappa \leq 2$.

q	Distribution of constituency sizes			
	(I) $\mathbf{U}(0, 10^8)$	**(II)** $\mathbf{U}(0, 10^8)$	**(III)** $\mathbf{U}(3 \cdot 10^6, 10^7)$	**(IV)** $\mathbf{U}(3 \cdot 10^6, 10^7)$
50%	**0.49**	**0.49**	**0.46**	**0.42**
	(4.85×10^{-15})	(4.28×10^{-15})	(3.46×10^{-14})	(9.93×10^{-13})
55%	**0.50**	**0.50**	**0.52**	**0.54**
	(2.05×10^{-15})	(1.58×10^{-15})	(2.33×10^{-14})	(3.23×10^{-14})
60%	**0.52**	**0.52**	**0.50**	**0.46**
	(9.00×10^{-15})	(9.52×10^{-15})	(4.84×10^{-14})	(4.41×10^{-13})
65%	**0.56**	**0.56**	**0.58**	**0.60**
	(5.40×10^{-14})	(3.48×10^{-14})	(1.36×10^{-13})	(4.12×10^{-14})
70%	**0.62**	**0.62**	**0.62**	**0.58**
	(2.16×10^{-13})	(1.50×10^{-13})	(2.60×10^{-13})	(8.56×10^{-14})
75%	**0.70**	**0.70**	**0.70**	**0.72**
	(5.79×10^{-13})	(4.10×10^{-13})	(6.95×10^{-13})	(2.52×10^{-13})
80%	**0.80**	**0.80**	**0.82**	**0.82**
	(1.32×10^{-12})	(9.36×10^{-13})	(1.72×10^{-12})	(5.68×10^{-13})

Table 1: Optimal value of α for constituency sizes from Uniform distributions $\mathbf{U}(a, b)$ (cumulative individual quadratic deviations from ideal probabilities in parentheses)

q	Distribution of constituency sizes			
	(I) $\mathbf{N}(10^7, 4 \times 10^6)$	**(II)** $\mathbf{N}(10^7, 4 \times 10^6)$	**(III)** $\mathbf{N}(10^7, 2 \times 10^6)$	**(IV)** $\mathbf{N}(10^7, 2 \times 10^6)$
50%	**0.50**	**0.50**	**0.40**	**0.50**
	(2.38×10^{-14})	(9.33×10^{-14})	(8.78×10^{-12})	(1.20×10^{-11})
55%	**0.50**	**0.50**	**0.64**	**0.65**
	(1.73×10^{-14})	(8.27×10^{-14})	(3.47×10^{-14})	(1.87×10^{-13})
60%	**0.52**	**0.52**	**0.50**	**0.50**
	(3.75×10^{-14})	(7.64×10^{-14})	(7.72×10^{-12})	(1.08×10^{-11})
65%	**0.55**	**0.55**	**0.68**	**0.72**
	(1.29×10^{-13})	(7.10×10^{-14})	(4.95×10^{-14})	(7.00×10^{-14})
70%	**0.60**	**0.62**	**0.50**	**0.56**
	(4.47×10^{-13})	(3.40×10^{-13})	(4.37×10^{-12})	(7.00×10^{-12})
75%	**0.68**	**0.70**	**0.80**	**0.84**
	(1.12×10^{-12})	(5.00×10^{-13})	(1.57×10^{-13})	(1.72×10^{-13})
80%	**0.80**	**0.80**	**0.66**	**0.72**
	(2.64×10^{-12})	(1.78×10^{-12})	(7.87×10^{-13})	(2.18×10^{-12})

Table 2: Optimal value of α for constituency sizes from Normal distributions $\mathbf{N}(\mu, \sigma)$ (cumulative individual quadratic deviations from ideal probabilities in parentheses)

q	Distribution of constituency sizes			
	(I) $\mathbf{P}(1.0, 500000)$	**(II)** $\mathbf{P}(1.0, 500000)$	**(III)** $\mathbf{P}(1.8, 500000)$	**(IV)** $\mathbf{P}(1.8, 500000)$
50%	**0.48**	**0.46**	**0.48**	**0.46**
	(1.96×10^{-12})	(7.46×10^{-12})	(3.37×10^{-13})	(1.86×10^{-11})
55%	**0.50**	**0.50**	**0.50**	**0.48**
	(2.59×10^{-13})	(2.44×10^{-12})	(4.34×10^{-13})	(2.55×10^{-12})
60%	**0.56**	**0.56**	**0.52**	**0.52**
	(1.44×10^{-11})	(2.94×10^{-11})	(3.99×10^{-13})	(4.59×10^{-11})
65%	**0.66**	**0.68**	**0.56**	**0.56**
	(1.22×10^{-10})	(2.56×10^{-10})	(1.95×10^{-12})	(2.67×10^{-10})
70%	**0.78**	**0.80**	**0.60**	**0.62**
	(5.22×10^{-10})	(9.72×10^{-10})	(1.50×10^{-11})	(5.73×10^{-10})
75%	**0.80**	**0.80**	**0.70**	**0.72**
	(1.62×10^{-9})	(2.74×10^{-9})	(7.69×10^{-11})	(7.71×10^{-10})
80%	**0.90**	**0.90**	**0.84**	**0.84**
	(2.99×10^{-9})	(4.31×10^{-9})	(2.30×10^{-10})	(8.53×10^{-10})

Table 3: Optimal value of α for constituency sizes from Pareto distributions $\mathbf{P}(\kappa, \underline{x})$ (cumulative individual quadratic deviations from ideal probabilities in parentheses)

EU Council of Ministers

The EU Council of Ministers decides the largest part of issues by qualified majority voting. A proposal is adopted if, first, 255 out of 345 votes (73.9%) are cast in its favor. The number of votes allocated to each member state roughly reflect the square root of population size. Additionally, the majority weight supporting a proposal must represent a simple majority of member states. Finally, any member state may ask for confirmation that the approving votes represent at least 62% of the EU's total population. The latter two requirements are, however, insignificant as they are in the great majority of cases fulfilled whenever the qualified majority is met (see Felsenthal and Machover 2001). With regard to EU decision-making the assumption of a status quo fixed to the left of the leftmost representative's ideal point is particularly interesting, because often not the direction of new legislation, but only the extent of change is subject to the process of legislative bargaining.

Figure 5 shows, for EU27 population data, the effect of a quota $q > 0.5$ on representation. The respective best value of $\alpha \in \{0, 0.02, \ldots, 0.98, 1\}$ is represented by the solid graph which is measured on the left vertical axis. The figure suggests that the optimal α is approximately a quadratic function of q. The right vertical axis measures the corresponding cumulative individual quadratic deviation. As it has its zero point in the upper right corner, the dashed graph can be interpreted as indicating the closeness between ideal egalitarian probabilities and the (estimated) pivot probabilities under the optimal α-rule. The drop in closeness (or rise in deviation) means that representation becomes increasingly

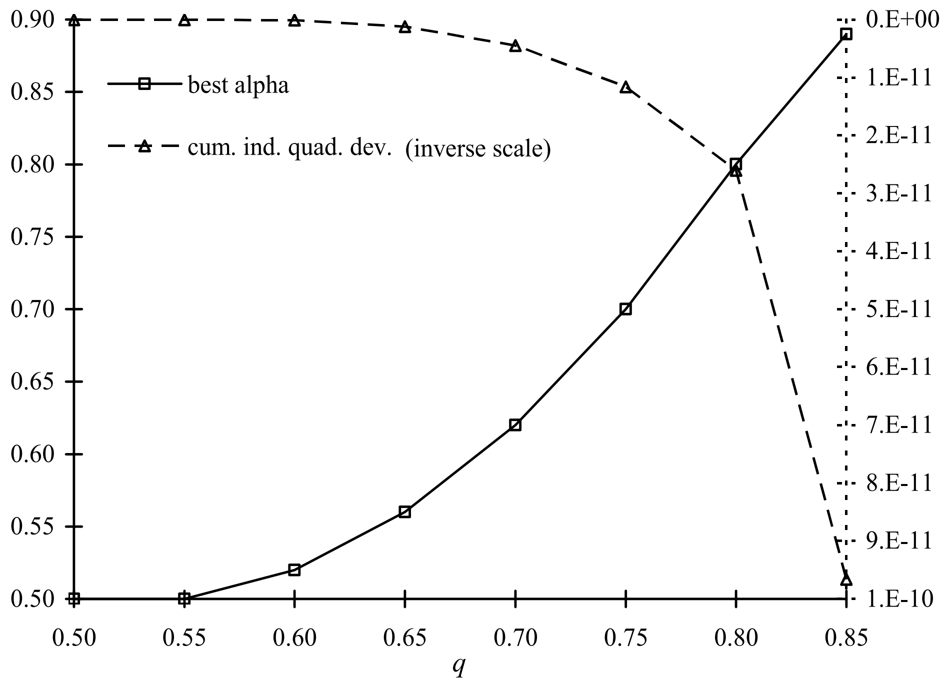

Figure 5: Effect of quota variation for EU27

unequal as the quota is increased. The following numbers may help to get an idea of the deterioration of fit: The ideal probabilities of Malta and Germany are 0.08% and 16.62%, respectively. Already for $q = 0.7$, however, their (estimated) pivot probabilities under the best weight assignment rule $w_j = n_j^{0.62}$ are 0.06% and 15.45%, which is 25% and 7% short of the ideal values.

6 Heterogeneity across constituencies

In contrast to the model considered so far, with i. i. d. ideal points of all individual voters, we now explore the idea that voters within a country are somewhat similar. The attitudes of citizens from one constituency could be systematically connected as the result of a voting with one's feet or sorting process à la Tiebout (1956). In Alesina and Spolaore (2003), preference homogeneity within a country is assumed to develop over time due to geographical proximity and national policies fostering cultural uniformity. The fact that citizens of one country usually share historical experience, traditions, language, communication etc. can be expected to induce some kind of common set of values or 'common belief' in them.[18] The existence of diverse 'common belief' systems is just the reason why

[18] A 'common belief' is also represented by Straffin's (1977) *homogeneity assumption* under which the probability of a voter 'affecting the outcome' coincides with the Shapley-Shubik index (Shapley and Shubik 1954).

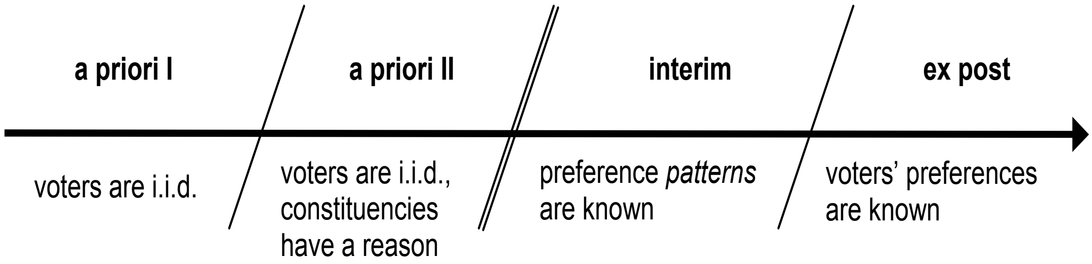

Figure 6: Increasing degree of information usage in the assessment of decision-making situations

countries differ in their population sizes and why they can hardly be redistricted so as to produce equal-size constituencies. Under the premise of Penrose's square root rule that the preferences of the individuals within one country are unconnected, there is no justification apart from the tough practical realization not to regroup citizens into purely administrative districts with equal numbers of voters. When all voters are i. i. d., what is the rationale for a committee of representatives instead of a single president-like decision-maker? We therefore argue that the mere assumption that there does exist some relationship between voter preferences within states, or more precisely, that preferences within a constituency exhibit a higher degree of correlation than across constituencies, is still behind the veil of ignorance and thus relevant in constitutional design. In Figure 6, which illustrates different degrees of using information in the assessment of voting situations, this assumption is referred to as 'a priori II'. The analysis based on the 'a priori II' assumption is to be distinguished from studies that model, particularly with regard to the EU, similarities or dissimilarities between countries based on economic or social dimensions (Widgrén 1995), size or geographical position (Beisbart and Hartmann 2006), and that possibly contribute to understand actual or *a posteriori* power.

We now generalize our earlier model by introducing constituency-specific distributions of individual ideal points. Given a policy issue, the ideal points λ_i^j of voters in constituency $\mathcal{C}_j$ come from an arbitrary identical distribution F_j with density f_j and *distribution median* $\tilde{\lambda}_j$.[19] It is assumed that, rather than being identical, the $\tilde{\lambda}_j$ are random variables with distribution $H_j = H$ for all $j = 1, \ldots, m$. The expected value of $\tilde{\lambda}_j$ is assumed to be zero, and the standard deviation is given by $\sigma_H > 0$, reflecting the degree of heterogeneity across constituencies. Let h denote the density function of H. Maaser and Napel (2007) dealt with the special case without heterogeneity across constituencies, i.e.,, $\sigma_H = 0$. Generally, the distribution F_j is specific to constituency $\mathcal{C}_j$, thus expressing the 'common belief' of that constituency. The ideal points of voters from different constituencies are then independent. In expectation, voter ideal points still have the same distribution, but group membership now makes a difference.

As we noted before, in the case of i. i. d. voters' ideal points the representative of a

[19] If the distributions of individual ideal points are symmetric, $\tilde{\lambda}_j$ could also refer to the mean of the distribution f_j.

larger constituency is on average more central in the electoral college, and given simple majority rule, more likely to be pivotal in it for a given weight allocation. In the light of the standard deviation of the population median λ_j as given by (3), it is plausible that slight differences in the countries' ideal point distributions suffice to make representatives' ideal points virtually identically distributed. The extent of the necessary perturbation depends on the population sizes involved. For example, according to (3), the largest standard deviation of the median ideal point in the EU27 (belonging to Malta as the smallest member state) is $\sigma_{\max} = 7.8 \times 10^{-4}$. Any amount of heterogeneity greater than that, say $\sigma_H = 0.001$, thus practically removes the greater centrality else implied by a larger population.

Let us make this intuition more precise. Given the distribution F_j of individual voters in constituency $\mathcal{C}_j$, the representative's (or median voter) ideal point λ_j is asymptotically normally distributed with mean $\mu_j = F_j^{-1}(0.5) = \tilde{\lambda}_j$, and standard deviation given by (3). For a specific realization of $\tilde{\lambda}_j$, let f_{λ_j} be the density of λ_j. This density is a shifted version of $f^0_{\lambda_j}$, defined by $f^0_{\lambda_j}(x) := f_{\lambda_j}(x \,|\, \tilde{\lambda}_j \equiv 0)$. In particular, it holds that $f_{\lambda_j}(x+\bar{\triangle}) = f^0_{\lambda_j}(x)$, where $\bar{\triangle}$ is a realization of a random variable $\triangle$ with distribution H. Then, the density $\tilde{f}_{\lambda_j}$ of representative j's ideal point λ_j under heterogeneity is given by

$$\tilde{f}_{\lambda_j}(x) = \int_{-\infty}^{+\infty} f_{\lambda_j}(x+\triangle)h(\triangle)d\triangle. \tag{14}$$

Consider first the case that $h(\triangle)$ is a uniform density on the interval $[-a, +a]$, $a > 0$. Then, (14) becomes

$$\tilde{f}_{\lambda_j}(x) = \frac{1}{2a}\int_{-a}^{+a} f_{\lambda_j}(x+\triangle)d\triangle,$$

so

$$\tilde{f}_{\lambda_j}(x) = \frac{1}{2a}\left[F_{\lambda_j}(x+a) - F_{\lambda_j}(x-a)\right].$$

As the standard deviation of F_{λ_j} is small for a constituency $\mathcal{C}_j$ with 'large' population, we have $F_{\lambda_j}(x+a) \approx 1$ and $F_{\lambda_j}(x-a) \approx 0$ if $a \gg 0$, and therefore $\tilde{f}_{\lambda_j}(x) \approx 1/(2a)$ in the 'center' of the interval $[-a, +a]$, irrespective of which constituency $\mathcal{C}_j$ one considers. For x 'close' to the boundaries, $\tilde{f}_{\lambda_j}(x)$ depends on the constituency-specific F_{λ_j}, and thus differs across constituencies. Figure 7 illustrates the above reasoning.

If, in equation (14), $h(\triangle)$ is a normal distribution with zero mean and standard deviation σ_H, then one gets

$$\begin{aligned}
\tilde{f}_{\lambda_j}(x) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_j}\exp\left(-\frac{1}{2}\frac{(x+\triangle)^2}{\sigma_j^2}\right)\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_H}\exp\left(-\frac{1}{2}\frac{\triangle^2}{\sigma_H^2}\right)d\triangle \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\sigma_j^2+\sigma_H^2}}\exp\left(-\frac{1}{2}\frac{x^2}{\sigma_j^2+\sigma_H^2}\right)\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\frac{\sqrt{\sigma_j^2+\sigma_H^2}}{\sigma_j\sigma_H}\exp\left(-\frac{1}{2}\frac{(\triangle+\frac{x\sigma_H^2}{\sigma_j^2+\sigma_H^2})^2}{\frac{\sigma_j^2\sigma_H^2}{\sigma_j^2+\sigma_H^2}}\right)d\triangle \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\sigma_j^2+\sigma_H^2}}\exp\left(-\frac{1}{2}\frac{x^2}{\sigma_j^2+\sigma_H^2}\right).
\end{aligned}$$

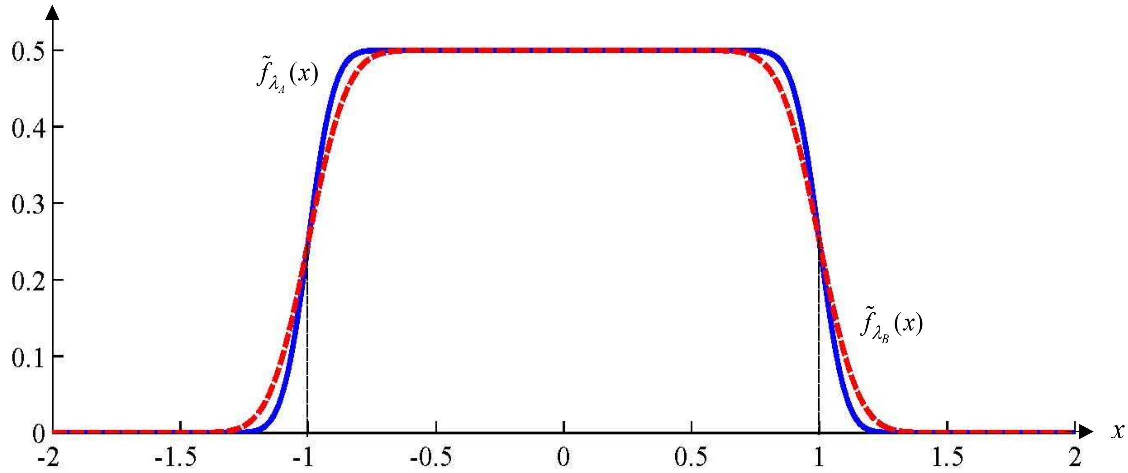

Figure 7: Densities $\tilde{f}_{\lambda_j}(x)$ of λ_j, $j = A, B$. The underlying f_{λ_j} are normal distributions with standard deviations $\sigma_A = 0.08$ (blue graph) and $\sigma_B = 0.12$ (red dashed graph). The heterogeneity function $h(x)$ is a uniform density over $[-1, +1]$.

This establishes that ideal point λ_j of the representative from constituency $\mathcal{C}_j$ is normally distributed with mean zero and standard deviation $(\sigma_j^2 + \sigma_H^2)^{1/2}$. For two constituencies $\mathcal{C}_j$ and $\mathcal{C}_k$ with large but different populations sizes, it holds that

$$\lambda_j \sim \mathbf{N}(0, \sqrt{\sigma_j^2 + \sigma_H^2}) \approx \mathbf{N}(0, \sqrt{\sigma_k^2 + \sigma_H^2}) \sim \lambda_k$$

under the condition that σ_j and σ_k are small compared to σ_H.

Figure 5 shows, for two constituencies A and B of different size, sample median distributions for seven realizations of $\tilde{\lambda}_j$ $(j = A, B)$, respectively. Due to A's larger population size, the f_{λ_A} are much more concentrated around the realizations of $\tilde{\lambda}_A$ than is the case for the f_{λ_B}. Ex ante, however, considering the random variables $\tilde{\lambda}_j$ rather than realizations of them, the density functions $\tilde{f}_{\lambda_A}$ and $\tilde{f}_{\lambda_B}$ practically coincide.

Generally, note that expression (14) is the distribution of the difference $Y_1 - Y_2$ between two random variables Y_1, Y_2 with $Y_1 \sim H$ and $Y_2 \sim F_{\lambda_j}$. If the variance of Y_2 is small compared to that of Y_1 (as is the case if F_{λ_j} is the distribution of the median of a large population), then the random variables Y_2 are almost constant. Hence, the variance of $Y_1 - Y_2$ is practically determined by the variance of Y_1.

The above arguments demonstrate that representatives' ideal points are virtually i. i. d. under the assumption of some heterogeneity between constituencies. Moreover, all $m!$ orderings of representatives are equiprobable. The chances π_j of any representative j to be the pivot at the top tier are then captured by the Shapley-Shubik index $\phi_j(w, q)$. Hence a simple rule ensuring equal representation emerges:

Shapley-Shubik linear rule (Sh-LR): With any amount of heterogeneity $\sigma_H \gg \max_j\{\sigma_j\}$ and for given decision quota q, the egalitarian weights satisfying (2) are defined implicitly

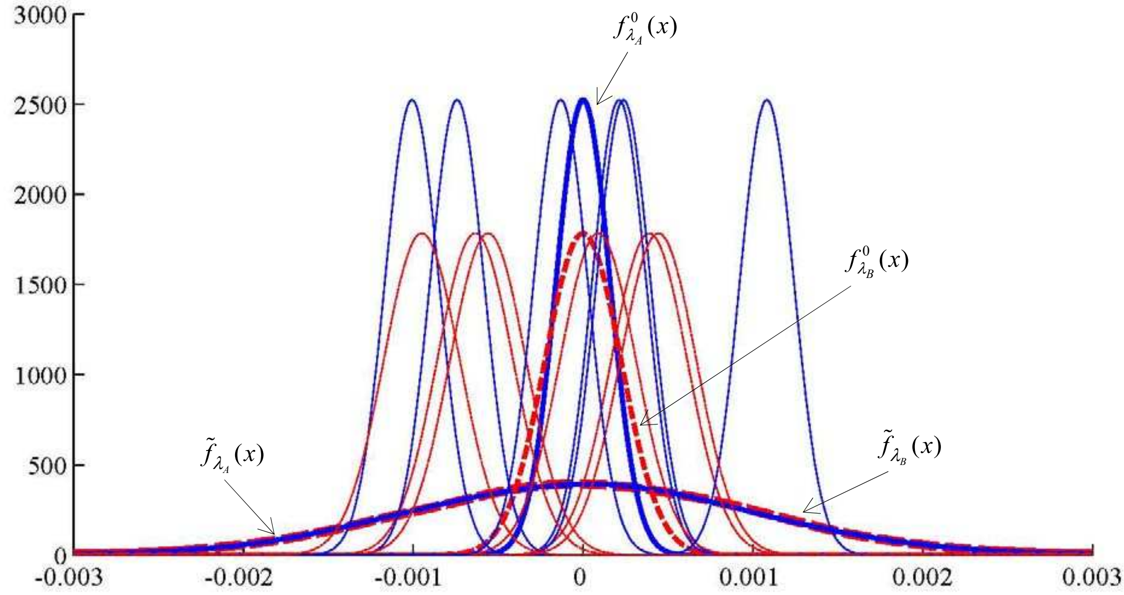

Figure 8: Density functions of median voter positions in large constituency (solid blue curves) and small constituency (dashed red curves) for varying median/mean voter positions. Uncertainty about the median/mean voter position results into flattened *ex ante* median densities.

as solutions to the inverse problem

$$\phi_j(w, q) = c \cdot n_j \qquad j = 1, \ldots, m, \;\; c > 0. \tag{15}$$

The finding that in the presence of heterogeneity weights should be chosen such that the Shapley-Shubik index is proportional to population size for each constituency is illustrated by Figure 9 for the EU Council of Ministers. Note that the proportional Shapley-Shubik rule holds for any quota used at the top tier, but the remarks concerning the inverse problem under high quotas still apply: Implementing the above rule requires a solution to the *inverse* problem of finding weights that yield the desired values. In general, only approximative solutions to this problem exist because the number of distinct voting games on the set of players M is finite, whereas the number of combinations of desired values is infinite. This technical problem is perceivable in Figure 9: when a 50%-majority rule is used, the Shapley-Shubik indices associated with best unconstrained weights are located nicely on the 45°-line, but under the 73.9%-quota they rather meander around that line.

In line with the above discussion, Figure 10 demonstrates that the transition from square root rule to a near-linear rule takes place very quickly. The square root rule survives $\sigma_H = 0.00001$, but already for $\sigma_H = 0.00005$ we get $\alpha^* = 0.58$. Given the small variation in the setting – in Maaser and Napel (2007) mean ideal points came from the degenerate normal distribution $N(\mu_U = 0, \sigma_H = 0)$ – the result differs strikingly from our previous finding $\alpha^* \approx 0.5$. The square root rule vanishes already for small degrees of heterogeneity.

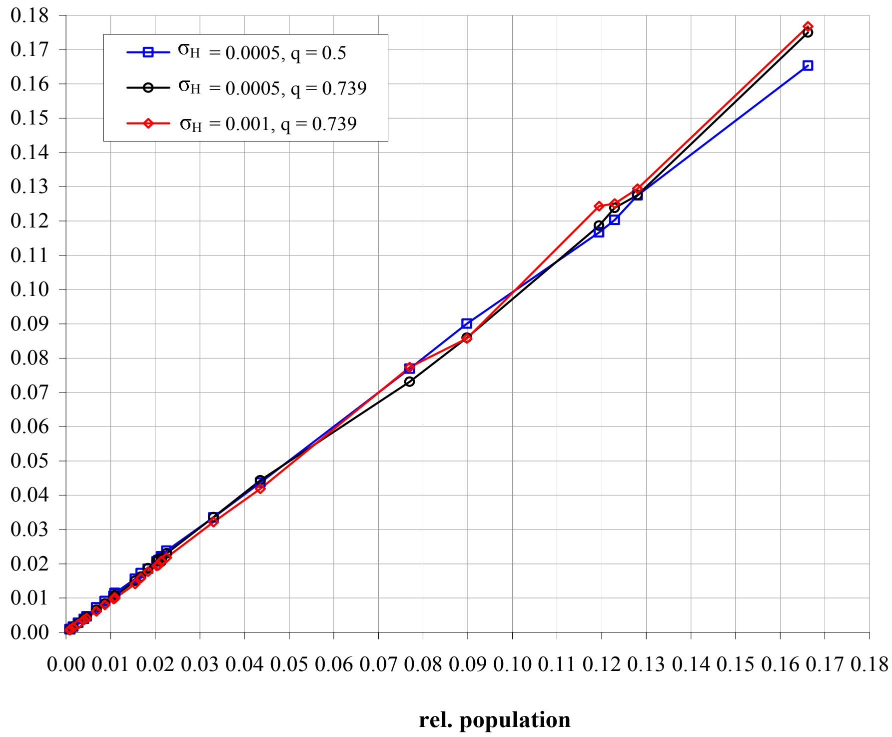

Figure 9: Shapley-Shubik linear rule: Shapley-Shubik indices for best unconstrained weights in EU27

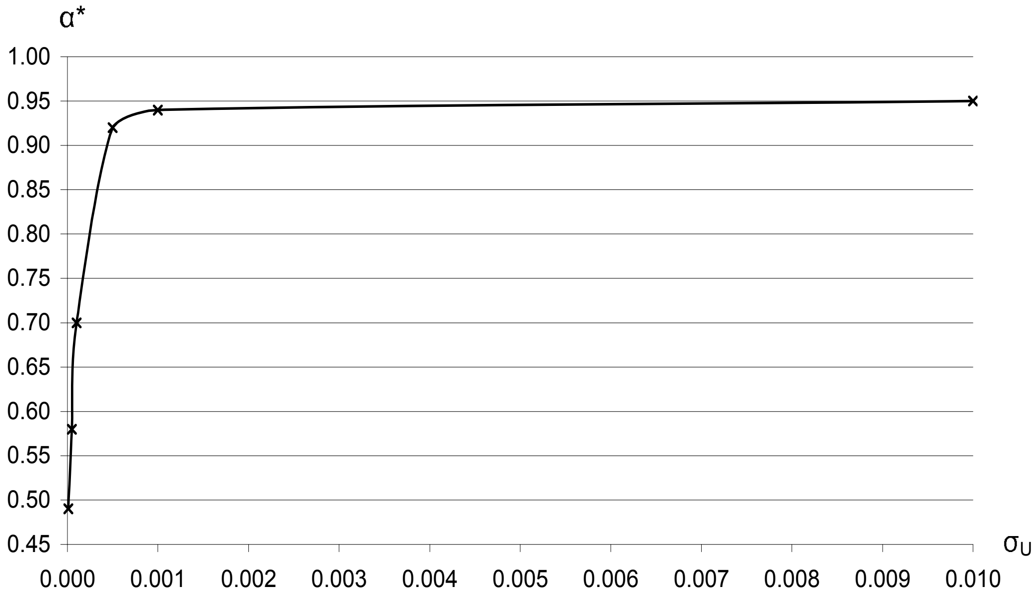

Figure 10: Rapid transition from square root rule to near-linear rule for EU27 ($q = 0.5$)

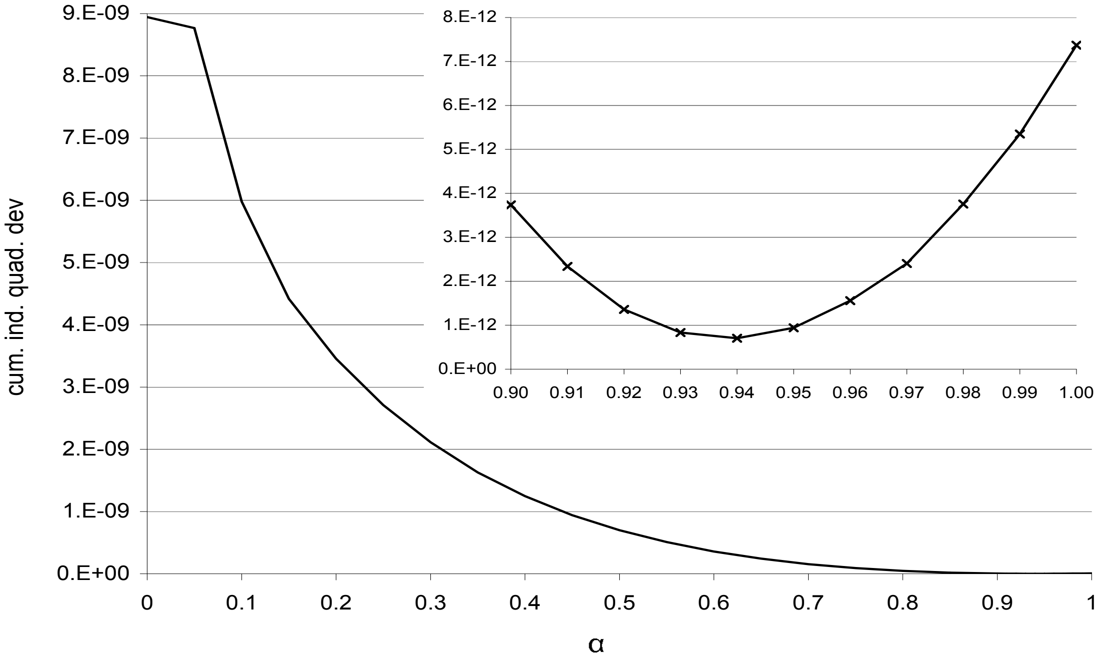

Figure 11: Cumulative individual quadratic deviation for EU27 ($\sigma_U = 0.001$, $q = 0.5$)

Figure 11 plots the objective criterion (4) versus α.[20] The mean ideal points μ_j of the 27 countries are drawn from the normal distribution $N(\mu_U = 0, \sigma_U = 0.001)$ and voters' ideal points in constituency j are uniformly distributed on $[\mu_j - 0.5, \mu_j + 0.5]$. Amongst all coefficients in $\{0, 0.01, \ldots, 1\}$ $\alpha = 0.94$ performs best, the cumulative individual quadratic deviation is 7.02×10^{-13}.

A second form of interim heterogeneity occurs if constituencies differ in their degree of "preference cohesion". In particular, it is conceivable that the strength of the 'common belief' decreases in the size of the society.[21] The variance σ_j^2 of the ideal point distribution F_j $(j = 1, \ldots, m)$ in constituency $\mathcal{C}_j$ can then be written as

$$\sigma_j^2 = g(n_j)$$

where $g(\cdot)$ is a monotonically increasing function.

Consider the simple case of proportionality, i.e., $g = a \cdot n_j$ where $a > 0$ is a constant, and assume that, in each constituency j, voters' ideal points are uniformly distributed on the interval $[x_{1j}, x_{2j}]$ with common mean $\mu_j = \mu$ for all for $j = 1, \ldots, m$. The density in

[20]Simulation input is Eurostat population numbers for EU27 countries as of 01/01/2007.

[21]The assumption that preferences are more heterogeneous in large populations is also made in Alesina and Spolaore (2003), and the trade-off between the costs of differences and the economies of scope in large jurisdictions determines nation size.

constituency j is

$$f_j(x) = \begin{cases} \frac{1}{x_{2j}-x_{1j}} & x_{1j} \le x \le x_{2j} \\ \\ 0 & \text{elsewhere,} \end{cases}$$

and the variance of the uniform distribution on $[x_{1j}, x_{2j}]$ is $\sigma_j^2 = (x_{2j} - x_{1j})^2/12$. Then, the length of the interval $[x_{1j}, x_{2j}]$ for constituency j with population n_j is proportional to $\sqrt{n_j}$. As the density $f_j(x)$ appears in (3), the standard deviation of the median ideal point is equalized for all constituencies and for all values of a. When the variance of voters' uniformly distributed ideal points increases proportionately in population size, the representatives' ideal points are i. i. d. For the reasons stated before, we arrive again at the recommendation to allocate weights such that each representative's Shapley-Shubik power is proportional to his constituency size.

Though the assumption that preferences are more widely spread in large constituencies could seem plausible at first glance, it is less convincing if one thinks about the policy space X as the carrier of individual preferences. Rather, preferences in a small society can be as varied as in a large society. Therefore the assumption that 'preference heterogeneity' is independent on population size, i.e., the variances σ_j^2 of the λ_i^j are taken to be identical for all constituencies.

It is interesting to note that the Sh-LR coincides with the 'neutral' voting rule that Laruelle and Valenciano (2008) obtain in the context of a bargaining committee. A bargaining committee consists of a voting rule $\mathcal{W}$ specifying the winning coalitions, and a m-person Nash bargaining problem $B = (U, d)$, where $U \subseteq \mathbb{R}^m$ is the set of feasible payoff vectors and d is the vector of status quo or disagreement payoffs. Unlike classical bargaining which is thought of as a unanimous decision process, an agreement in a bargaining committee only needs the support of a winning coalition to be implemented. Under the condition that the Shapley-Shubik index ϕ is accepted as a valid measure of bargaining power, Laruelle and Valenciano (2007) axiomatically derive a solution $F(B, \mathcal{W})$ to the bargaining committee problem $(B, \mathcal{W})$:

$$F(B, \mathcal{W}) = \mathrm{Nash}^{\phi(\mathcal{W})}(B) = \arg\max_{\substack{u \in U, \\ u \ge d}} \prod_{j=1}^{m} (u_j - d_j)^{\phi_j(\mathcal{W})}, \tag{16}$$

that is, an asymmetric Nash solution with weights given by the Shapley-Shubik indices of the committee members under the voting rule. Solution (16) can be regarded as a reasonable expectation of the utility levels when a general agreement is achieved in the bargaining committee.

A voting rule is called 'neutral' if every citizen in every constituency $\mathcal{C}_j$ is indifferent between bargaining for himself with all other citizens in the union or leaving the bargaining to a representative who bargains on behalf of $\mathcal{C}_j$ in the bargaining committee. As shown by Laruelle and Valenciano (2008), a rule with this property exists if citizens' preferences are $\mathfrak{C}$-symmetric. This property requires that all citizens in constituency $\mathcal{C}_j$ have a common status quo payoff d_j, and that the set of payoff vectors which are attainable for them

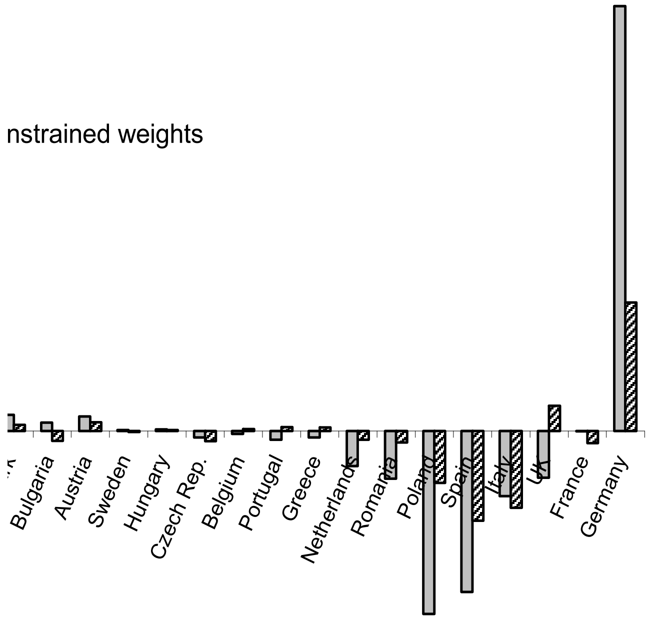

Figure 12: Absolute deviations of estimated from ideal probabilities for EU27 ($\sigma_U = 0.001$, $q = 0.5$)

must be symmetric for any fixed distribution of payoffs among non-members of $\mathcal{C}_j$. Under $\mathfrak{C}$-symmetry, citizens within the same constituency have the same bargaining characteristics. This makes it possible to condense the bargaining problem that would be faced by the 'committee of the whole', i.e., by all citizens bargaining directly, into the m-person bargaining problem $B = (U, d)$. Let u_i^{dNB} denote the payoff to citizen i under direct (and unweighted) Nash bargaining. Now suppose that i is a member of constituency $\mathcal{C}_j$. It is quite obvious that i would get a payoff equal to u_i^{dNB} if, for all j, the weight of representative j in an asymmetric Nash bargaining solution is proportional to the number of citizens in $\mathcal{C}_j$. In view of (16), this implies that a voting rule $\mathcal{W}$ is neutral if the Shapley-Shubik index of representative j under $\mathcal{W}$ is proportional to his constituency's population number n_j.

Neutrality of the voting rule can be interpreted in terms of equivalence between direct and indirect democracy. In the double median model, it requires that the outcome of the two-tiered decision-making process should equal the outcome of bargaining with assembly N, which corresponds to the median ideal point of all citizens. Generally, these two outcomes differ, but finding a weight allocation rule that is 'most neutral' in the sense of minimizing the discrepancy seems a worthwhile topic for future research.

7 Concluding remarks

Square root rules have been found to lack robustness in other contexts before. Yet, the literature so far has considered only the binary voting model and other forms of correlation among voters' preferences. For example, considering a binary voting model, Kirsch (2007) finds that square root weights minimize the difference between the margin of representatives accepting or rejecting a proposal and the size of the popular margin.[22] Yet, if each constituency exhibits a 'collective bias' (similar to Straffin's (1977) homogeneity assumption) whose strength is independent of constituency size, the optimal weights with respect to the above minimization problem are proportional to population numbers rather than to the square roots of the latter. Investigating the ideals of maximizing and equalizing expected utility, respectively, Beisbart and Bovens (2005) come to basically the same conclusion: with i. i.d. voters and simple majority rule, both ideals are met by simple square root weights. But if correlations of individual utilities within each constituency are introduced such that an individual is correlated with more others the larger the constituency is, then the square root rule quickly makes way for a proportional rule if the aim is maximizing expected utility, and to an equal weight allocation is the aim is equalizing expected utilities.

Proportional rules have been derived in various models, yet so far, they have not been shown to produce equal representation as it is understood here. Barberà and Jackson (2006) show in their 'fixed-number-of-blocks model' that, if each constituency consists of a given number of blocks of identical voters, setting weights proportional to population sizes maximizes expected utility irrespective of the voting threshold. Whereas all above findings pertain to binary decision-making, Laruelle and Valenciano (2008) study a committee whose members bargain over a convex space of alternatives. They demonstrate that a voting rule ensures 'neutral representation' in the sense that all individual citizens are indifferent between bargaining themselves or putting bargaining in the hands of a representative if each representative is provided with bargaining power proportional to his constituency's population size.

References

Alesina, A. and E. Spolaore (2003). *The Size of Nations*. Cambridge, MA: The MIT Press.

Algaba, E., J. M. Bilbao, J. R. Fernández García, and J. J. López (2003). Computing power indices in weighted multiple majority games. *Mathematical Social Sciences 46*, 63–80.

Ansolabehere, S., A. Gerber, and J. Snyder (2002). Equal votes, equal money: Court-ordered redistricting and the distribution of public expenditures in the American

[22]The difference between these two quantities is very similar to the mean majority deficit which is also minimized under square root weights (Felsenthal and Machover 1999).

States. *American Political Science Review 96*(4), 767–777.

Arnold, B. C., N. Balakrishnan, and H. N. Nagaraja (1992). *A First Course in Order Statistics*. New York: John Wiley & Sons.

Avriel, M. (1976) *Nonlinear programming: analysis and methods*. Englewood Cliffs, NJ: Prentice-Hall.

Baldwin, R. E., E. Berglöf, F. Giavazzi, and M. Widgrén (2001). *Nice Try: Should the Treaty of Nice Be Ratified?* Monitoring European Integration 11. London: Center for Economic Policy Research.

Banks, J. S. and J. Duggan (2000). A Bargaining Model of Social Choice. *American Political Science Review 94*(1), 73–88.

Barberà, S. and M. O. Jackson (2006). On the weights of nations: Assigning voting weights in a heterogeneous union. *Journal of Political Economy 114*(2), 317–339.

Beisbart, C. and L. Bovens (2005). Why degressive proportionality? An argument from cartel formation. mimeo, University of Dortmund and London School of Economics.

Beisbart, C. and L. Bovens (2007). Welfarist evaluations of decision rules for boards of representatives. *Social Choice and Welfare 29*(4), 581–608.

Beisbart, C., L. Bovens, and S. Hartmann (2005). A utilitarian assessment of alternative decision rules in the Council of Ministers. *European Union Politics 6*(4), 395–419.

Beisbart, C. and S. Hartmann (2006). Welfarism and the Assessment of Social Decision Rules. In U. Endriss and J. Lang (eds.), *Computational Social Choice 2006*, pp. 35–48. Amsterdam.

Bilbao, J. M., F. Bobay, W. Kirsch, M. Machover, I. McLean, B. Plechanovová, F. Pukelsheim, W. Słomczyński, K. Życzkowski, and many others. (2004). Letter to the governments of the EU member states. URL (consulted last in April 2006): http://www.esi2.us.es/∼mbilbao/pdffiles/letter.pdf.

Buchanan, J. M. and G. Tullock (1962). *The Calculus of Consent*. Ann Arbor, MI: University of Michigan Press.

Caplin, A. and B. Nalebuff (1988). On 64%-Majority Rule. *Econometrica 56*(4), 787–814.

Chamberlain, G. and M. Rothschild (1981). A note on the probability of casting a decisive vote. *Journal of Economic Theory 25*, 152–162.

Chang, P.-L., V. C. Chua, and M. Machover (2006). L S Penrose's limit theorem: Tests by simulation. *Mathematical Social Sciences 51*(1), 90–106.

Felsenthal, D. and M. Machover (1998). *The Measurement of Voting Power – Theory and Practice, Problems and Paradoxes*. Cheltenham: Edward Elgar.

Felsenthal, D. and M. Machover (1999). Minimizing the mean majority deficit: The second square-root rule. *Mathematical Social Sciences 37*(1), 25–37.

Felsenthal, D. and M. Machover (2001). The Treaty of Nice and qualified majority voting. *Social Choice and Welfare 18*(3), 431–464.

Felsenthal, D. and M. Machover (2004). Analysis of QM rules in the draft Constitution for Europe proposed by the European Convention, 2003. *Social Choice and Welfare 23*(1), 1–20.

Gelman, A., J. N. Katz, and J. Bafumi (2004). Standard voting power indexes don't work: An empirical analysis. *British Journal of Political Science 34*(1133), 657–674.

Gelman, A., J. N. Katz, and F. Tuerlinckx (2002). The mathematics and statistics of voting power. *Statistical Science 17*, 420–435.

Gerber, E. R. and J. B. Lewis (2004). Beyond the median: Voter preferences, district heterogeneity, and political representation. *Journal of Political Economy 112*(6), 1364–1383.

Good, I. and L. S. Mayer (1975). Estimating the efficacy of a vote. *Behavioral Science 20*, 25–33.

Goodin, R. E. and C. List (2006). Special Majorities Rationalized. *British Journal of Political Science 36*(2), 213–241.

Harstad, B. (2005). Majority rules and incentives. *Quarterly Journal of Economics 120*(4), 1535–1568.

Hollander, M. and D. A. Wolfe (1999). *Nonparametric Statistical Methods*(2nd ed.). New York: Wiley.

Kirsch, W. (2007). On Penrose's Square-root Law and Beyond. *Homo Oeconomicus 24*(3/4), 357–380.

Laruelle, A. and F. Valenciano (2007). Bargaining in committees as an extension of Nash's bargaining theory. *Journal of Economic Theory 132*(1), 291–305.

Laruelle, A. and F. Valenciano (2008). Bargaining in committees of representatives: The 'neutral' voting rule. *Journal of Theoretical Politics 20*(1), 93–106.

Leech, D. (2002). Designing the voting system for the EU Council of Ministers. *Public Choice 113*(3-4), 437–464.

Leech, D. (2003). Power Indices as an Aid to Institutional Design: The Generalised Apportionment Problem. In M. J. Holler, H. Kliemt, D. Schmidtchen, and M. E. Streit (Eds.), *Jahrbuch für Neue Politische Ökonomie 22*, pp. 107–121. Tübingen: Mohr Siebeck.

Leech, D. and M. Machover (2003). Qualified Majority Voting: The Effect of the Quota. In M. J. Holler, H. Kliemt, D. Schmidtchen, and M. E. Streit (Eds.), *Jahrbuch für Neue Politische Ökonomie 22*, pp. 127–143. Tübingen: Mohr Siebeck.

Lindner, I. and M. Machover (2004). L. S. Penrose's limit theorem: Proof of some special cases. *Mathematical Social Sciences 47*, 37–49.

Maaser, N. and S. Napel (2007). Equal Representation in Two-tier Voting Systems. *Social Choice and Welfare 28*(3), 401-420.

May, K. O. (1952). A Set of Independent, Necessary and Sufficient Conditions for Simple Majority Decision. *Econometrica 20*(4), 680–684.

McGann, A. J. (2004). The tyranny of the supermajority: How majority rule protects minorities. *Journal of Theoretical Politics 16*(1), 53–77.

Miller, N. (1983). Social Choice and Pluralism. *American Political Science Review 77*(3), 734–747.

Nitzan, S. and J. Paroush (1984). Are qualified majority rules special? *Social Choice and Welfare 18*(3), 431–464.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society 109*, 53–57.

Penrose, L. S. (1952). *On the Objective Study of Crowd Behaviour.* London: H. K. Lewis & Co.

Rae, D. (1975). The Limits of Consensual Decision Making. *American Political Science Review 69*(4), 1270–1294.

Riker, W. H. (1961). Voting and the summation of preferences: An interpretive bibliographical review of selected developments during the last decade. *American Political Science Review 55*(4), 900–911.

Shapley, L. S. and M. Shubik (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review 48*(3), 787–792.

Słomczyński, W. and K. Życzkowski (2007). From a Toy Model to the Double Square Root Voting System. *Homo Oeconomicus 24*{(3/4), 381–399.

Straffin, P. D. Jr. (1977). Homogeneity, independence, and power indices. *Public Choice 30*(1), 107–118.

Tiebout, C. (1956). A Pure Theory of Local Expenditure. *Journal of Political Economy 64*, 416–424.

Wicksell, K. (1969[1896]). *Finanztheoretische Untersuchungen nebst Darstellung und Kritik des Steuerwesens Schwedens.* Aalen: Scientia Verlag.

Widgrén, M. (1995). Probabilistic voting power in the EU Council: The cases of trade policy and social regulation. *Scandinavian Journal of Economics 97*(2), 345–356.