Equal Representation in Two-tier Voting Systems*

Nicola Maaser
University of Hamburg
Department of Economics
Von-Melle-Park 5
D-20146 Hamburg
nicolamaaser@yahoo.de

Stefan Napel
University of Hamburg
Department of Economics
Von-Melle-Park 5
D-20146 Hamburg
napel@econ.uni-hamburg.de

July 7, 2005

—— PRELIMINARY AND INCOMPLETE DRAFT ——

Abstract

The paper investigates how voting weights should be assigned to differently sized constituencies of an electoral college if the objective is to give voters in all constituencies equal indirect influence on decisions. In contrast to existing literature, decisions are elements of a one-dimensional convex policy space. Individual a priori power is identified with the joint probability that a voter is the median at the level of the constituency *and* that his or her constituency is pivotal in the electoral college. Extensive Monte-Carlo simulations demonstrate that the *square root rule*, which ensures approximate power equalization for dichotomous decisions, comes close to equalizing voters' influence also for much richer decision environments, provided that a 50% quota is used. The paper also addresses the question of how (un)equal voters' representation is in the EU's Council of Ministers and US Electoral College.

Keywords: voting power, equal representation, fair voting systems

---

*x

1 Introduction

Free and fair elections are a cornerstone of democracy and indispensable to its long-run survival. Whether or not elections are regarded as fair is largely determined by the voting and decision-making procedures. In particular, the principle of "one person, one vote" provides a litmus test for fairness of voting. However, there exist other important criteria for the design of a voting system. Efficiency considerations or historical reasons can, for example, provide a strong rationale for construction of a two-tier voting system in which people's policy preferences are first aggregated in multiple constituencies, e.g. districts or states of a federation or members of a supranational institution, and then again among representatives from the constituencies. Unless all constituencies contain an equal number of voters, implementation of the one-person-one-vote principle then becomes non-trivial and, possibly, only a second-best approximation of it may be feasible. In particular, assigning each voter one vote in his or her constituency and each constituency one vote at the aggregate level results in greater average influence for voters coming from smaller constituencies.

Penrose's (1946) *square root rule* (also see Felsenthal and Machover 1998) gives an answer to the problem of (approximately) equal representation in two-tier electoral systems. It rests on a particular model of decision making. Namely, it considers dichotomous 'yes' or 'no' (0 or 1) decisions and assumes that all bottom-tier voters cast their vote independently of each other and with equal probability for both alternatives.[1] From these assumptions it follows that constituencies' voting power at the top-tier (as measured by the Penrose-Banzhaf index[2]) should be proportional to the *square root* of the respective constituency size.

Variations of Penrose's model of decision making can be expected to lead to different optimal weight assignment. This issue has received surprisingly little attention. Some recent empirical work, such as Gelman, Katz, and Tuerlinckx (2002), questions the realism of Penrose's equiprobability assumption (or its generalizations, see fn. 1) and argues in favor of something closer to direct proportionality of population and weight. Recent theoretical work like Barberà and Jackson (2004) or Beisbart, Bovens, and Hartmann (2004) investigates the optimal assignment of weights for a utilitarian rather than egalitarian objective function.

To our knowledge, this paper is the first to investigate equal representation when decisions are elements of a one-dimensional convex policy space, e.g. elements of interval $[0, 1]$. To us, this seems at least as relevant and realistic as decisions concerning the set

[1] More generally, one can consider a doubly-probabilistic behavioral model in which any voter i independently of all other voters accepts an unspecified proposal with a probability p_i which is in turn the realization of a random variable P_i distributed independently across voters with mean $\mu_i = 0.5$.

[2] The Penrose-Banzhaf index (Penrose 1946; Banzhaf 1965) is approximately proportional to voting weight if the number of constituencies is large and constituency sizes are not too different (see Lindner and Machover 2004). Therefore, calling for proportionality of square root of population and *voting weight* (rather than *voting power*) entails only small imprecision e. g. for the EU's Council of Ministers or US Electoral College.

$\{\text{yes}, \text{no}\}$ or $\{0, 1\}$. Supported by various political and economic models, we assume that the policy advocated by the top-tier representative of any given constituency $\mathcal{C}_j$ coincides with the ideal point of the constituency's bottom-tier *median voter*. The decision taken at the top tier corresponds to the position of the pivotal representative, where pivotality is determined by the weights assigned to (representatives of) the constituencies and a 50% quota. A two-tier voting system then ensures equal representation and obeys the one-person-one-vote principle if and only if the joint probability of any bottom-tier voter determining his or her constituency's position *and* of this constituency being pivotal at the top-level is equal for all voters.

We assume that ideal points of bottom-tier voters are a priori independently identically distributed. Then any voter from constituency $\mathcal{C}_j$ with population n_i has probability $1/n_j$ to be pivotal at the bottom level. So equal representation requires that the probability of constituency $\mathcal{C}_j$ being pivotal is proportional to n_j. Finding or at least characterizing a suitable weight assignment which ensures proportionality is non-trivial because the size n_j influences the a priori distribution of $\mathcal{C}_j$'s representative's position. In particular, variance of $\mathcal{C}_j$'s ideal point is larger, the smaller is n_j. Assigning weights proportional to n_j therefore would give citizens of large countries too much power, while uniform weights give them too little power.

Analytical treatment of the problem is hard and we conjecture a neat and general result similar to Penrose's square root rule to be impossible except (perhaps) for special limit situations. Therefore, we conduct extensive Monte-Carlo simulations. We vary the weight assignment and search for rules that minimize the (cumulative quadratic) deviation from equal representation in computer-generated as well as real-world two-tier electoral systems. It turns out that, based on a decision quota of 50% at the top tier, weights that are proportional to the square root of population are close to optimal for many population distributions (but not all).

The paper thus demonstrates that *if* constituencies are sufficiently similar in size (no constituency would become a dummy or dictator if the square root rule were applied) and a 50% quota is used, Penrose's square root rule is a quite robust norm for egalitarian design of two-tier voting systems. It is known to be the right reference point in the context of 'yes'-or-'no'-referenda (with independence and equiprobability as a focal formalization of a *veil of ignorance* guiding constitutional design). We show that it is also the right reference point in to us more realistic settings with finely graded policy alternatives on a continuous left–right, high tax–low tax, environment friendly–business friendly, etc. scale.

The remainder of the paper is structured as follows: Section 2 introduces our model and points out the difficulties in providing analytical results. Section 3 then discusses our Monte-Carlo approach and presents results for randomly generated population configurations as well as EU's Council of Ministers and US Electoral College. Section 4 concludes.

2 Model

Consider a large population of *voters* partitioned into m *constituencies* $\mathcal{C}_1, \ldots, \mathcal{C}_m$, with $n_j = |\mathcal{C}_j| > 0$ members each. Voters' preferences are single-peaked with ideal point λ_i^j (for $i = 1, \ldots, n_j$ and $j = 1, \ldots, m$) in the convex one-dimensional *policy space* $X \equiv [0, 1]$. Assume that all n_j are odd numbers.

Let $\cdot : n_j$ denote the permutation of voter numbers in constituency j such that

$$\lambda^j_{1:n_j} \leq \ldots \leq \lambda^j_{n_j:n_j}$$

holds. In other words, $\lambda^j_{k:n_j}$ denotes the *k-th order statistic* of $\lambda^j_1, \ldots, \lambda^j_{n_j}$, i. e. the k-th leftmost ideal point in $\mathcal{C}_j$.

A policy $x \in X$ is decided on by an *electoral college* $\mathcal{E}$ consisting of one representative from each constituency. It is assumed that the representative of $\mathcal{C}_j$, denoted by j, adopts the ideal point of his constituency's *median voter*, denoted by $\lambda^j \equiv \lambda^j_{(n_j+1)/2:n_j}$. Let $\lambda^{k:m}$ denote the k-th order statistic of $\lambda^1, \ldots, \lambda^m$.

In the electoral college $\mathcal{E}$, 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 $q \equiv \frac{1}{2} \sum_{j=1}^m w_j$, i. e. half of the total weight, can implement a policy $x \in X$. We assume that the agreed policy is – e. g. in the equilibrium of some non-cooperative game which reflects the decision procedure applied in $\mathcal{E}$ – equal to the ideal point of the *pivotal representative*, $P : m$, in $\mathcal{E}$. The random pivotal position P is defined by

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

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.[3]

Given this setting, consider the following egalitarian norm: *Each voter in any constituency should have an equal chance to determine the policy implemented by the electoral college.* More formally, there should exist a constant $c > 0$ such that

$$\forall j \in \{1, \ldots, m\} \colon \forall i \in \mathcal{C}_j \colon \Pr\left(\lambda^j_i = \lambda^j = \lambda^{P:m}\right) \equiv c. \tag{1}$$

We would like to answer the following question: which allocation of weights $w_1, \ldots, w_m$ satisfies this norm (at least approximately) for an arbitrary given partition of the total population into m constituencies? In other words we search for an analogue of *Penrose's square root rule* (see e.g. Felsenthal and Machover 1998), which concerns a considerably simpler binary model of collective decision-making rather than arbitrary policies in X.[4]

[3] Things are more complicated if $q > \frac{1}{2} \sum_{j=1}^m w_j$ is assumed. Then, the complement of a losing coalition need no longer be winning. In this case there may not exist *any* policy $x \in X$ which beats all alternatives $x' \neq x$ despite unidimensionality of X and single-peakedness of preferences.

[4] Given Penrose's assumption of independent equiprobable random 'yes'-or-'no' votes by each voter in

Probability $\Pr\left(\lambda_i^j = \lambda^j = \lambda^{P:m}\right)$ depends on the distribution of all voters' ideal points. Though in practice ideal points in different constituencies may come from different distributions on X and may exhibit various dependencies, it is appealing from the normative viewpoint of constitutional design to presume that the ideal points of all voters in all constituencies are *independently and identically distributed* (i. i. d.).

If voters' ideal points in constituency $\mathcal{C}_j$ are i. i. d., each voter $i \in \mathcal{C}_j$ has the same probability to be its median.[5] Hence,

$$\forall j \in \{1, \ldots, m\} \colon \forall i \in \mathcal{C}_j \colon \; \Pr\left(\lambda_i^j = \lambda^j\right) = \frac{1}{n_j}.$$

Using that the events $\{\lambda^j = \lambda^{P:m}\}$ and $\{\lambda_i^j = \lambda^j\}$ are independent, we can write (1) as

$$\forall j \in \{1, \ldots, m\} \colon \forall i \in \mathcal{C}_j \colon \frac{\Pr\left(\lambda^j = \lambda^{P:m}\right)}{n_j} \equiv c. \tag{2}$$

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.

Suppose for a moment that *representatives'* ideal points $\lambda^1, \ldots, \lambda^m$ are independently and also identically distributed. Then, $\Pr\left(\lambda^j = \lambda^{P:m}\right)$ is simply the Shapley-Shubik index (SSI) value, $\phi_j(w, q)$, of representative j in voting body $\mathcal{E}$ defined by weight vector $w = (w^1, \ldots, w^m)$ and quota q (see Shapley and Shubik 1954). Therefore, making the i. i. d. assumption at the level of representatives implies a *linear rule* (based on the SSI) as the replacement for Penrose's square root rule in our context of spatial rather than binary voting. In other words, w has to be chosen such that $\phi_j(w, q)$ is directly proportional to population size n_j for all constituencies $j = 1, \ldots, m$.

However, making the in our view more plausible and egalitarian i. i. d. assumption at the level of *individual voters*, representatives' ideal points $\lambda^1, \ldots, \lambda^m$ are independently but – except in the trivial case $n_1 = \ldots = n_m$ – *not* identically distributed: If, for example, all voters' ideal points are *uniformly* distributed, then in any constituency $\mathcal{C}_j$ the median position λ^j is *beta distributed* with mean $\mu^j = 1/2$ and standard deviation $\sigma^j = 1/\sqrt{4(n_j + 2)}$. The latter approaches zero for $n_j \to \infty$.

More generally, voter ideal points may come from an (identical) *arbitrary* distribution F with density f. In this case (see e.g. Arnold et al. 1992), $\mathcal{C}_j$'s median position is

every constituency – which deterministically define each constituency's 'yes'-or-'no' vote in the electoral college by simple majority –, quadrupling a constituency's size approximately (invoking Stirling's formula) halves a given voter's chance to be pivotal in it. The voter must hence be compensated by weights that double the constituency's probability of being pivotal at the top tier. This probability (=the Penrose-Banzhaf index) is usually *not* proportional to voting weight, in particular for a small number of constituencies. Hence, implementing the square root rule as well as possible hence requires numerical solution of the *inverse problem* of finding weights which induce a desired Penrose-Banzhaf power distribution (see Leech 2002).

[5] We disregard the possibility that two ideal points coincide, in which case the median voter – in contrast to the median policy – is not well-defined. This is innocuous for any continuous ideal point distribution.

asymptotically normally distributed with mean

$$\mu^j = F^{-1}(0.5)$$

and standard deviation

$$\sigma^j = \frac{1}{2\, f(F^{-1}(0.5))\sqrt{n_j}}.$$

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 (assumed to be identical for all λ_i^j). This makes the representative of a larger constituency on average more central in the electoral college and more likely to be pivotal in it, given any fixed weight allocation. Weights and SSI of large constituencies hence should be less than in the case considered in the previous paragraph, i. e. the correct rule must be increasing but less than linearly.

It is in our view an important observation that the assumption of *collective preferences* of each constituency having an identical a priori distribution is inconsistent with the assumption that all *individual preferences* are a priori identically distributed. The intuitively appealing linear rule of giving twice the voting power to a constituency double the size violates the one-person-one-vote principle if one makes the latter assumption. In the context of constitutional design we find it considerably more convincing and will assume i. i. d. ideal points for individual voters throughout this paper.

Probability $\Pr\left(\lambda^j = \lambda^{P:m}\right)$ in (2) depends both on the different distributions of representatives' ideal points, i. e. the different standard deviations σ^j determined by constituency sizes n_j, and the voting weight assignment. Unfortunately, computation of the probability of a given constituency $\mathcal{C}_j$ being pivotal is already a complex numerical task even for simple majority voting with *uniform weights*. In this most simple case, the representative of $\mathcal{C}_j$ with *median* ideal point is always pivotal in $\mathcal{E}$, i. e. $P \equiv (m+1)/2$. Define $N^j \equiv \{1, \ldots, j-1, j+1, \ldots, m\}$ as the index set of all constituencies except $\mathcal{C}_j$. Then, the probability of constituency $\mathcal{C}_j$ being pivotal is

$$\begin{aligned}
\Pr\left(\lambda^j = \lambda^{(m+1)/2:m}\right) &= \Pr\left(\text{exactly } \tfrac{m-1}{2} \text{ of the } \lambda^k, k \neq j, \text{ satisfy } \lambda^k < \lambda^j\right) \\
&= \sum_{\substack{S \subset N^j, \\ |S|=(m-1)/2}} \prod_{k \in S} F_k(\lambda^j) \cdot \prod_{k \in N_j \setminus S} (1 - F_k(\lambda^j)) \\
&= \int \sum_{\substack{S \subset N^j, \\ |S|=(m-1)/2}} \prod_{k \in S} F_k(x) \cdot \prod_{k \in N_j \setminus S} (1 - F_k(x)) \cdot f_j(x)\, dx,
\end{aligned} \tag{3}$$

where f_j and F_j denote the density and cumulative density functions of λ^j $(j = 1, \ldots, m)$.

It seems feasible (but is beyond the scope of this paper) to provide an asymptotic approximation for this probability as a function of constituency sizes $n_1, \ldots, n_m$ for special cases, e. g. when $n_2 = \ldots = n_m$ (hence $F_2 = \ldots = F_m$) and only n_1 varies (resp., F_1 and f_1). However, we doubt the existence of such an approximation for general configurations

$(n_1, \ldots, n_m)$. And even if one could obtain one, it would allow to compute a measure of the inequity associated with uniform weights, but be of little help in finding weights that achieve equity.

The weighted voting analogue of (3) is even more intractable. For a *given* realization p of random variable P, the pivotal member of $\mathcal{E}$, one might be able to approximate

$$\Pr(\lambda^j = \lambda^{p:m}) \;=\; \int \sum_{S \subset N^j,\ |S|=P-1} \prod_{k \in S} F_k(x) \cdot \prod_{k \in N^j \setminus S} (1 - F_k(x)) \cdot f_j(x)\, dx.$$

But events $\{P = p\}$ and $\{\lambda^j = \lambda^{p:m}\}$ are not independent for non-uniform weights and so, typically,

$$\Pr(\lambda^j = \lambda^{P:m}) \;\neq\; \sum_{p=1}^{m} \Pr(P = p) \cdot \Pr(\lambda^j = \lambda^{p:m}).$$

To see this, consider the extreme case of representative j having weight $w_j > 0.5$ even though all constituencies are of equal size, so that ideal points λ^k $(k = 1, \ldots, m)$ are i. i. d. Since j is a dictator, $\Pr(\lambda^j = \lambda^{P:m}) = 1$. But $\Pr(P = p) = 1/m$ for all p and $\Pr(\lambda^j = \lambda^{p:m}) = 1/m$ for given p and all j.

To sum up, a purely analytical investigation of the model is unlikely to offer much insight. The following section therefore uses Monte-Carlo simulation in order to approximate the probability of any constituency $\mathcal{C}_j$ being pivotal for partition $\{\mathcal{C}_1, \ldots, \mathcal{C}_m\}$ and a fixed weight vector $(w_1, \ldots, w_m)$ and, based on this, to find weights $(w_1^*, \ldots, w_m^*)$ which approximately satisfy the equal representation conditions (1) or (2).

3 Simulation

The probability $\pi_j \equiv \Pr\left(\lambda^j = \lambda^{P:m}\right)$ can also be viewed as an *expected value*, namely of the random variable $H_j \equiv g_j^w(\lambda^1, \ldots, \lambda^m)$ which equals 1 if $\lambda^j = \lambda^{P:m}$ holds for given weight vector w and realized median ideal points $\lambda^1, \ldots, \lambda^m$, and 0 otherwise. The *Monte-Carlo method* then exploits that the empirical average of s independent draws of H_j

$$\bar{h}_j^s = \frac{1}{s} \sum_{l=1}^{s} h_j^l$$

converges to H_j's theoretical expectation

$$\mathbf{E}(H_j) = \pi_j$$

by the *law of large numbers*. The speed of convergence in s can be assessed by the sample variance of $h_1, \ldots, h_s$. Using the *central limit theorem*, it is thus possible to obtain estimates of π_j with a desired precision (referring e. g. to a 95%-confidence interval) if one generates and analyzes a sufficiently large number s of realizations of H_j.

To obtain a realization h_j^l of H_j, first, we draw[6] m random numbers $\lambda^1, \ldots, \lambda^m$ from distributions $F_1, \ldots, F_m$. Throughout our analysis, we take F_j to be a *beta distribution* with parameters $\big((n_j+1)/2, (n_j+1)/2\big)$. This corresponds to the median of n_j independently $[0,1]$-uniformly distributed voter ideal points, i. e. all individual voter positions are i. i. d. uniformly.[7] Second, the realized constituency positions are sorted and the pivotal position p (the smallest r such that aggregate weight of the r leftmost constituencies exceeds half the total weight) is determined. Constituency $\mathcal{C}_{p:m}$ is thus identified as the pivotal player of $\mathcal{E}$. It follows that $h_j^l = 1$ for $j = p:m$, and 0 for all other constituencies.

Our goal is to identify a simple rule for assigning voting weights to constituency sizes which – if it exists – approximately satisfies equal representation conditions (1) or (2) for various numbers m of constituencies and population configurations $\{\mathcal{C}_1, \ldots, \mathcal{C}_m\}$. A natural starting point is the investigation of *power laws*

$$w_j = n_j{}^{\alpha} \tag{4}$$

with $\alpha \in [0,1]$, including (for big m approximately) Penrose's square root rule as the special case $\alpha = 0.5$.

For any given m and population configuration $\{\mathcal{C}_1, \ldots, \mathcal{C}_m\}$ under consideration, we fix α, then approximate π_j by its empirical average by a run of 10 million iterations, and finally compute the cumulative quadratic deviation of estimated probabilities $\hat{\pi}_j$ and the egalitarian or ideal probability $\pi_j^* = n_j / \sum_{k=1}^m n_k$. This is repeated for different values of α, ranging from 0 to 1 with different step sizes, in order to find the power law or coefficient α which comes closest to implying equal representation under our distributional assumptions for the given configuration. Section 3.1 first looks at computer-generated random environments with constituency numbers between 10 and 50; we investigate several population configurations for each m to check the robustness of the optimal α. Sections 3.2 and 3.3

3.1 Randomly generated configurations

Table 1 reports the optimal values of α that were obtained for four sets of configurations $\{\mathcal{C}_1, \ldots, \mathcal{C}_m\}$.[8] For $m \in \{10, 15, 20, 25, 30, 40, 50\}$, constituency sizes $n_1, \ldots, n_m$ were independently drawn from a uniform distribution over $[0.5 \cdot 10^6,\ 99.5 \cdot 10^6]$. Numbers in column (I) are the the optimal $\alpha \in \{0, 0.1, \ldots, 0.9, 1\} \subset [0,1]$, where probabilities π_j were estimated from a simulation with 10 mio. iterations and where – throughout this paper – our optimality criterion is minimal cumulative quadratic deviation from the egalitarian norm at the level of constituencies.[9] Cumulative quadratic deviations for optimal α's are

[6]We use a self-made computer program. The *Java* source code is available upon request.

[7]The mentioned asymptotic results for order statistics imply that only F's median position and density at the median matter when constituency sizes are large. So below findings are not specific to the assumption of uniform distributions at the bottom tier.

[8]The configuration draws are independent across different values of m. Thus, the table actually reports optimal values obtained for 28 *independent* configurations.

[9]This weights deviations for all constituencies equally. We expect no qualitative changes for an objective function that weights deviations of $\hat{\pi}_j$ from $n_j / \sum n_k$ e. g. with n_j.

# const	(I)	(II)	(III)	(IV)
10	0.5	0.6	0.39	0.00
	(6.38×10^{-4})	(4.27×10^{-4})	(7.47×10^{-5})	(1.31×10^{-3})
15	0.5	0.5	0.49	0.48
	(5.65×10^{-6})	(4.07×10^{-6})	(1.19×10^{-6})	(3.18×10^{-6})
20	0.5	0.5	0.49	0.49
	(2.39×10^{-6})	(4.68×10^{-6})	(3.43×10^{-7})	(6.52×10^{-7})
25	0.5	0.5	0.49	0.49
	(4.09×10^{-7})	(6.93×10^{-7})	(1.63×10^{-7})	(3.29×10^{-7})
30	0.5	0.5	0.49	0.49
	(6.19×10^{-7})	(3.36×10^{-7})	(1.64×10^{-7})	(9.42×10^{-8})
40	0.5	0.5	0.49	0.49
	(1.63×10^{-7})	(2.39×10^{-7})	(1.50×10^{-7})	(1.51×10^{-7})
50	0.5	0.5	0.50	0.50
	(1.58×10^{-7})	(2.06×10^{-7})	(1.74×10^{-7})	(1.36×10^{-7})

Table 1: Optimal value of α for uniformly distributed constituency sizes (cumulative squared deviations from ideal probabilities in parentheses)

shown in brackets. Column (II) reports values obtained for a second set of independent realizations of constituency sizes; columns (III) and (IV) do likewise but using the finer grid $\alpha \in \{0, 0.01, 0.02, \ldots, 0.99, 1\}$.[10]

While results for $m = 10$ are inconclusive, $\alpha \approx 0.5$ is the very robust ideal parameter for larger number of constituencies. The reported cumulative quadratic deviations are so small that even if the power laws assumed in (4) should not contain the theoretically best rule for equal representation in our median-voter context (e. g., because constituencies' weights are not the right reference point, but rather their Penrose-Banzhaf or Shapley-Shubik index values), the latter seem to allow for a sufficiently good approximation for all practical purposes.

Results in Table 1 are suggesting that (an approximation of) Penrose's square root rule actually holds also in the context of policy decisions in $[0, 1]$. But one may wonder if optimality of $\alpha \approx 0.5$ is not an artifact of considering uniformly distributed constituency sizes $n_1, \ldots, n_m$ for any given m. Uniformity supposes that small constituencies are as

[10] Hence column (III), for example, reports on $101 \cdot 7$ simulation runs with 10 mio. iterations each.

likely to occur as large ones.

Constituency sizes at the national level are a matter of history or design. In the latter case, one might expect them to be clustered around some intermediate level; this would make a (truncated) normal distribution around some value $\bar{n}$ rather than uniformity focal for our experiments.[11] In the former case, when historical boundaries determine a population partition, a natural theoretical benchmark is a power law distribution for n_j. As a convenient approximation of distributions such as *Zipf's law* (or *zeta distribution*), which has big empirical support in a variety of contexts,[12] we consider the Pareto distribution with cumulative density and density functions

$$\begin{aligned} G(x|k,\, x_m) &= Pr(X \le x) = 1 - \left(\frac{x}{x_m}\right)^{-k} \\ g(x|k,\, x_m) &= k\,\frac{x_m^k}{x^{k+1}} \end{aligned}$$

and support for $x \in [x_m, \infty]$. Parameter x_m provides a lower bound on n_j and parameter k determines how quickly the probability of drawing a large (rather than small or medium-sized) constituency approaches 0.

Table 2 reports simulations with constituency sizes drawn from the Pareto densities $g(x|1.0,\, 0.1)$, $g(x|1.8,\, 0.1)$, $g(x|3.4,\, 0.1)$, and $g(x|5.0,\, 0.1)$, where numbers refer to million inhabitants. As long as the distribution is only moderately skewed (small k), findings correspond to those for the uniform distribution: A power law with $\alpha = 0.5$ gets close to ensuring equal representation provided that the number of constituencies is sufficiently large. However, for a distribution of constituency sizes that is heavily skewed, corresponding to mostly small constituencies and only one or very few large constituencies (reminiscent of atomic players in an otherwise oceanic game), other values of α – coming close to giving all constituencies an equal size-independent weight – perform better.[13] We conclude that $\alpha = 0.5$ does very well as long as m is not too small and constituency sizes $n_1, \ldots, n_m$ do not have isolated outliers.

3.2 EU Council of Ministers

Table 3 reports the empirical frequencies of individual members of the European Union's *Council of Ministers* being pivotal under weight assignment according to $w_j = {n_j}^b$ and

[11] Preliminary simulations using normally distributed constituency sizes support $\alpha \approx 0.5$ for a sufficiently big number of constituencies.

[12] Examples for which (approximative) power-law behavior has been claimed include sizes of human settlements (Gabaix 1999; Reed 2004), frequencies of words in long sequences of text, visits to web sites, the value of oil reserves in oil fields, and the size of meteor impacts on the moon. Explanations for this widespread regularity are based on ideas such as self-organized criticality and highly optimized tolerance (see e. g. Newman 2000).

[13] At this point, the fact that we do not weight quadratic deviation from the ideal probabilities with e. g. constituency size in our objective function may play a significant role. Choosing the right penalty function for the approximation of perfectly equal representation is a normative question worthy of more discussion (also in the context of the square root rule in Penrose's original 0-1 setting).

	Number of constituencies				
k	**10**	**20**	**30**	**40**	**50**
1.0	**0.5**	**0.5**	**0.5**	**0.5**	**0.5**
	(8.06×10^{-4})	(5.23×10^{-5})	(8.33×10^{-6})	(1.11×10^{-4})	(9.34×10^{-4})
1.8	**0.5**	**0.5**	**0.5**		
	(4.96×10^{-4})	(3.09×10^{-5})	(1.67×10^{-5})		
3.4	**0.0**	**0.5**	**0.5**	**0.5**	**0.5**
	(4.76×10^{-4})	(2.11×10^{-5})	(1.03×10^{-5})	(9.81×10^{-7})	(7.86×10^{-7})
5.0	**0.0**	**0.05**	**0.1**	**0.05**	**0.1**
	(1.86×10^{-3})	(5.48×10^{-4})	(2.18×10^{-5})	(5.97×10^{-6})	(2.95×10^{-5})

Table 2: Optimal values of α for Pareto distributed constituency sizes (cumulative squared deviations from ideal probabilities in parentheses)

a decision quota of $q = 50\%$. Values are averages of six simulation runs with 10 million iterations each, and exact to the third digit (based on the 95% confidence interval).

Numbers are illustrated for selected values of α in Figure 2. It nicely shows how close the probability of country j being pivotal comes to that required for (a priori) perfectly equal representation if weights are proportional to square root of population and if a 50% quota is used.

With the exceptions of Germany, Spain and Poland, the weights agreed in the Treaty of Nice roughly correspond to square root of populations. Figure 2 confirms that if a quota of 50 % were used in the Council of Ministers, estimated probabilities π_j would be rather close to their egalitarian values (with the mentioned exceptions).[14] However, a qualified majority of 72.2% of the weight (and additionally two majority requirements regarding the total population represented by Council members supporting a motion and the number of supporters) is needed for most Council decisions. For a qualified majority rule, greater centrality of median opinion in large countries such as Germany or France no longer provides greater chances of being pivotal in the Council. It actually reduces them. As illustrated in Figure 3, not only is representation even more biased against German voters but now also French, British, and Italian representatives are less often pivotal than would be necessary to give individual voters an equal chance of indirectly determining the Council's aggregate policy position.

This illustrates that a high quota is not only challenging the efficiency of a decision-making body such as the Council of Ministers in terms of the probability that random proposals are passed in the classical 0-1 setting (see Felsenthal and Machover 2001, Bald-

[14] This ignores the cumulative population and number-of-members requirements also agreed in the Treaty.

	b										
	0.0	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
M	.0074	.0051	.0032	.0021	.0013	.0008	.0005	.0003	.0002	.0001	.0001
L	.0078	.0054	.0035	.0023	.0015	.0010	.0006	.0004	.0002	.0002	.0001
CY	.0105	.0076	.0052	.0036	.0025	.0017	.0012	.0008	.0005	.0004	.0002
EST	.0139	.0106	.0077	.0056	.0041	.0029	.0021	.0015	.0011	.0008	.0005
SLO	.0166	.0131	.0099	.0075	.0057	.0042	.0031	.0023	.0017	.0012	.0009
LV	.0182	.0146	.0112	.0087	.0066	.0050	.0038	.0029	.0022	.0016	.0012
LT	.0222	.0185	.0148	.0119	.0095	.0075	.0059	.0046	.0036	.0028	.0021
IRL	.0235	.0198	.0161	.0131	.0105	.0084	.0067	.0053	.0042	.0033	.0025
FIN	.0272	.0235	.0197	.0165	.0137	.0113	.0092	.0075	.0060	.0049	.0039
SK	.0277	.0239	.0202	.0170	.0141	.0117	.0095	.0078	.0064	.0051	.0041
DK	.0277	.0240	.0202	.0170	.0142	.0117	.0096	.0078	.0064	.0052	.0041
A	.0339	.0306	.0269	.0235	.0204	.0175	.0149	.0127	.0108	.0091	.0075
S	.0355	.0324	.0288	.0254	.0222	.0193	.0166	.0143	.0122	.0104	.0086
H	.0378	.0349	.0314	.0281	.0249	.0219	.0191	.0166	.0144	.0124	.0104
CZ	.0380	.0351	.0316	.0284	.0251	.0221	.0193	.0168	.0146	.0126	.0106
B	.0381	.0353	.0318	.0286	.0254	.0223	.0195	.0170	.0148	.0128	.0107
P	.0381	.0353	.0319	.0285	.0254	.0223	.0195	.0170	.0148	.0127	.0107
GR	.0392	.0365	.0331	.0299	.0267	.0236	.0207	.0182	.0159	.0138	.0117
NL	.0474	.0460	.0436	.0409	.0380	.0350	.0320	.0292	.0266	.0239	.0209
PL	.0706	.0757	.0797	.0824	.0840	.0847	.0846	.0833	.0810	.0781	.0758
E	.0720	.0775	.0822	.0856	.0876	.0888	.0890	.0882	.0865	.0843	.0825
I	.0831	.0933	.1038	.1126	.1204	.1270	.1326	.1372	.1404	.1430	.1447
GB	.0841	.0949	.1060	.1156	.1240	.1314	.1378	.1429	.1471	.1503	.1526
F	.0842	.0951	.1062	.1158	.1244	.1319	.1382	.1435	.1478	.1511	.1534
D	.0951	.1114	.1311	.1495	.1678	.1860	.2040	.2221	.2405	.2599	.2802

Table 3: Monte Carlo estimates of π_j for EU25.

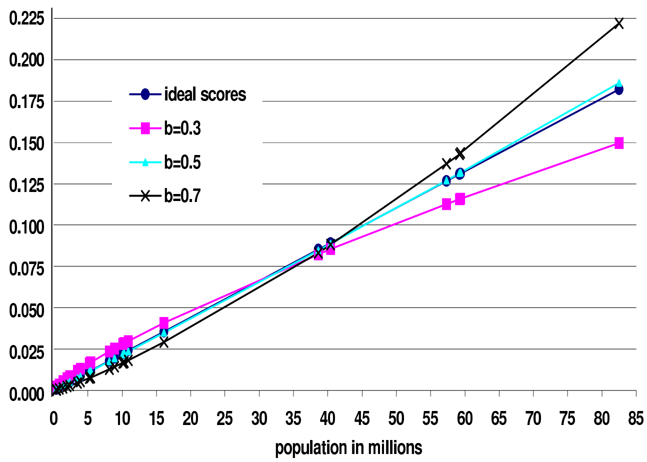

Figure 1: EU25 with weights $w_j = {n_j}^{\alpha}$ compared to ideal probabilities

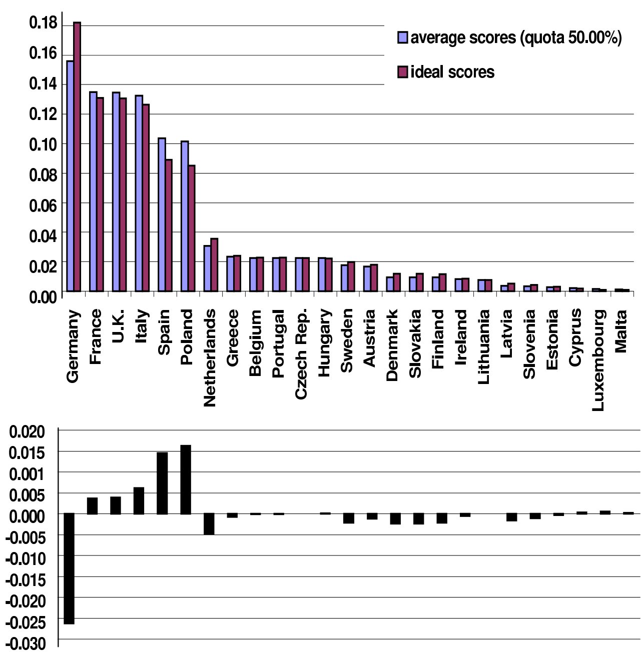

Figure 2: EU25 with Nice weights and hypothetical quota of 50%

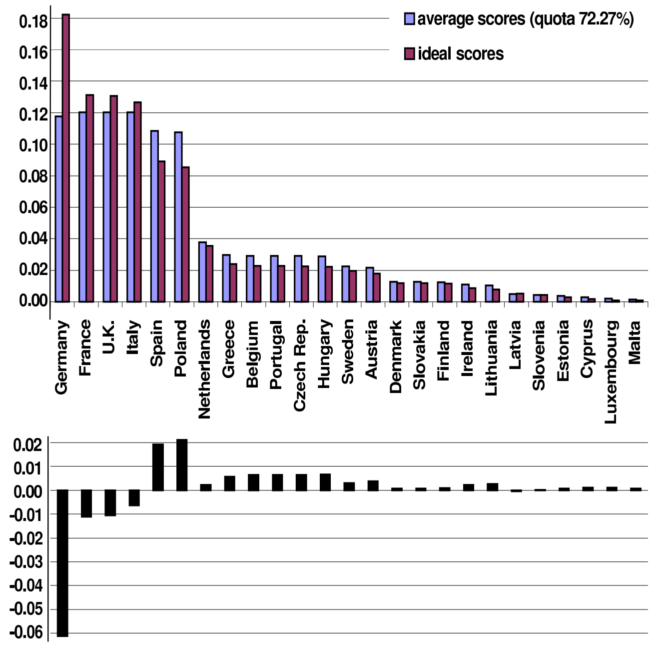

Figure 3: EU25 with Nice weights and 72.2% quota

win et al. 2001). In the [0,1]-spatial voting setting considered in this paper, it also has important implications for equality of representation and hence the legitimacy of decisions.

3.3 US Electoral College

Figure 4 reports the cumulative quadratic deviations from ideal probabilities of US states being pivotal in the *US Electoral College*. Again, $\alpha = 0.5$ does best amongst all power laws for weight assignment. Moreover, as illustrated by Figure 5, the square root rule is extremely successful in ensuring equal representation.

4 Concluding remarks

How should voting weights in an electoral college be assigned to constituencies of different sizes in order to reflect the one-person-one-vote principle? This important question has to the best of our knowledge so far only been addressed for the quite specific case of dichotomous choices. A possible reason for why richer settings such as the – still comparatively simple – one considered in this paper, have been neglected is that they do not seem to allow an easy *analytical* answer. However, our numerical investigation produces straightforward arguments in favor of assigning voting weights proportional to the *square root* of constituencies' populations. For all practical purposes, i. e. ignoring the rather small difference between Penrose-Banzhaf index and voting weight for large number of constituencies, this coincides with Penroses's square root rule.

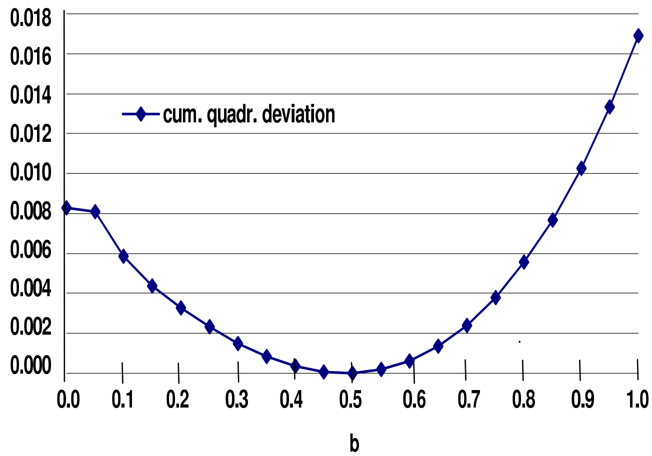

Figure 4: Cumulative quadratic deviation for US Electoral College

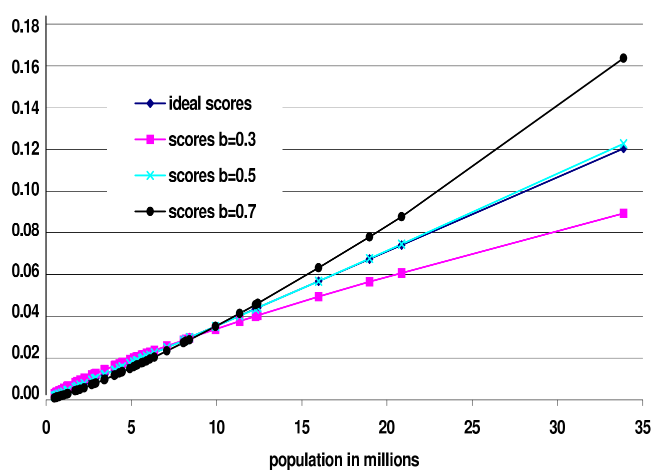

Figure 5: US Electoral College with weights $w_j = {n_j}^{\alpha}$ compared to ideal probabilities

This is a rather surprising result. Different assumptions about the decision environment can a priori be expected to produce different rules for weight assignment, trying to achieve a fixed norm such as equal indirect power of voters. Apart from the 'veil of ignorance' assumption of a priori identical but independent voters, our setting is quite unrelated to the classical 0-1 model. That the square root rule is at least close to optimal in it is comforting: there is a single norm which (approximately) ensures equal representation in two-tier voting systems under very different assumptions.

The robustness of the square root rule can be challenged and has been challenged within the 0-1 setting.[15] That it turns up again in our very distinct decision environment indicates that the square root rule is *not* an artifact of a very particular behavioral model, but rather a robust principle for constitutional design. A caveat to this is that while the quota does not matter for representation in the 0-1 setting, it becomes important for decisions concerning a richer set of alternatives. The square root rule does well when the quota is 50%, but large constituencies need greater weight under qualified majority rule.

References

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

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.

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review 19*(2), 317–343.

Barberà, S. and M. O. Jackson (2004). On the weights of nations: Assigning voting weights in a heterogeneous union. mimeo, CODE, Universitat Autonoma de Barcelona and California Institute of Technology.

Beisbart, C., L. Bovens, and S. Hartmann (2004). A utilitarian assessment of alternative decision rules in the Council of Ministers. mimeo, Dept. of Philosophy, London School of Economics.

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

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

[15]For example, Good and Mayer (1975) and Chamberlain and Rothschild (1981) give up the assumption of equiprobable 'yes' and 'no'-decisions (or alternatively a probability p_i which is the realization of a random variable P_i distributed *independently* across voters with mean $\mu_i = 0.5$) and find that the probability of a voter being pivotal in a given constituency j no longer falls as the square root of n_j. Then, Penrose's square root rule can actually produce quite unequal representation.

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

Gabaix, X. (1999). Zipfs law for cities: An explanation. *Quarterly Journal of Economics 114*, 739–767.

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

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

Leech, D. (2002). Power indices as an aid to institutional design. Warwick Economic Research Papers 646, University of Warwick.

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

Newman, N. (2000). The power of design. *Nature 405*, 412–413.

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

Reed, W. J. (2004). On the rank-size distribution for human settlements. *Journal of Regional Science 42*, 1–17.

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.