

The fate of the square root law for correlated voting

Werner Kirsch* · Jessica Langner†

July 2, 2012

Abstract

We consider two-tier voting system and try to determine optimal weights for a fair representation in such systems. A prominent example of such a voting system is the Council of Ministers of the European Union. Under the assumption of independence of the voters, the square root law gives a fair distribution of power (based on the Penrose-Banzhaf power index) and a fair distribution of weights (based on the concept of the majority deficit), both given in the book by Felsenthal and Machover. In this paper, special emphasis is given to the case of correlated voters. The cooperative behaviour of the voters is modeled by suitable adoptions of spin systems known from statistical physics. Under certain assumptions we are able to compute the optimal weights as well as the average deviation of the council's vote from the public vote which we call the democracy deficit.

Acknowledgement This paper will be presented at the Leverhulme Trust sponsored Voting Power in Practice Symposium held at the London School of Economics, 20–22 March 2011.

1 Introduction

In this paper, we consider two-tier voting systems. The first level of such a systems usually consists of the voters in a country or an association of countries. The voters in each constituency (or member country) are represented by a delegate in the second level voting

*Werner Kirsch, ✉Fakultät für Mathematik und Informatik, FernUniversität Hagen, D-58095 Hagen, Germany, werner.kirsch@fernuni-hagen.de

†Jessica Langner, ✉Fakultät für Mathematik und Informatik, FernUniversität Hagen, D-58095 Hagen, Germany, jessica.langner@fernuni-hagen.de

system, the council. Delegates in the council are given a voting weight which as a rule depends on the population of the constituency they represent.

Examples of such two-tier voting systems are the Council of Ministers of the European Union, the Electoral College in the USA and the ‘Bundesrat’, the state chamber of Germany’s parliamentary system. In each case we assume that the representatives vote according to the majority vote in their respective constituency.

What is a fair voting weight for a delegate in a council? This question arises immediately in all these examples. It seems self-evident that for a fair voting system the voting outcome in the council should agree with the result of a popular vote. The US presidential elections 2000 show that this is not always the case. While Al Gore won the public vote the majority in the Electoral College elected Georg W. Bush as the 43rd president of the USA. The difference between the voting result in the council and the public vote is called the ‘democracy deficit’.

In fact, it is not hard to see, that *no* voting system for the council can *guarantee* that the vote in the council and the public vote agree. In other words, no matter how we choose the voting weights for the council members, the democracy deficit cannot be zero for *all* possible distributions of ‘yes’- and ‘no’- votes among the voters. Thus, the best one can do is to minimize the expected democracy deficit, i. e., the difference between the vote in the council and popular vote. Obviously, the term ‘expected’ needs a careful interpretation. If one assumes that all voters cast their votes independently of each other then one can show that the expected democracy deficit is minimized if the voting weight of a representative is chosen proportional to the *square root* $\sqrt{N_\nu}$ of the population (N_ν) of the respective country (with number ν).

This is (one version of) the celebrated ‘square root law’ by Penrose (see [Felsenthal and Machover 1998] and [Penrose 1946]). In this paper, we go beyond the square root law by dropping the assumption of the voters’ independence. We apply two different schemes to model the correlation between the voters. In our main model we assume that the voters are influenced by a ‘common belief’ of the society or -which is the same, technically speaking- by a strong group of opinion makers. We call this system the CBM (for ‘common belief model’ or ‘collective bias model’) (see [Kirsch 2007]). The CBM can be looked upon as a generalization of a model proposed by Straffin [Straffin 1977] in connection with the Shapley-Shubik power index (see [Shapley and Shubik 1954]). The other model we look at takes into account that voters influence each other. It is based on a model (the Curie-Weiss Model) for ferromagnetic behaviour taken from statistical physics (see [Kirsch 2007] and cf. [Ellis 1985, Thompson 1972]).

If we assume that the voters in different countries vote independently of each other, we can compute the optimal voting weights in terms of the expected margins of the voting outcome in the countries. For the CBM the optimal weights are proportional to the pop-

ulation N_ν . We also compute the expected democracy deficit for these models (for large N_ν).

Under the assumption that the voters influence each other also across country borders (according to the CBM) we can also compute the expected democracy deficit asymptotically. It turns out that in this case any voting weight is as good as any other one. In other words, on an asymptotical scale *any* distribution of voting weights is close to optimal.

2 The General Model

We consider a situation where M states (countries, constituencies) form a federation. The states are labeled by Greek characters, e. g., ν, κ, \dots . The number of voters (population) of the state ν is denoted by N_ν . Consequently, the total population of the union is given by $N = \sum_{\nu=1}^M N_\nu$.

We represent the vote of the voter i in state ν by $X_{\nu i}$. This voter may vote either ‘yes’, in which case we set $X_{\nu i} = 1$ or ‘no’ encoded as $X_{\nu i} = -1$. Consequently, the result of a simple majority voting in the state ν is represented by the sum $S_\nu = \sum_{i=1}^{N_\nu} X_{\nu i}$. A voting in that state is affirmative if $S_\nu > 0$. For the simplicity of notation and to avoid nonsignificant technicalities we assume that all N_ν are odd numbers, this excludes a draw described by $S_\nu = 0$.

We denote the voting decision in the state ν by $\chi_\nu = \chi_\nu(S_\nu)$ which we set equal to 1 if $S_\nu > 0$ and equal to -1 if $S_\nu \leq 0$. Thus, the representative of state ν will vote ‘yes’ if $\chi_\nu = 1$ and ‘no’ if $\chi_\nu = -1$. For later use we note that $\chi_\nu S_\nu = |S_\nu|$.

If we denote the voting weight for state ν in the council by g_ν then the voting result in the council is given by

$$C = \sum_{\nu=1}^M g_\nu \chi_\nu. \quad (1)$$

This voting result has to be compared with the popular vote given by

$$P = \sum_{\nu=1}^M S_\nu. \quad (2)$$

We call the absolute value of the difference between C and P the *democracy deficit* and denote it by Δ

$$\Delta = |C - P| \quad (3)$$

$$= \left| \sum_{\nu=1}^M g_\nu \chi_\nu - \sum_{\nu=1}^M S_\nu \right|. \quad (4)$$

The democracy deficit Δ depends explicitly on the voting weights g_1, \dots, g_M . The voting weights should be chosen in such a way that the democracy deficit is as small as possible.

The voting results $X_{\nu i}$ are the voter's reaction on a particular proposal ω . Hence, the democracy deficit Δ depends on the given proposal ω as well. It is easy to choose the weights g_ν such that Δ vanishes for a *given* proposal. But our goal is to optimize the weights in such a way that Δ is small for *most* proposals. Thus, we look at the *expected* value of Δ^2 , denoted by

$$\mathbb{D} := \mathbb{E}(\Delta^2). \quad (5)$$

We will call \mathbb{D} the *expected democracy deficit* in the following (instead of the correct but clumsy 'expected square of the democracy deficit').

By looking at expectation values we regard the proposals as random input to the voting system. Hence the probability that the next proposal to the system is a particular proposal ω is determined by a probability rule. We assume that there is no bias to certain proposals, in particular any proposal and its counterproposal have the same probability.

The voting system reacts in a deterministic (and rational) way to this random input. The voting results as well as the democracy deficit are therefore (otherwise deterministic) functions of the random input, the proposal. The voting outcome is a vector in the space $\Omega = \{-1, 1\}^N$, where N is the total number of voters and the probability distribution of the proposals equips Ω with probability distribution \mathbb{P} as well, namely the probability of a given outcome (X_1, \dots, X_N) is the probability of all proposals ω that lead to that outcome. Since the voters react rationally they vote -1 on the opposite to a proposal they would favour and vice versa. Hence the probability distribution \mathbb{P} satisfies

$$\mathbb{P}(X_1, \dots, X_N) = \mathbb{P}(-X_1, \dots, -X_N). \quad (6)$$

We call such a measure a *voting measure*. For any voting measure we have $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$, but probabilities concerning more than one voter, like $\mathbb{P}(X_1 = 1 \text{ and } X_2 = 1)$ cannot be computed from the mere assumption that \mathbb{P} is a voting measure. Such events concern the correlation structure of the measure and they have yet to be fixed depending on the situation at hand. One possible specification is the assumption that all voters act independently of each other. This leads to the property that

$$\mathbb{P}(X_1 = 1 \text{ and } X_2 = 1) = \mathbb{P}(X_1 = 1) \cdot \mathbb{P}(X_2 = 1) = \frac{1}{4}.$$

More generally, under the assumption of independence we have

$$\mathbb{P}(X_1 = \xi_1, X_2 = \xi_2, \dots, X_N = \xi_N) = \frac{1}{2^N} \quad (7)$$

for any $\xi_1, \dots, \xi_N \in \{-1, 1\}$. The voting measure describes the mutual influence of the voters on each other, mathematically speaking it describes the correlation structure of the voting system. The above example describes *independent* voters - in some sense the classical case of the theory. An extreme case is given by the measure \mathbb{P}_u

$$\begin{aligned} \mathbb{P}_u(X_1 = 1, X_2 = 1, \dots, X_N = 1) &= \mathbb{P}_u(X_1 = -1, X_2 = -1, \dots, X_N = -1) \\ &= \frac{1}{2}. \end{aligned} \quad (8)$$

For this (rather boring) voting measure the only possible outcomes are the unanimous votes, it represents total (positive) correlation.

If \mathbb{P} is a voting measure, we denote the expectation value with respect to \mathbb{P} by \mathbb{E} , as was already anticipated in (5). Since we assume that the numbers N_ν are odd, it follows that $S_\nu \neq 0$. From this we conclude that $\mathbb{E}(\chi_\nu) = 0$ for any voting measure.

3 Optimal Weights for Independent States

We begin by determining optimal weights, under the assumption that voters in different states are independent. Thus, we assume that the random variables $X_{\nu i}$ and $X_{\kappa j}$ are independent for $\nu \neq \kappa$.

We want to minimize the function

$$\begin{aligned} \mathbb{D}(\gamma_1, \dots, \gamma_M) &= \mathbb{E}(\Delta(\gamma_1, \dots, \gamma_M)^2) \\ &= \sum_{\nu, \kappa=1}^M \left(\gamma_\nu \gamma_\kappa \mathbb{E}(\chi_\nu \chi_\kappa) - 2\gamma_\nu \mathbb{E}(\chi_\nu S_\kappa) + \mathbb{E}(S_\nu S_\kappa) \right). \end{aligned} \quad (9)$$

The function $\mathbb{D}(\gamma_1, \dots, \gamma_M)$ is a measure for the expected democracy deficit for voting weights $\gamma_1, \dots, \gamma_M$.

By the assumption of independent *states* we can conclude that

$$\mathbb{E}(\chi_\nu \chi_\kappa) = \mathbb{E}(\chi_\nu) \mathbb{E}(\chi_\kappa) = 0 \quad \text{for } \nu \neq \kappa, \quad (10)$$

$$\mathbb{E}(\chi_\nu S_\kappa) = \mathbb{E}(\chi_\nu) \mathbb{E}(S_\kappa) = 0 \quad \text{for } \nu \neq \kappa, \quad (11)$$

and

$$\mathbb{E}(S_\nu S_\kappa) = \mathbb{E}(S_\nu) \mathbb{E}(S_\kappa) = 0 \quad \text{for } \nu \neq \kappa. \quad (12)$$

Moreover, we have $\chi_\nu^2 = 1$ and $\chi_\nu S_\nu = |S_\nu|$, thus

$$\mathbb{D}(\gamma_1, \dots, \gamma_M) = \sum_{\nu=1}^M \left(\gamma_\nu^2 - 2\gamma_\nu \mathbb{E}(|S_\nu|) + \mathbb{E}(S_\nu^2) \right). \quad (13)$$

It is not hard to find the minimizing weights g_ν (by the usual procedure: Find the zeros of the derivative), in fact: The weights g_1, \dots, g_M which minimize the function \mathbb{D} are given by

$$g_\nu = \mathbb{E}(|S_\nu|). \quad (14)$$

This result has a very intuitive interpretation. The quantity S_ν is the difference between the ‘yes’-votes and the ‘no’-votes, so $|S_\nu|$ describes the margin of the voting outcome, i. e., the surplus of votes of the winning party. Therefore, the optimal weights g_ν for the state ν are given by the expected margin of a vote in that state. In fact, the delegate of state ν does not represent the opinion of *all* voters in this state, but only those who agree with the majority, he or she acts against the will of the minority, so as a net result the delegate just represents the margin.

We can also compute the expected democracy deficit \mathbb{D} for the optimal weights g_1, \dots, g_M

$$\mathbb{D}(g_1, \dots, g_M) = \sum_{\nu=1}^M \left(\mathbb{E}(|S_\nu|^2) - \mathbb{E}(|S_\nu|)^2 \right) = \sum_{\nu=1}^M \mathbb{V}(|S_\nu|) \quad (15)$$

where $\mathbb{V}(|S_\nu|)$ denotes the variance of the random quantity $|S_\nu|$.

We emphasize that we did not yet make assumptions about the correlation structure of voters *inside* a country. Of course, the numerical evaluation of the optimal weights and minimal democracy deficit requires further assumptions on the correlation between voters.

4 Independent Voters

In this section we assume that all voters act independently of each other, in mathematical terms: all random variables $X_{\nu i}$ are independent of each other. Under this assumption we can compute the optimal weight $g_\nu = \mathbb{E}(|S_\nu|)$ as well as the minimal expected democracy deficit.

For the independent random variables $X_{\nu i}$ we have the central limit theorem, namely the weighted sums

$$\frac{1}{\sqrt{N_\nu}} S_\nu := \frac{1}{\sqrt{N_\nu}} \sum_{i=1}^{N_\nu} X_{\nu i} \quad (16)$$

are asymptotically distributed for large N_ν according to a standard normal distribution (cf. Lamperti [Lamperti 1996]). From this it follows that for large N_ν

$$\mathbb{E}(|S_\nu|) \approx \frac{\sqrt{2}}{\sqrt{\pi}} \sqrt{N_\nu}, \quad (17)$$

$$\mathbb{E}(|S_\nu|^2) \approx \sqrt{N_\nu}, \quad (18)$$

and

$$\mathbb{V}(|S_\nu|) \approx \frac{\pi - 2}{\pi} N_\nu. \quad (19)$$

We conclude that the optimal weight for independent voters is proportional to the square root of the population. This is exactly the content of the square root law by Penrose (see [Penrose 1946] and [Felsenthal and Machover 1998]).

The above formulae also allow us to evaluate the minimum of the expected democracy deficit

$$\mathbb{D}(g_1, \dots, g_M) \approx \frac{\pi - 2}{\pi} N. \quad (20)$$

This implies that the *expected democracy deficit per voter*, namely

$$\mathbb{E} \left(\left(\frac{\Delta}{N} \right)^2 \right) \quad (21)$$

converges to zero as N becomes large (with convergence rate $\frac{1}{N}$).

5 The Collective Bias Model

Now, we introduce and discuss a model for collective behaviour of voters. The basic idea is that there is a mainstream opinion, e. g., a common belief due to the country's tradition or the influence of opinion makers. For a given proposal ω we model this 'common belief' by a value $\zeta \in [-1, 1]$ which depends on the proposal at hand. The value $\zeta = 1$ means there is such a strong common belief in favor of the proposal that all voters will vote 'yes', $\zeta = -1$ means all voters will vote 'no'. In general, ζ denotes the expected outcome of the voting, i. e., $\mathbb{E}(X_{\nu i})$. The voting results $X_{\nu i}$ themselves fluctuate around this value randomly.

Let us be more precise about this. Suppose the voting results are X_1, \dots, X_N (where we dropped the index ν for notational simplicity). Let μ be a measure on $[-1, 1]$, which

is the distribution of the common belief value ζ , that is $\mu(\cdot|a, b)$ is the probability that the value ζ is between a and b . Let P_ζ be the probability measure on $\{-1, 1\}$ with

$$P_\zeta(X_1 = 1) = p_\zeta = \frac{1}{2}(1 + \zeta),$$

so that

$$E_\zeta(X_1) := P_\zeta(X_1 = 1) - P_\zeta(X_1 = -1) = p_\zeta - (1 - p_\zeta) = \zeta.$$

For a given value of ζ we set

$$\mathcal{P}_\zeta(\xi_1, \dots, \xi_N) = \prod_{i=1}^N P_\zeta(\xi_i). \quad (22)$$

For any $\zeta \in [-1, 1]$ the expression \mathcal{P}_ζ is a probability distribution on $\Omega = \{-1, 1\}^N$. We define the collective bias measure \mathbb{P}_μ with respect to μ as

$$\mathbb{P}_\mu(X_1 = \xi_1, \dots, X_N = \xi_N) := \int \mathcal{P}_\zeta(\xi_1, \dots, \xi_N) d\mu(\zeta). \quad (23)$$

Note, that \mathcal{P}_ζ is *not* a voting measure (unless $\zeta = \frac{1}{2}$). However \mathbb{P}_μ is a voting measure if μ is invariant under sign change, i. e., $\mu(\cdot|a, b) = \mu(\cdot|-b, -a)$. We call μ the *bias measure*.

If the measure μ is concentrated in 0, then \mathbb{P}_μ makes the voting results X_i independent, thus we are in the case of section 4. If μ is the uniform distribution on $[-1, 1]$ (that is every point is equally likely), then the corresponding measure was already considered by Straffin [Straffin 1977] where he established an intimate connection of this model to the Shapley-Shubik index. In a similar way, the Penrose-Banzhaf measure is connected with the model of independent voters.

The CBM can be looked upon as a model for spins in statistical mechanics. There the voters are replaced with elementary magnets (spins) which can be directed upwards ($X_i = 1$) or downwards ($X_i = -1$). In this language the Collective Bias Model describes spins which do not interact with each other but are influenced by an exterior magnetic field, namely the collective bias ζ .

In the papers [Kirsch 2007], [Kirsch and Langner 2012] and [Langner 2012] we investigate also another model for collective voting behaviour which comes directly from statistical physics, the Curie-Weiss Model (CWM). In this model the spins (voters) influence each other by an interaction which makes spins to prefer to be directed parallel to the others. For voting this means that voters prefer to agree to the other voters. The Curie-Weiss Model is a very interesting tool to investigate collective behaviour. However, it is

technically more involved than the other models we discuss. Therefore, we will mention it only rather briefly and refer to the papers mentioned above for more details.

Let us define

$$H(X_1, \dots, X_N) = -\frac{1}{N} \left(\sum_{i=1}^N X_i \right)^2. \quad (24)$$

This is the *energy function* for the spin configuration X_1, \dots, X_N . We use this to define measures

$$Q_\beta(X_1, \dots, X_N) = e^{-\beta H(X_1, \dots, X_N)} \quad (25)$$

where $\beta \in]0, \infty[$ is the inverse temperature in statistical physics. As a rule, Q_β is not a probability measure, so we normalize it by dividing through its total mass Z and set

$$P_\beta(X_1, \dots, X_N) = \frac{e^{-\beta H(X_1, \dots, X_N)}}{Z}. \quad (26)$$

This is the Curie-Weiss measure for inverse temperature β . The parameter β measures the strength of the interaction between the voters. The extreme case $\beta = 0$ corresponds to the model of independent voters, the other extreme $\beta = \infty$ describes the case of the measure \mathbb{P}_u defined in (8) for unanimous voting.

6 Optimal weights for the Collective Bias Model

Let us now suppose that voters in different countries are independent, but voting inside the countries follows the CBM with bias measure μ . According to section 3 in this case the optimal weights are given by

$$g_\nu = \mathbb{E}_\mu(|S_\nu|). \quad (27)$$

For large N_ν we have

$$g_\nu = \mathbb{E}_\mu(|S_\nu|) = \mu_1 N_\nu \quad (28)$$

where $\mu_1 = \int |\zeta| d\mu(\zeta)$ is the first absolute moment of μ . Note, that for any probability measure μ the quantity μ_1 is non zero, except for the case $\mu = \delta_0$, the measure is concentrated at the point 0. This means that the optimal weights for a council are proportional to the population of the respective country if the voters can be described by a CBM. This also includes the Straffin case (μ is the uniform distribution), which corresponds to the Shapley-Shubik power index.

The only exception from proportionality is the case $\mu = \delta_0$ corresponding to independent voting (the Penrose-Banzhaf case), where the square root law applies.

We mention that there is a ‘phase transition’ for the Curie-Weiss Model if we vary β from 0 to ∞ , namely

$$g_\nu = \mathbb{E}_\beta(|S_\nu|) = \begin{cases} \frac{\sqrt{2}}{\sqrt{\pi}\sqrt{1-\beta}} \sqrt{N_\nu}, & \text{for } \beta < 1; \\ C N_\nu^{\frac{3}{4}}, & \text{for } \beta = 1; \\ C(\beta) N_\nu, & \text{for } \beta > 1. \end{cases} \quad (29)$$

The constant $C(\beta)$ converges to 0 as $\beta \searrow 1$ and to 1 as $\beta \nearrow \infty$.

7 Democracy Deficit for the Collective Bias Model

Given the optimal weights (28) for the CBM (and independent states) we can compute (the asymptotic behaviour of) the expected democracy Deficit \mathbb{D}_μ

$$\mathbb{D}_\mu = \sum_{\nu=1}^M \mathbb{V}(|S_\nu|) \approx (\mu_2 - \mu_1^2) N^2 \quad (30)$$

where $\mu_1 = \int |\zeta| d\mu(\zeta)$ and $\mu_2 = \int |\zeta|^2 d\mu(\zeta)$. Note that $\mu_2 - \mu_1^2 \neq 0$ unless μ is concentrated in at most two points. It follows that the expected democracy deficit per voter, i. e.,

$$\mathbb{E}_\mu \left(\left(\frac{\Delta}{N} \right)^2 \right)$$

converges to a *positive* constant as the N_ν tend to infinity (in a uniform way, i. e., $N_\nu = \alpha_\nu N$).

It is interesting to remark that the expected democracy deficit per voter converges also to a constant if we choose a *non optimal* voting weight, like for instance $g_\nu \sim \sqrt{N_\nu}$ or $g_\nu = 1$ for all ν . This constant will in general be larger than the one for the optimal weights, but the order of magnitude of \mathbb{D} is *not* changed.

For the Curie-Weiss Model the expected democracy deficit per voter converges to zero (for $\beta \neq 1$ even with rate $\frac{1}{N}$).

8 A Model with Global Collective Behaviour

So far we have always assumed that voter in different states act independently. In this section we consider the case of collective behaviour across country borders. We assume

that *all* voters act according to the Collective Bias measure \mathbb{P}_μ . This means there is a common belief, expressed through the measure μ , for all voters in the union.

Then, the formulae (10) – (13) are no longer valid. In fact, determining the optimal voting weights requires to solve a rather complicated system of M dependent linear equations. Instead of doing this we try to look at the democracy deficit directly. It turns out that for large N_ν we have for any ν, κ

$$\mathbb{E}_\mu(\chi_\nu \chi_\kappa) \approx 1, \quad (31)$$

$$\mathbb{E}_\mu(\chi_\nu S_\kappa) \approx \mathbb{E}_\mu(|S_\kappa|) \approx \mu_1 N_\kappa \quad (32)$$

and

$$\mathbb{E}(S_\nu S_\kappa) \approx \mu_2 N_\nu N_\kappa. \quad (33)$$

Inserting these terms into the expression for \mathbb{D} we obtain

$$\begin{aligned} \mathbb{D}(g_1, \dots, g_M) &= \sum_{\nu, \kappa=1}^M \mathbb{E}_\mu(\chi_\nu \chi_\kappa) g_\nu g_\kappa - 2 \sum_{\nu=1}^M g_\nu \sum_{\kappa=1}^M \mathbb{E}_\mu(\chi_\nu S_\kappa) + \sum_{\nu, \kappa=1}^M \mathbb{E}_\mu(S_\nu S_\kappa) \\ &\approx \sum_{\nu, \kappa=1}^M g_\nu g_\kappa - 2 \sum_{\nu=1}^M g_\nu \sum_{\kappa=1}^M \mu_1 N_\kappa + \sum_{\nu, \kappa=1}^M \mu_2 N_\nu N_\kappa \\ &= \left(\sum_{\nu=1}^M g_\nu \right)^2 - 2\mu_1 \left(\sum_{\nu=1}^M g_\nu \right) N + \mu_2 N^2 \\ &= G^2 - 2\mu_1 G + \mu_2 N^2. \end{aligned} \quad (34)$$

This last expression depends only on the sum $G = \sum_{\nu=1}^M g_\nu$ of the voting weights and *not* on the single weight g_ν . This means that for large N_ν the asymptotic value of \mathbb{D} does not depend on the way the weights are distributed among the member states of the union. The minimal value of \mathbb{D} is obtained by choosing $G = \mu_1 N$ independent of the values of the particular weight g_ν . We also note that the value of G has no real meaning, since we don't change the voting system at all if we multiply all weights (and the quota) with the same number $C > 0$.

Finally, we remark that the somewhat hand waving arguments in (34) need a careful mathematical interpretation. A precise formulation gives:

$$\lim_{N \rightarrow \infty} \mathbb{E}_\mu \left(\left(\frac{\Delta(g_1, \dots, g_M)}{N} \right)^2 \right) = \mu_2 - \mu_1^2 \quad (35)$$

for $G = \sum_{\nu=1}^M g_{\nu} = \mu_1 N$ and

$$\liminf_{N \rightarrow \infty} \mathbb{E}_{\mu} \left(\left(\frac{\Delta(g_1, \dots, g_M)}{N} \right)^2 \right) \geq \mu_2 - \mu_1^2 \quad (36)$$

for any arbitrary choice of g_{ν} . This result can be interpreted in the following way: If there is a strong common belief in the union across border lines then it doesn't matter how one distributes the voting weights in the council.

References

- [Banzhaf 1965] Banzhaf JF (1965) Weighted voting doesn't work: a mathematical analysis. *Rutgers Law Review* 19: 317–343.
- [Ellis 1985] Ellis R (1985) Entropy, large deviations and statistical mechanics. *Grundlehren der mathematischen Wissenschaften*, vol. 271. Berlin Heidelberg New York: Springer.
- [Felsenthal and Machover 1995] Felsenthal DS, Machover M (1995) Postulates and paradoxes of relative voting power - A critical re-appraisal. *Theory and Decision* 38 (2): 195–229.
- [Felsenthal and Machover 1998] Felsenthal DS, Machover M (1998) The measurement of voting power: Theory and practice, problems and paradoxes. Edward Elgar, Cheltenham, forthcoming.
- [Kirsch 2004] Kirsch W (2004) What is a Fair Distribution of Power in the Council of Ministers of the EU? http://www.ceps.be/Article.php?article_id=360.
- [Kirsch 2007] Kirsch W (2007) On Penrose's Squareroot Law and Beyond. *Homo Oeconomicus* 24 (3,4): 357–380.
- [Kirsch and Langner 2012] Kirsch W, Langner J (2012) Mathematical Theory of Correlated Voting, in preparation.
- [Lamperti 1996] Lamperti JW (1996) *Probability: A Survey of the Mathematical Theory*, 2nd Edition, Wiley.
- [Langner 2012] Langner J (2012) PhD-thesis, in preparation.
- [Penrose 1946] Penrose LS (1946) The elementary statistics of majority voting. *Journal of the Royal Statistical Society* 109: 53–57.
- [Shapley and Shubik 1954] Shapley LS, Shubik M (1954) A Method for Evaluating the Distribution of Power in a Committee System. *American Political Science Review* 48: 787–792.
- [Straffin 1977] Straffin PD (1977) Homogeneity, Independence, and Power Indices. *Public Choice* 30: 107–118.
- [Thompson 1972] Thompson C (1972) *Mathematical Statistical Mechanics*. Princeton University Press.