Homo Oeconomicus 24(3/4): 381–399 • (2007)

www.accedoverlag.de

# From a Toy Model to the Double Square Root Voting System

Wojciech Słomczyński
Institute of Mathematics, Jagiellonian University, Kraków, Poland
(eMail: Wojciech.Slomczynski@im.uj.edu.pl)

Karol Życzkowski
Institute of Physics, Jagiellonian University, Kraków, Poland
Center for Theoretical Physics, Polish Academy of Sciences, Poland
(eMail: karol@tatry.if.uj.edu.pl)

**Abstract** We investigate systems of indirect voting based on the law of Penrose, in which each representative in the voting body receives the number of votes (voting weight) proportional to the square root of the population he or she represents. For a generic population distribution, the quota required for the qualified majority can be set in such a way that the voting power of any state is proportional to its weight. For a specific distribution of population the optimal quota has to be computed numerically. We analyse a toy voting model for which the optimal quota can be estimated analytically as a function of the number of members of the voting body. This result, combined with the normal approximation technique, allows us to design a simple, efficient, and flexible voting system, which can be easily adopted for varying weights and number of players.

*Keywords* power indices, weighted voting games, optimal quota, Penrose square root law, normal approximation

## 1. Introduction

A game theory approach proved to be useful to analyse voting rules implemented by various political or economic bodies. Since the pioneering contributions of Lionel Penrose (1946) who originated the mathematical theory of voting power just after the World War II, this subject has been studied by a number of researchers, see, e.g. Felsenthal and Machover (1998, 2004a) and


ISBN 978-3-89265-66-7 ISSN 0943-0180

references therein.

Although the current scientific literature contains several competing definitions of voting indices, which quantitatively measure the voting power of each member of the voting body, one often uses the original concept of Penrose. The *a priori voting power* in his approach is proportional to the probability that a vote cast by a given player in a hypothetical ballot will be decisive: should this country decide to change its vote, the winning coalition would fail to satisfy the qualified majority condition. Without any further information about the voting body it is natural to assume that all potential coalitions are equally likely. This very assumption leads to the concept of *Penrose-Banzhaf index* (*PBI*) called so after John Banzhaf (1965), who introduced this index independently.

Recent research on voting power was partially stimulated by the political debate on the voting system used in the Council of Ministers of the European Union (EU). The *double majority* system endorsed in 2004 by The Treaty Establishing a Constitution for Europe, based on 'per capita' and 'per state' criteria, was criticized by several authors (e.g. Paterson and Silárszky (2003), Baldwin and Widgrén (2004), Bilbao (2004), Bobay (2004), Cameron (2004), Kirsch (2004), Plechanovová (2004), Życzkowski and Słomczyński (2004), Plechanovová (2006), Taagepera and Hosli (2006), Algaba et al. (2007)), who pointed out that it is favourable to the most and to the least populated EU countries at the expense of all medium size states. Ironically, a similar conclusion follows from a book written fifty years earlier by Penrose, who also discovered this drawback of a 'double majority' system.[1]

In search for an optimal two-tier voting system (where a set of constituencies of various size elect one delegate each to a decision-making body) Penrose (1946) considered first a direct election in a state consisting of $N$ voters and proved that the voting power of a single citizen decays as $1/\sqrt{N}$, provided that the votes are uncorrelated. To compensate this effect he suggested that the a priori *voting power* of each representative in the voting body should behave proportionally to $\sqrt{N}$ making the citizens' voting power in all states equal and so the whole system *representative* (the *Penrose square root law*).

Systems, where the *voting weight* of each state is proportional to the square root of its population were discussed by several authors in the EU context, see Laruelle and Widgrén (1998), Hosli (2000), Tiilikainen and Widgrén (2000), Felsenthal and Machover (2001, 2002), Laruelle and Valenciano (2002), Moberg (2002), Mabille (2003), Widgrén (2003), College of Europe (2004), Felsenthal and Machover (2004b), Hosli and Machover (2004), Plechanovová (2004), Widgrén (2004). Different experts have proposed different quotas for a square root voting systems, usually varying from

[1] Penrose (1952) wrote: '[...] if two votings were required for every decision, one on a *per capita* basis and the other upon the basis of a single vote for each country. This system [...] would be inaccurate in that it would tend to favour large countries.'

60% to 74%. Clearly, the choice of an appropriate decision-taking quota (threshold) $q$ affects both the distribution of voting power in the Council (and thus also the representativeness of the system) and the voting system's efficiency and transparency.

However, the assertion that the voting weight of each country should be proportional to the square root of its population does not entirely solve the problem of distributing the power. Kirsch (2004) expressed this as follows: 'The square root law tells us how the power should be distributed among the countries. It is, however not clear at a first glance how to implement it in terms of voting weights, as the voting weights do not give the power indices immediately'. Accordingly, the question arise: how to solve the *inverse problem*, i.e. how to allocate weights and how to set *quota* (threshold) for qualified majority (the Council reaches a decision when the sum of the weights of the Member States voting in favour exceeds the threshold) to obtain required distribution of power, see Laruelle and Widgrén (1998), Sutter (2000), Leech (2002), Lindner and Machover (2004), Widgrén (2004), Pajala (2005), Aziz et al. (2007).

The answer we proposed in (Życzkowski and Słomczyński (2004)) is surprisingly simple: one should choose the weights to be also proportional to the square root of the population and then find such an optimal quota $q_*$ that would produce the maximally *transparent* system, that is a system under which the voting power of each Member State would be approximately equal to its voting weight, or more precisely, the mean discrepancy $\Delta$ between the voting power of each state and the rescaled root of its population would be minimal. Then the Penrose law would be practically fulfilled, and the potential influence of every citizen of each Member State on the decisions taken in the Council would be almost the same.

For a concrete distribution of population in the EU consisting of 25 (resp. 27) member states it was found in (Życzkowski and Słomczyński(2004), Życzkowski et al. (2006), see also Feix et al. (2007)) that the discrepancy exhibits a sharp minimum around a critical quota $q_* \sim 62\,\%$ (resp. $61.5\,\%$) falling down to a negligible value. Therefore, the Penrose square root system with this quota is optimal, in the sense that every citizen in each member state of the Union has the same voting power (measured by the Penrose-Banzhaf index), i.e. the same influence on the decisions taken by the European Council. Such a voting system occurs to give a larger voting power to the largest EU states than the Treaty of Nice but smaller than the draft European Constitution, and thus has been dubbed by the media as the 'Jagiellonian Compromise'.

The existence of such a critical quota $q_*$ for which the rescaled PBIs of all states are approximately equal to their voting weights, is not restricted to this particular distribution of population in the EU. On the contrary, it seems to be a rather generic behaviour which was found by means of numerical simulations for typical random distributions of weights in the voting body generated with respect to various probability measures, see Życzkowski and Słomczyński (2004), Chang et al. (2006), Słomczyński and Życzkowski (2006). The value of $q_*$ depends to some extent on a given realization of the random population distribution, but more importantly, it varies considerably with the number $M$ of the member states. In the limit $M \to \infty$ the optimal quota seems to tend to 50%, in consistence with the so called *Penrose limit theorem* (see Lindner (2004), Lindner and Machover (2004)), which claims that for the quota 50% the relative power of two voters tends asymptotically to their relative voting weight.

Working with random probability distributions it becomes difficult to get any analytical prediction concerning the functional dependence of $q_*$ on the number $M$ of the members of the voting body. Therefore in this work we propose a toy model in which an analytical approach is feasible. We compute the PBIs for this model distribution of population consisting of $M$ states and evaluate the discrepancy $\Delta$ as a function of the quota $q$. The optimal quota $q_*$ is defined as the value at which the quantity $\Delta$ achieves its minimum. This reasoning performed for an arbitrary number of states $M$ allows us to derive an explicit dependence of the optimal quota on $M$. Results obtained analytically for this particular model occur to be close to these received earlier in numerical experiments for random samples. The normal approximation of the number of votes achieved by all possible coalitions provides another estimate of the optimal quota as a function of the quadratic mean of all the weights. The efficiency of voting systems with optimal quota does not decrease when the number of players $M$ increases.

Applying these results we are tempted to design a simple scheme of indirect voting (the *double square root voting systems*) based on the square root law of Penrose supplemented by a rule setting the approximate value of the optimal quota $q_*$ as a function either of the square root of the number of players or the square root of the sum of their weights. Such systems are representative, transparent and efficient.

This work is organized as follows. In Sect. 2 we recall the definition of Penrose-Banzhaf index and define the optimal quota. Sect. 3 provides a description of the toy model of voting in which one player is $c$ times stronger than all other players. We describe the dependence of the optimal quota in this model on the number of voters for $c = 2$ and $c = 3$. In Sect. 4 we discuss the optimal quota applying an alternative technique of normal approximation. The paper is concluded in Sect. 5, where we design a complete voting system. The heuristic proof of the validity of the normal approximation method is given in the Appendix.

## 2. A priori voting power and critical quota

Consider a set of $M$ members of the voting body, each representing a state with population $N_k$, $k = 1, \ldots, M$. Let us denote by $w_k$ the voting weight attributed to the niedziela, marzec 30, 2008 at 1:29 pm$k$-th representative. We work with renormalised quantities, so that $\sum_{i=1}^{M} w_i = 1$, and we assume that the decision of the voting body is taken if the sum of the weights of all members of the coalition exceeds the given *quota* $q \in [0.5, 1]$, i.e. we consider so called (*canonical*) *weighted majority voting game* $[q; w_1, \ldots, w_M]$, see Felsenthal and Machover (1998).

To analyse the voting power of each member one has to consider all $2^M$ possible coalitions and find out the number $\omega$ of winning coalitions which satisfy the qualified majority rule adopted. The quantity $A := \omega/2^M$ measures the *decision-making efficiency* of the voting body, i.e. the probability that it would approve a randomly selected issue. Coleman (1971) called this quantity the *power of a collectivity to act*. For a thorough discussion of this concept, see Lindner (2004).

The *absolute* (or *probabilistic*) *Penrose–Banzhaf index* (*PBI*) $\psi_k$ of the $k$–th state is defined as the probability that a vote cast by $k$–th representative is decisive. This happens if $k$ is a *critical voter* in a coalition, i.e. the winning coalition with $k$ ceases to fulfil the majority requirements without $k$. Assuming that all $2^M$ coalitions are equally likely, we see that the PBI of the $k$–th state depends only on the number $\omega_k$ of winning coalitions that include this state. Namely, the number $\eta_k$ of coalitions where a vote of $k$ is decisive is given by:

$$\eta_k = \omega_k - (\omega - \omega_k) = 2\omega_k - \omega \, . \tag{1}$$

Moreover, the absolute Penrose-Banzhaf index of the $k$–th state is equal to $\psi_k = \eta_k/2^{M-1}$. To compare these indices for decision bodies consisting of different number of players, it is convenient to define the *normalised PBIs*:

$$\beta_k := \frac{\psi_k}{\sum_{i=1}^{M} \psi_i} = \frac{\eta_k}{\sum_{i=1}^{M} \eta_i} \tag{2}$$

($k = 1, \ldots, M$) fulfilling $\sum_{i=1}^{M} \beta_i = 1$.

In the *Penrose voting system* one sets the voting weights proportional to the square root of the population of each state, i.e. $w_k := \sqrt{N_k} / \sum_{i=1}^{M} \sqrt{N_i}$ for $k = 1, \ldots, M$. For any level of the quota $q$ one may compute numerically the power indices $\beta_k$. The Penrose rule would hold perfectly if the voting power of each state was proportional to the square root of its population. Hence, to quantify the overall representativeness of the voting system one can use the *mean discrepancy* $\Delta$, defined by

$$\Delta := \sqrt{\frac{1}{M}\sum_{i=1}^{M}(\beta_i - w_i)^2} \,. \tag{3}$$

The *optimal quota* $q_*$ is defined as the quota for which the mean discrepancy $\Delta$ is minimal. Note that this quota is not unique and usually there is a whole interval of optimal points. However, the length of this interval decreases with increasing number of voters.

Studying the problem for a concrete distribution of population in the European Union, as well as using a statistical approach and analyzing several random distributions of population we found (Życzkowski and Słomczyński (2004), Życzkowski et al.(2006)) that in these cases all $M$ ratios $\beta_k/w_k$ ($k = 1,\dots,M$), plotted as a function of the quota $q$, cross approximately near a single point $q_*$, i.e.

$$\beta_k\,(q_*) \approx w_k\,(q_*) \tag{4}$$

for $k = 1,\dots,M$. In other words, the discrepancy $\Delta$ at this *critical quota* $q_*$ is negligible. The existence of the critical quota was confirmed numerically in a recent study by Chang et al. (2006). (This does not contradict the fact that there is a wide range of quotas, where the mean discrepancy is relatively small, see Widgrén (2004), Pajala (2005).) In the next section we propose a toy model for which a rigorous analysis of this numerical observation is possible.

## 3. Toy model

Consider a voting body of $M$ members and denote by $w_k$, $k = 1,\dots,M$ their normalized voting weights. Assume now that a single *large* player with weight $w_L := w_1$ is the strongest one, while remaining $m := M-1$ players have equal weights $w_S := w_2 = \cdots = w_M = (1-w_L)/m$. We may assume that $w_L \le 1/2$, since in the opposite case, for some values of $q$, the strongest player would become a 'dictator' and his relative voting power would be equal to unity. Furthermore, we assume that the number of *small* players $m$ is larger than two, and we introduce a parameter $c := w_L/w_S$ which quantifies the difference between the large player and the other players. Thus we consider the weighted voting game $\left[q; \frac{c}{m+c}, \frac{1}{m+c}, \dots, \frac{1}{m+c}\right]$, where the population distribution is characterized by only two independent parameters, say, the number of players $M$ and the ratio $c$. Sometimes it is convenient to use as a parameter of the model the weight $w_L$, which is related with the ratio $c$ by the formula $c = mw_L/(1-w_L)$. On the other hand, the qualified majority quota $q$, which determines the voting system, is treated as a free parameter and will be optimized to minimize the discrepancy (3). Note that a similar model has been analysed in Dubey and Shapley (1978), and Merrill (1982).

To avoid odd-even oscillations in the discrepancy $\Delta(q)$ we assume that $c \geq 2$. To compute the PBIs of all the players we need to analyse three kinds of possible winning coalitions. The vote of the large player is decisive if he forms a coalition with $k$ of his colleagues, where $k < mq/(1-w_L)$ and $k \geq m(q-w_L)/(1-w_L)$. Using the notion of the *roof*, i.e. the smallest natural number larger than or equal to $x$, written as $\lceil x \rceil := \min\{n \in \mathbb{N} : n \geq x\}$, we may put

$$j_1 := \left\lceil \frac{m(q-w_L)}{1-w_L} \right\rceil - 1 \tag{5}$$

and

$$j_2 := \left\lceil \frac{mq}{1-w_L} \right\rceil - 1\,, \tag{6}$$

and recast the above conditions into the form

$$j_1 + 1 \leq k \leq j_2\,. \tag{7}$$

On the other hand, there exist two cases where the vote of a small player is decisive. He may form a coalition with $j_2$ other small players, or, alternatively, he may form a coalition with the large player and $j_1$ small players.

With these numbers at hand, we may write down the absolute Penrose–Banzhaf indices for both players. The a priori voting power of the larger player can be expressed in terms of binomial symbols:

$$\psi_L := \psi_1 = 2^{-m} \sum_{k=j_1+1}^{j_2} \binom{m}{k}\,, \tag{8}$$

while the voting power for all the small players is equal and reads:

$$\psi_S := \psi_2 = \cdots = \psi_M = 2^{-m} \left[ \binom{m-1}{j_1} + \binom{m-1}{j_2} \right]. \tag{9}$$

It is now straightforward to renormalize the above results according to (2) and use the normalized indices $\beta_L$ and $\beta_S$ to write an explicit expression for the discrepancy (3), which depends on the quota $q$. Searching for an 'ideal' system we want to minimize the discrepancy

$$\begin{aligned}\Delta(q) &= \frac{1}{\sqrt{M}}\sqrt{\left(\beta_L - w_L\right)^2 + m\left(\beta_S - w_S\right)^2} \\ &= \frac{1}{\sqrt{M}}\sqrt{\left(\beta_L - w_L\right)^2 + m\left(\frac{1-\beta_L}{m} - \frac{1-w_L}{m}\right)^2} \qquad (10)\\ &= \sqrt{\frac{1+1/m}{M}}\left|\beta_L - w_L\right| \qquad (11)\\ &= \frac{1}{\sqrt{m}}\left|\beta_L - \frac{c}{m+c}\right| \\ &= \frac{1}{\sqrt{m}}\left|\frac{\sum_{k=j_1+1}^{j_2}\binom{m}{k}}{\sum_{k=j_1+1}^{j_2}\binom{m}{k} + m\left(\binom{m-1}{j_1}+\binom{m-1}{j_2}\right)} - \frac{c}{m+c}\right| \\ &= \frac{1}{\sqrt{m}}\left|\frac{\sum_{k=\lceil d-c\rceil}^{\lceil d\rceil-1}\binom{m}{k}}{\sum_{k=\lceil d-c\rceil}^{\lceil d\rceil-1}\binom{m}{k} + m\left(\binom{m-1}{\lceil d-c\rceil-1}+\binom{m-1}{\lceil d\rceil-1}\right)} - \frac{c}{m+c}\right|, \qquad (12)\end{aligned}$$

where $d := mq/\left(1-w_L\right) = (m+c)\,q$.

In theory, one may try to solve this problem looking first for the optimal $d$ and then computing the optimal quota $q_*$, but due to the roof in the bounds of the sum the general case is not easy to work with.

The problem simplifies significantly if we set $c = 2$, considering the $M$–point weight vector $(w_L, w_L/2, \ldots, w_L/2)$, where $w_L = 2/\left(M+1\right)$.

In such a case, (12) becomes

$$\begin{aligned}\Delta(q) &= \frac{1}{\sqrt{m}}\left|\frac{\binom{m}{r-2}+\binom{m}{r-1}}{\binom{m}{r-2}+\binom{m}{r-1} + m\left(\binom{m-1}{r-3}+\binom{m-1}{r-1}\right)} - \frac{2}{m+2}\right| \\ &= \frac{1}{(m+2)\sqrt{m}}\left|\frac{m^2-4mr+5m+4r^2-12r+8}{m^2-2mr+4m+2r^2-6r+5}\right|, \qquad (13)\end{aligned}$$

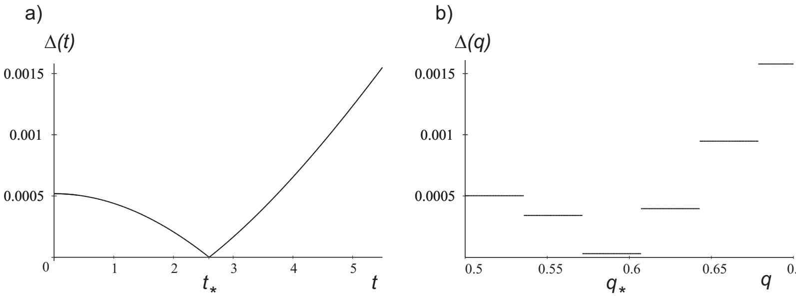
where $r := \lceil d\rceil = \lceil (M+1)\,q\rceil$. To analyse this dependence we introduce a new variable

$$t := r - M/2 - 1 = \lceil (M+1)\,q\rceil - M/2 - 1, \qquad (14)$$

obtaining

$$\begin{aligned}\Delta(t) &= \frac{2}{(M+1)\sqrt{M-1}}\frac{\left|M-4t^2\right|}{M^2+4t^2} \\ &= \frac{4}{(M+1)\sqrt{M-1}}\frac{\sqrt{M}+2t}{M^2+4t^2}\left|\sqrt{M}/2 - t\right|. \qquad (15)\end{aligned}$$

Fig. 1 — a) The 'mean discrepancy' $\Delta(t)$ as a function of the parameter $t$;
b) The mean discrepancy $\Delta(q)$ as a function of the parameter $q$
(in both cases $c = 2$, $M = 27$)




In principle, one can minimize this expression finding $\min \Delta(t) = 0$ for $t_* = \sqrt{M}/2$, see Fig.1a. However, due to the presence of the roof function in (14), $\Delta(q)$ is not a continuous function of the quota, and, consequently, the optimization problem $\min \Delta(q)$ does not have a unique solution and the minimal value may be greater than 0, see Fig.1b. Nevertheless, applying (14) and (15), one can show that there exists an optimal quota $q_*(M)$ in the interval

$$\frac{M+\sqrt{M}}{2(M+1)} \leq q_*(M) \leq \frac{2+M+\sqrt{M}}{2(M+1)} . \tag{16}$$

This means that for a large number $M$ of players the optimal quota behaves exactly as

$$q_*(M) \simeq q_s(M) := \frac{1}{2}\left(1+\frac{1}{\sqrt{M}}\right) . \tag{17}$$

Although this is an asymptotic formula, it works also for a moderate number of states. Moreover, it follows from (13) and (16) that the minimal mean discrepancy $\Delta(q_*(M)) \leq 8/M^3$.

Surprisingly, the efficiency of the system given by

$$\begin{aligned} A(q_s(M)) &= \frac{\sum_{k=r(M)-2}^{M-1}\binom{M-1}{k} + \sum_{k=r(M)}^{M-1}\binom{M-1}{k}}{2^M} \\ &= \frac{\sum_{k=r(M)}^{M}\binom{M}{k} + \binom{M-1}{r(M)-2}}{2^M} , \end{aligned} \tag{18}$$

where $r(M) := \lceil (M+1)\, q_s(M) \rceil$, does not decrease with the number of players to 0. On the contrary, it is always larger than $15/128 \approx 0.117$ and, according to the central limit theorem, it tends to $1-\Phi(1) \approx 0.159$ for $M \to \infty$.

Analogous considerations for $c = 3$ give similar result:

$$\frac{1 + M + \sqrt{M}}{2(M+1)} \leq q_* (M) \leq \frac{3 + M + \sqrt{M}}{2(M+1)}, \tag{19}$$

and so, also in this case, $q_* (M) \simeq \frac{1}{2}(1 + 1/\sqrt{M})$.

## 4. Normal approximation

Let us have a closer look at the approximate formula (17) for the optimal quota. In the limit $M \to \infty$ the optimal quota tends to 1/2 in agreement with the Penrose limit theorem, see Lindner (2004), Lindner and Machover (2004). Numerical values of the approximate optimal quota $q_s$ obtained in our toy model for $c = 2$ and $c = 3$ are consistent, with an accuracy up to two per cent, with the data obtained numerically by averaging quotas over a sample of random weights distributions (generated with respect to the statistical measure, i.e. the symmetric Dirichlet distribution with Jeffreys' priors), see Życzkowski and Słomczyński (2004), Życzkowski et al. (2006).[2] Furthermore, the above results belong to the range of values of the quota for qualified majority, which have been used in practice or recommended by experts on designing the voting systems.

Consider now a voting body of $M$ members and denote by $w_k$, $k = 1, \ldots, M$, their normalized voting weights fulfilling $\sum_{i=1}^{M} w_i = 1$. Feix et al. (2007) proposed (also in the EU context) yet another method of estimating the optimal quota for the weighted voting game $[q; w_1, \ldots, w_M]$, where $q \in [0.5, 1]$. They considered the histogram $n$ of the sum of weights (number of votes) achieved by all possible coalitions

$$n(z) = \frac{\operatorname{card}\{I \subset \{1, \ldots, M\} : \sum_{i \in I} w_i = z\}}{2^M} \tag{20}$$

and assumed that it allows the normal approximation with the mean value $m = \frac{1}{2} \sum_{i=1}^{M} w_i = \frac{1}{2}$ and the variance $\sigma^2 = \frac{1}{4} \sum_{i=1}^{M} w_i^2$, i.e.

$$\mathcal{N}(q) := \sum_{z \leq q} n(z) \approx \int_{-\infty}^{q} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx = \Phi\left(\frac{q-m}{\sigma}\right), \tag{21}$$

where $\Phi$ is the standard normal cumulative distribution function. The authors argued that for the quota close to the inflection point $q_n := m + \sigma$ of the

[2] Nevertheless, one can construct an artificial model with different values of optimal quota. In this aim, it is enough to consider one 'small' state and an even number of 'large' states with equal population (i.e. $c < 1$ in our toy model), see Lindner (2004), Lindner and Machover (2004). As Lindner stressed: 'experience suggests that such counter-examples are atypical, contrived exceptions'.

normal curve, where the 'density' of the histogram is approximately linear, the ratios $\beta_k/w_k$ $(k = 1, \dots, M)$ are close to 1. In other words, the quota $q_n$ is close to the optimal quota $q_*$. In Appendix we show how this fact follows from the normal approximation formula for the absolute Banzhaf indices. In particular we use heuristic arguments to demonstrate that in this case

$$\psi_k \approx \sqrt{\frac{2}{\pi e}} \frac{w_k}{\sqrt{\sum_{i=1}^M w_i^2}} \tag{22}$$

and, in consequence,

$$\beta_k \approx w_k \tag{23}$$

for $k = 1, \dots, M$. The validity of this method depends on the accuracy of the normal approximation for the absolute Banzhaf indices (see Appendix). The condition for the latter, which follows from the Berry-Esseen inequality, is

$$\max_{j=1,\dots,M} w_j \ll \sqrt{\sum_{i=1}^M w_i^2} \; . \tag{24}$$

For the thorough discussion of the problem see Owen (1975), Leech (2003), Lindner (2004), Feix et al. (2007), O'Donnell and Servedio (2008). For the Penrose voting system, where $w_k \sim \sqrt{N_k}$ $(k = 1, \dots, M)$, (24) is equivalent to

$$\max_{j=1,\dots,M} N_j \ll \sum_{i=1}^M N_i \; , \tag{25}$$

which means that the population of each country is relatively small when compared with the total population of all countries. One can easily check that it is more likely that (24) holds in this case than when the weights are proportional to the population.

Approximating the optimal quota $q_*$ by the inflection point of the normal distribution, $q_n = m + \sigma$, we arrive at an explicit weights-dependent formula for the optimal quota:

$$q_* \simeq q_n(w_1, \dots, w_M) := m + \sigma = \frac{1}{2}\left(1 + \sqrt{\sum_{i=1}^M w_i^2}\right) . \tag{26}$$

This approximation of the optimal quota can be directly compared with the approximation (17) obtained for the toy model. Since $\sum_{i=1}^M w_i = 1$ implies $\sum_{i=1}^M w_i^2 \geq 1/M$, it follows that

$$q_s(M) = \frac{1}{2}\left(1 + \frac{1}{\sqrt{M}}\right) \leq \frac{1}{2}\left(1 + \frac{1}{\sqrt{M_{eff}}}\right) = q_n \; , \tag{27}$$

Table 1 — Comparison of optimal quotas for the Penrose voting system applied to the EU ($q_*$) and for two approximations ($q_s$, $q_n$).

| $M$ | 15 | 25 | 27 |
|---|---|---|---|
| year | 1995 | 2004 | 2007 |
| $q_s$ [%] | 62.9 | 60.0 | 59.6 |
| $q_*$ [%] | 64.4 | 62.0 | 61.5 |
| $q_n$ [%] | 64.9 | 62.2 | 61.6 |

*Note* Calculations are based on data from: *50 Years of Figures on Europe. Data 1952-2001.* Office for Official Publications of the European Communities: Luxembourg, 2003, and on data from: EUROSTAT: Lanzieri G. Population in Europe 2005: First Results. Statistics in Focus. *Population and social conditions* 2006; 16: 1–12.

where $M_{eff} := 1/\sum_{i=1}^{M} w_i^2$ is equal to the *effective number of players*. (This quantity was introduced by Laakso and Taagepera (1979) and is the inverse of the more widely used *Herfindahl–Hirschman index of concentration* (Hirschman (1945), Herfindahl (1950), see also Feld and Grofman (2007).) The equality in (27) holds if and only if all the weights are equal. For the Penrose voting system we have

$$q_n = \frac{1}{2}\left(1 + \frac{\sqrt{\sum_{i=1}^{M} N_i}}{\sum_{i=1}^{M} \sqrt{N_i}}\right), \tag{28}$$

where $N_k$ stands for the population of the $k$-th country. For the toy model we get $q_n = \frac{1}{2}\left(1 + \frac{\sqrt{M+c^2-1}}{M+c-1}\right) \simeq q_s(M)$ for large $M$.

Both approximations $q_s$ and $q_n$ are consistent with an accuracy up to two per cent, with the optimal quotas $q_*$ obtained for the Penrose voting system applied retrospectively to the European Union (see Tab. 1 above). Observe that in this case the approximation of the optimal quota $q_*$ by $q_n$ is better for larger number of states, where the normal approximation of the histogram is more efficient.

Applying the normal approximation one can easily explain why the efficiency $A$ of our system does not decrease when the number of players $M$ increases. We have

$$A(q_s) \geq A(q_n) \approx 1 - \mathcal{N}(q_n) \approx \int_{m+\sigma}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx\,. \tag{29}$$

The right-hand side of this inequality depends neither on $m$ nor on $\sigma$, and it equals $1 - \Phi(1) \approx 15.9\%$, where $\Phi$ is the standard normal cumulative distribution function.

## 5. Double square root voting system

We shall conclude this paper proposing a complete voting system based on the Penrose square root law. The system consists of a single criterion only and is determined by the following two rules:

*A*. The voting weight attributed to each member of the voting body of size $M$ is proportional to the square root of the population he or she represents;

*B*. The decision of the voting body is taken if the sum of the weights of members of a coalition exceeds the quota $q_s = (1 + 1/\sqrt{M})/2$.

These rules characterize the *double square root* system: On one hand, the weight of each state is proportional to the square root of its population, on the other hand, the quota decreases to 0.5 inversely proportionally to the square root of the size of the voting body. If the weights $w_i$ are fixed, one can set the quota to $q_n = (1 + (\sum_{i=1}^{M} w_i^2)^{1/2})/2$, or just take the optimal quota $q_*$ which, however, requires more computational effort.

Such a voting system is extremely simple, since it is based on a single criterion. It is objective and so cannot a priori handicap a given member of the voting body. The quota for qualified majority is considerably larger than 50% for any size of the voting body of a practical interest. Thus the voting system is also moderately conservative. Furthermore, the system is representative and transparent: the voting power of each member of the voting body is (approximately) proportional to its voting weight. However, as a crucial advantage of the proposed voting system we would like to emphasize its extendibility: if the size $M$ of the voting body changes, all one needs to do is to set the voting weights according to the square root law and adjust the quota. The system is also moderately efficient: as the number $M$ grows, the efficiency of the system does not decrease.

The formulae for the quotas $q_s(M)$ and $q_n$ can be also applied in other weighted voting games. Note that for a fixed number of players the quota $q_s(M)$ does not depend on the particular distribution of weights in the voting body. This feature may be relevant, e.g. for voting bodies in stock companies where the voting weights of stockholders depend on the proportion of stock that investors hold and may vary frequently.

Although the limiting behaviour $M \to \infty$ may not necessarily be interesting for politicians, our work seems to have some practical implications for the on-going debate concerning the voting system in the Council of the EU. Since the number of Member States is not going to be explicitly provided in the text of the European Constitution, one should rather avoid to include any specific threshold for the qualified majority. In fact the optimal quota depends on the number of members of the voting body, so there should be a possibility to adjust it in future without modifying the European Constitution.

## Acknowledgements

It is a pleasure to thank W. Kirsch, A. Pajala, T. Sozański, and T. Zastawniak for fruitful discussions, and to F. Bobay, M. Machover, A. Moberg, and E. Ratzer for beneficial correspondence on voting systems. We also would like to thank an anonymous referee for helping us to clarify several important points of our reasoning. This work was supported in part by the Marie Curie Actions Transfer of Knowledge project COCOS (contract MTKD-CT-2004-517186) and the MNiSW grant no 1 H02E 052 30.

*Note added.* After this work was completed we learned about a related work by Lindner and Owen (2007), in which the same toy model was investigated.

## Appendix: Optimal quota for the normal approximation

In this appendix we show that in vicinity of the inflection point $q_n = m + \sigma$ of the density of the normal distribution the relative Banzhaf indices $\beta_j$ are close to the weights $w_j$. This reasoning holds in particular for the Penrose voting system, for which the weights are proportional to the square root of the populations.

Consider a weighted voting game $[q; w_1, \ldots, w_M]$, where $q \in [0.5, 1]$ and $\sum_{i=1}^{M} w_i = 1$. Set $m := \frac{1}{2}\sum_{i=1}^{M} w_i = \frac{1}{2}$ and $\sigma^2 := \frac{1}{4}\sum_{i=1}^{M} w_i^2$. Let $j = 1, \ldots, M$. We put $m_j := m - w_j/2$ and $\sigma_j^2 := \sigma^2 - w_j^2/4$.

The absolute Banzhaf index

$$\psi_j = \Pr\left(\left\{I \subset \{1, \ldots, M\} : q - w_j \leq \sum_{i \in I, i \neq j} w_i < q\right\}\right) \tag{A.1}$$

is equal to the difference of the number of wining coalitions formed with and without the $j$–th player divided by $2^{M-1}$. A key step in our reasoning is to assume that the sum of weights of the members of a coalition can be approximated by the normal distribution. This assumption implies that the Banzhaf index $\psi_j$ is approximately equal to the difference of two normal cumulative distribution functions taken at two points shifted by the corresponding weight $w_j$,

$$\psi_j \approx \Phi\left(q; m_j, \sigma_j\right) - \Phi\left(q - w_j; m_j, \sigma_j\right) . \tag{A.2}$$

Here $\Phi\left(x; \mu, d\right) = \Phi\left((x - \mu)/d\right)$ stands for the normal cumulative distribution function with mean $\mu$ and standard deviation $d$. Therefore

$$\psi_j \approx \Phi\left(\frac{q - m_j}{\sigma_j}\right) - \Phi\left(\frac{q - w_j - m_j}{\sigma_j}\right) . \tag{A.3}$$

We are going to analyse the behaviour of the power indices at the inflection point, $q = q_n := m + \sigma$. In such a case,

$$\begin{aligned}\psi_j &\approx \Phi\left(\frac{m+\sigma-m_j}{\sigma_j}\right) - \Phi\left(\frac{m+\sigma-w_j-m_j}{\sigma_j}\right)\\ &= \Phi\left(\frac{\sigma+\frac{1}{2}w_j}{\sigma_j}\right) - \Phi\left(\frac{\sigma-\frac{1}{2}w_j}{\sigma_j}\right)\\ &= \Phi\left(\sqrt{\frac{1+\nu_j}{1-\nu_j}}\right) - \Phi\left(\sqrt{\frac{1-\nu_j}{1+\nu_j}}\right), \qquad \text{(A.4)}\end{aligned}$$

where $\nu_j := w_j/2\sigma = w_j/\sqrt{\sum_{i=1}^M w_i^2}$. If $w_j \ll \sqrt{\sum_{i=1}^M w_i^2}$, then $\nu_j \ll 1$, and both $\sqrt{(1+\nu_j)/(1-\nu_j)}$ and $\sqrt{(1-\nu_j)/(1+\nu_j)}$ are close to 1. Near this point the standard normal density function $\Phi'$ is almost linear and so

$$\Phi\left(\frac{\sigma+\frac{1}{2}w_j}{\sigma_j}\right) - \Phi\left(\frac{\sigma-\frac{1}{2}w_j}{\sigma_j}\right) \approx \Phi'\left(\frac{\sigma}{\sigma_j}\right)\frac{w_j}{\sigma_j}. \qquad \text{(A.5)}$$

From (A.4) and (A.5) we deduce that

$$\begin{aligned}\psi_j &\approx \Phi'\left(\frac{\sigma}{\sigma_j}\right)\frac{w_j}{\sigma_j}\\ &= \frac{1}{\sqrt{2\pi}}\frac{w_j}{\sigma_j}\exp\left(-\frac{\sigma^2}{2\sigma_j^2}\right)\\ &= \sqrt{\frac{2}{\pi}}\frac{w_j}{\sqrt{\left(\sum_{i=1}^M w_i^2\right)-w_j^2}}\exp\left(-\frac{\sum_{i=1}^M w_i^2}{2\left(\left(\sum_{i=1}^M w_i^2\right)-w_j^2\right)}\right)\\ &= \sqrt{\frac{2}{\pi}}\frac{\nu_j}{\sqrt{1-\nu_j^2}}\exp\left(-\frac{1}{2\left(1-\nu_j^2\right)}\right)\\ &= \sqrt{\frac{2}{\pi e}}\nu_j + o\left(\nu_j^4\right). \qquad \text{(A.6)}\end{aligned}$$

Consequently,

$$\psi_j \approx \sqrt{\frac{2}{\pi e}}\frac{w_j}{\sqrt{\sum_{i=1}^M w_i^2}} + o\left(\nu_j^4\right), \qquad \text{(A.7)}$$

and so

$$\beta_j/w_j \approx 1. \qquad \text{(A.8)}$$

The above reasoning shows that if we select the inflection point $q_n$ for the quota $q$, then all the normalised Penrose–Banzhaf indices $\beta_j$ are approximately equal to the weights $w_j$, and so $q_n$ must be close to the critical quota $q_*$. The accuracy of this approximation depends highly on the accuracy of the normal approximation in (A.2), see condition (24).

Note that for the quota $q = m = 1/2$ we get (see Lindner (2004), Lindner and Machover (2004) for the formal proof)

$$\begin{aligned}\psi_j &\approx \Phi'(0)\,\frac{w_j}{\sigma_j} \\ &= \sqrt{\frac{2}{\pi}}\,\frac{w_j}{\sqrt{\left(\sum_{i=1}^{M} w_i^2\right) - w_j^2}} \\ &= \sqrt{\frac{2}{\pi}}\,\frac{w_j}{\sqrt{\sum_{i=1}^{M} w_i^2}} + o\left(v_j^2\right)\ . \end{aligned} \tag{A.9}$$

In this case, the second order terms in $v_j$ are present and, in consequence, the indices $\beta_j$ need not be as close to $w_j$ as for $q = q_n$, where (A.7) holds up to corrections of order four in $v_j$. Moreover, it is interesting to note, that increasing the threshold from $m = 1/2$ to $q_n = m + \sigma = 1/2 + \sigma$ causes the decrease of the absolute power indices $\phi_j$ by the factor $1/\sqrt{e} \approx 0.607$.

## References

Algaba, E., Bilbao, J.M. and Fernández, J.R. (2007), The distribution of power in the European Constitution, *European Journal of Operational Research* 176: 1752–1766.

Aziz, H., Paterson, M. and Leech, D. (2007), *Efficient Algorithm for Designing Weighted Voting Games*. Preprint. http://www.dcs.warwick.ac.uk/reports/cs-rr-434.pdf

Baldwin, R.E. and Widgrén, M. (2004), *Council Voting in the Constitutional Treaty: Devil in the Details*. (CEPS Policy Briefs No. 53; Centre for European Policy Studies, Brussels)

Banzhaf, J.F. (1965), Weighted voting does not work: A mathematical analysis, *Rutgers Law Review* 19: 317–343.

Bilbao, J.M. (2004), *Voting Power in the European Constitution*. Preprint. http://www.esi2.us.es/~mbilbao/pdffiles/Constitution.pdf

Bobay, F. (2004), Constitution européenne: redistribution du pouvoir des États au Conseil de l'UE, *Économie et Prévision* 163: 101–115.

Cameron, D.R. (2004), The stalemate in the constitutional IGC over the definition of a qualified majority, *European Union Politics* 5: 373–391.

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

Coleman, J.S. (1971), Control of collectivities and the power of a collectivity to act, in: B. Lieberman (ed.), *Social Choice*, Gordon and Breach: New York, pp. 269–

300. Reprinted in: J.S. Coleman, 1986, *Individual Interests and Collective Action*, Cambridge University Press: New York.

College of Europe (2004), *Making Europe Work: A Compromise Proposal on Voting in the Council*. College of Europe: Warsaw and Bruges. http://www.coleurop.be/content/publications/pdf/MakingEuropeWork.pdf

Dubey, P. and Shapley, L.S. (1979), Mathematical properties of the Banzhaf power index, *Mathematics of Operations Research* 4: 99–131.

Feix, M.R., Lepelley, D., Merlin, V. and Rouet, J.L. (2007), On the voting power of an alliance and the subsequent power of its members, *Social Choice and Welfare* 28: 181–207.

Feld, S.L. and Grofman, B. (2007), The Laakso-Taagepera index in a mean and variance framework, *Journal of Theoretical Politics* 19: 101–106.

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

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

Felsenthal, D.S. and Machover, M. (2002), *Enlargement of the EU and Weighted Voting in its Council of Ministers* [online]. LSE Research Online: London. http://eprints.lse.ac.uk/407/01/euenbook.pdf

Felsenthal, D.S. and Machover M. (2004a), A priori voting power: What is it all about? *Political Studies Review* 2: 1–23.

Felsenthal, D.S. and Machover M. (2004b), Qualified majority voting explained, *Homo Oeconomicus* 21: 573–576.

Herfindahl, O.C. (1950), *Concentration in the Steel Industry*. PhD Dissertation; Columbia University.

Hirschman, A.O. (1945), *National Power and Structure of Foreign Trade*. University of California Press: Berkeley.

Hosli, M.O. (2000), *Smaller States and the New Voting Weights in the Council*. (Working Paper, Netherlands Institute of International Relations, Clingendael, July 2000) http://www.clingendael.nl/publications/2000/20000700_cli_ess_hosli.pdf

Hosli, M.O. and Machover, M. (2004), The Nice Treaty and voting rules in the Council: a reply to Moberg (2002), *Journal of Common Market Studies* 42: 497–521.

Kirsch, W. (2004), *The New Qualified Majority in the Council of the EU. Some Comments on the Decisions of the Brussels Summit*. Preprint. http://www.ruhr-uni-bochum.de/mathphys/politik/eu/Brussels.pdf

Laakso, M. and Taagepera, R. (1979), Effective number of parties: a measure with application to West Europe, *Comparative Political Studies* 12: 3–27.

Laruelle, A. and Valenciano, F. (2002), Inequality among EU citizens in the EU's Council decision procedure, *European Journal of Political Economy* 18: 475–498.

Laruelle, A. and Widgrén, M. (1998), Is the allocation of voting power among the EU states fair? *Public Choice* 94: 317–339.

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

Leech, D. (2003), Computing power indices for large voting games, *Management Science* 49: 831–838.

Lindner, I. (2004), *Power Measures in Large Weighted Voting Games. Asymptotic Properties and Numerical Methods*. PhD Dissertation; University of Hamburg. http://deposit.ddb.de/cgi-bin/dokserv?idn=972400516

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

Lindner, I. and Owen, G. (2007), Cases where the Penrose limit theorem does not hold, *Mathematical Social Sciences* 53: 232–238.

Mabille, L. (2003), *Essai sur une juste pondération des voix au Conseil de l'Union européenne*. Preprint; see also: Dubois, N., Pondération des voix: la preuve par "27", *Liberation* 26/11/2003.
http://pageperso.aol.fr/lcmabille/

Merrill III, S. (1982), Approximations to the Banzhaf index of voting power, *American Mathematical Monthly* 89: 108–110.

Moberg, A. (2002), The Nice Treaty and voting rules in the Council, *Journal of Common Market Studies* 40: 259–82.

O'Donnell, R. and Servedio, R.A. (2008), *The Chow Parameters Problem*. Preprint. http://www.cs.cmu.edu/~odonnell/chow-params.pdf

Owen, G. (1975), Multilinear extensions and the Banzhaf value, *Naval Research Logistics Quarterly* 22: 741–750.

Pajala, A. (2005), *Maximal Proportionality between Votes and Voting Power: The Case of the Council of the European Union* [online]. LSE Research Online: London. http://www.lse.ac.uk/collections/VPP/VPPpdf/VPPpdf_Wshop4/pajala.pdf

Paterson, I. and Silárszky, P. (2003), *Draft Constitution and the IGC: Voting can be Simple and Efficient – without introducing the Massive Transfer of Power implied by the Convention's Double Majority Proposal*. (Institute for Advanced Studies, Vienna, December 2003).
http://www.ihs.ac.at/publications/lib/forum2ihsdec2003.pdf

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*. H.K. Lewis & Co: London.

Plechanovová, B. (2004), *Draft Constitution and the Decision-Making Rule for the Council of Ministers of the EU – Looking for Alternative Solution*. (European Integration online Papers (EIoP), Vol. 8, No. 12)
http://eiop.or.at/eiop/pdf/2004-012.pdf

Plechanovová, B. (2006), Je rozhodovací procedura v Radě Evropské unie spravedlivá? *Mezinárodní vztahy* 41/1: 5–22.

Słomczyński, W. and Życzkowski, K. (2006), Penrose voting system and optimal quota, *Acta Physica Polonica* 37: 3133–3143.

Sutter, M. (2000), Fair allocation and re-weighting of votes and voting power in the EU before and after the next enlargement, *Journal of Theoretical Politics* 12: 433–449.

Taagepera, R. and Hosli, M.O. (2006), National representation in international organizations: the seat allocation model implicit in the European Union Council and Parliament, *Political Studies* 54: 370–398.

Tiilikainen, T. and Widgrén, M. (2000), Decision-making in the EU: a small country perspective, *The Finnish Economy and Society* 4: 63–71.

Widgrén, M. (2003), *Power in the Design of Constitutional Rules*. (European Economy Group, University of Madrid, Working Papers No. 23)
http://www.ucm.es/info/econeuro/documentos/documentos/dt232003.pdf

Widgrén, M. (2004), Enlargements and the Principles of Designing EU Decision-Making Procedures, in: Blankart, C.B. and Mueller, D.C. (eds.), *A Constitution for the European Union*, MIT Press: Cambridge, MA, pp. 85–108.

Życzkowski, K. and Słomczyński, W. (2004), *Voting in the European Union: The Square Root System of Penrose and a Critical Point.* Preprint cond-mat.0405396; May 2004.
http://arxiv.org/ftp/cond-mat/papers/0405/0405396.pdf

Życzkowski, K., Słomczyński, W. and Zastawniak, T. (2006), Physics for fairer voting, *Physics World* 19: 35–37.