Minimizing the threat of a positive majority deficit in two-tier voting systems with equipopulous units

Article *in* Public Choice · January 2011
DOI: 10.1007/s11127-011-9810-2

2 authors:

Claus Beisbart
Universität Bern

Luc Bovens
The London School of Economics and Political Science

Minimizing the Thread of a Positive Majority Deficit in Two-Tier Voting Systems with Equipopulous Units

Luc Bovens and Claus Beisbart

Abstract

Let us suppose that there is a two-tiered voting system in a company, a country, or a federation with *n* people partitioned in *m* equipopulous constitutive units. We institute simple majority voting at both levels and postulate a Bernoulli voting model.

We are interested in the sensitivity and the mean majority deficit for different types of partitions. Clearly, if there is only one unit or if there are as many units as there are people, then we have a one-tiered voting system with maximal sensitivity and a mean majority deficit of zero. Our question is: What is the partition that yields the undesirable feature of minimal sensitivity and maximal mean majority deficit?

We find the following results: For odd *n*, minimal sensitivity and maximal mean majority deficit occurs when *m* is close to the square root of *n*. For even *n*, minimal sensitivity and maximal mean majority deficit occurs when *m* is 2 or *n*/2.

We relax the assumption of equipopulous units and conclude with a discussion of the political relevance of our findings.

1. Introduction

A vote is taken in a population – in a company, in a university or in a nation. A binary issue is on the table, say whether to adopt a certain strategy, or whether to elect one of two candidates for a job. One may decide the issue by means of popular majority vote. Alternatively, there may be historical or other reasons to organise the vote as a two-tier voting system. The population is composed of *n* units. A vote is taken in each unit, a representative of the unit will convey the outcome of the majority vote in her unit to a board of representatives, and a majority vote in that board will decide the issue. The latter procedure may have certain advantages. To name one such advantage, people identify with their unit and may feel less alienated from the political process if their vote is recorded at the level of their unit and this vote carries weight at the federal level. But it has a clear disadvantage. A two-tier voting system may yield a different outcome than the outcome of the popular vote – although less people voted for a motion, say, the motion was accepted. We say that there is a positive *majority deficit* in this case. This raises questions of legitimacy and the greater the majority deficit, the stronger such concerns are. This was precisely the problem of the presidential elections in the US in 2000. Bush had won the electoral college, but Gore received 543,895 more votes than Bush.[1]

We are concerned with the majority deficit in a two-tier voting procedure. Let the majority deficit for a particular vote be zero when the outcomes of the popular vote and the two-tier voting system coincide and let it take on a positive value i when the outcomes disagree and there is a margin of i voters. It is better to avoid a (positive) majority deficit and if it is non-zero, then it is better for it to be minimal. The *mean majority deficit* (MMD) is the expectation value of the majority deficit. That is, the MMD equals the sum of the products of the probability that the majority deficit takes on the value n times n. (Felsenthal and Machover, 1998: 60)

There are many ways to set up a two-tier voting procedure. Often it is the case that units are pre-existing, e.g. when the company is a company of independent countries, such as the European Union. Much work has been done on how to design a voting system for a board of representatives for such a company. But we will ask a different question here. We will assume that we have a free hand in splitting up the population into units—only the units should be equal-sized. We need to decide into *how many* units we will split the population. In doing so, we need to monitor the MMD. How does our choice of the number of units affect the MMD? We will investigate this question under the Bernoulli model, i.e. each voter is equally likely to vote one way or the other and there is probabilistic independence between the votes. We will also assume that there is simple-majority voting at both tiers of the voting system – the board of representatives as well as the units take their votes following simple majority voting.

Let us suppose that our population contains n citizens and that we split it up into m units. Now suppose that $m = 1$ or that $m = n$. These are limiting cases of two-tier voting procedures, which just coincide with a popular vote and hence the MMD equals 0. For what m is the MMD maximal? That is the question of this paper. For simplicity, we will cast our discussion in terms of a company.

[1] http://www.infoplease.com/ipa/A0876793.html

2. The MMD under Equipopulous Units

Let there be n voters in the population. We assume that the population can be split in exactly m units with n/m people, each. Here, m and n/m are assumed to be integer-valued. Every unit has one representative whose vote is determined on the basis of simple majority voting within the unit. Voters make a binary choice with outcomes Y and N. The outcome is x iff more than one half of the representatives vote x for x = Y, N.

MD denotes the majority deficit and E[MD] is its expectation value. We calculate E[MD] on the basis of the Bernoulli model, i.e. every possible voting profile has the same probability of $1/2^n$.

Our results rest on a theorem in the voting power literature. We follow Felsenthal and Machover (1998: 60) who make reference to Dubey and Shapley (1979). The theorem relates E[MD] to *sensitivity* S, which is defined as the sum of the non-normalised Banzhaf voting powers β_i' for all voters $i = 1,\ldots, n$. (Felsenthal and Machover, 1998: 39) According to the theorem E[MD] is a linear function of S:

(1) $$E[MD] = \frac{S_n - S}{2} \quad ,$$

where S_n is the sensitivity for simple majority voting with n voters and is a constant for a given n. The theory implies that maximising sensitivity is equivalent to minimising the E[MD]. Since the optimal value for E[MD] is zero, maximal sensitivity obtains for simple majority voting when S equals S_n.

Let us now calculate the sensitivity S for our two-tier voting system. The voting power of a single person i is the probability that i is doubly pivotal. Under the Bernoulli model, the probability factorises in the probability that i is pivotal in her unit and that her representative is pivotal in the board of representatives. The probability that a voter is pivotal in a simple majority vote with k voters, P_k, is

(2) $$P_k = Binomial[(k-1),[k/2]]/2^{k-1} \quad ,$$

where $[k]$ is the largest integer l with $l \leq k$ and Binomial is the binomial coefficient. The probability of pivotality within one's unit (*unit*) thus equals

(3) $P(i \rightarrow unit) = Binomial\ [(n/m-1), [n/(2m)]]/2^{n/m-1}$.

Likewise, the probability that a representative of a particular unit is pivotal in the board of representatives (*br*) equals

(4) $P(unit \rightarrow br) = Binomial\ [m-1, [m/2]]/2^{m-1}$.

Since each voter has the same voting power, the sensitivity *S* equals

(5) $S = n \times P(i \rightarrow unit) \times P(unit \rightarrow br)$.

By the same reasoning, the sensitivity for the popular vote equals

(6) $S_n = n \times Binomial\ [n-1, [n/2]]/2^{n-1}$.

From Felsenthal and Machover (1998: 56), we know that S_n approaches $\sqrt{2n/\pi}$ for large *n*. We will sometimes use this approximation. Not much hinges on that, because our primary concern – the *m*-dependence of *E*[MD] – is not at all affected by that approximation. Combining Eqs. (1) and (3) – (6), we can evaluate the mean majority deficit for every pair of *n*- and *m*-values.

We now have to distinguish between several cases.

2.1. First case: odd *n*

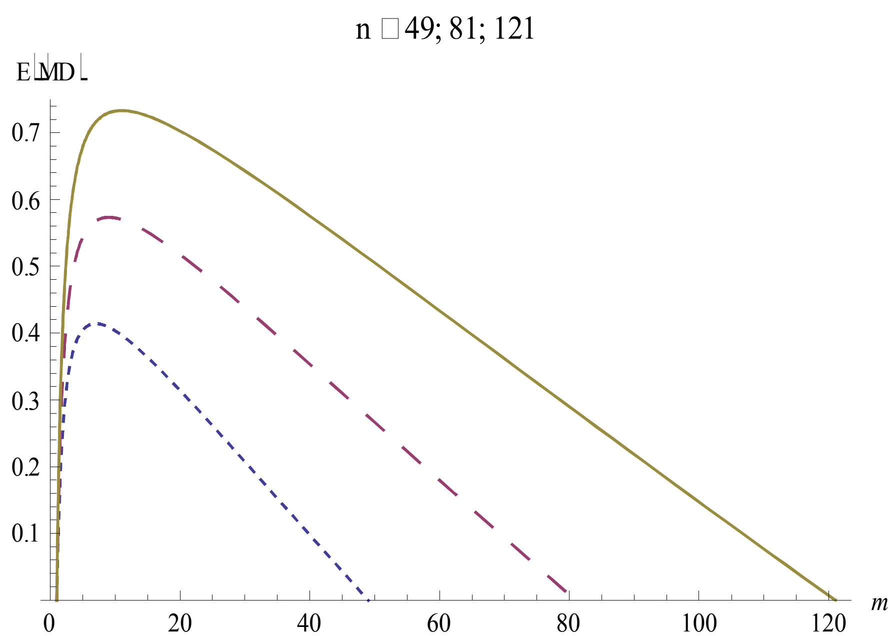
If the population size is odd, both the numbers of units and the number of voters within each unit have to be odd as well. As a consequence, the voting power of a voter is

(7) $Binomial\ [(n/m-1), (n/m-1)/2]/2^{n/m-1} \times Binomial\ [(m-1), (m-1)/2]/2^{m-1}$.

Strictly speaking, this result holds only for *m* and *n/m* being integers. But expressing the Binomials in terms of factorials and using the Gamma function as an extension of the faculty: ($n! = \Gamma(n+1)$), we may think of the voting power and of *E*[MD] as functions that are defined on the whole interval [1, *n*]. Call this the real-extended voting power and the real-extended *E*[MD].

We show the real-extended E[MD] for three different values of n as a function of m (solid line).

Figure 1. The real-extended mean majority deficit E[MD] for two-tier voting systems with n = 49 (short-dashed blue line); 81 (long-dashed, pink); 121 (solid green) voters as a function of the number of the units m.



In Fig. 1, the curves for the different n-values have the same shape: They start at 0, increase, as m increases, reach a maximum and go down to zero again. For each curve, the maximum is located at $m = \sqrt{n}$. The same holds for other values of n, even for n-values that cannot be thought of as the square of an integer.

How can we understand this result? We have not been able to prove analytically that, for any n, the curve for the mean majority deficit follows the shape that can be seen in Fig. 1. But we can provide some analytic results that help us understand and generalize what we observe in Fig. 1. A first general result is the following theorem:

Theorem 1. For any integer $n,$ the real-extended E[MD] has an extremum at $m = \sqrt{n}$.

The proof is in Appendix 1.

Furthermore, by plotting the second derivative of the real-extended E[MD] at $m = \sqrt{n}$ for a broad range of values of n, we observe that that extremum is a maximum.

But that does not yet settle the question how E[MD] behaves as a function of m, because there may be other minima or maxima. In particular, we can not yet infer that the maximum that we have found is a global one.

For further analytical argument, we consider approximations of E[MD]. E[MD] contains binomials and hence factorials. Accordingly, the real-extended E[MD] contains Gamma functions. As is well known, the Gamma functions may be approximated by what is sometimes called the Stirling series. The Stirling series starts (see Morse and Feshbach 1953, p. 443):

$$(8) \qquad k! = \Gamma(k+1) \approx \sqrt{2\pi k}\, e^{k\ln(k)-k}\left(1+\frac{1}{12k}+\ldots\right).$$

Unfortunately, this series does not converge. Care is therefore required in using Eq. (8) for approximations.

We consider tentatively a few approximations by keeping the first r addends in the bracket of the Stirling series. We consider $r = 1$ and $r = 2$. For $r = 1$, we obtain Stirling's formula:

$$(\mathrm{a1}) \qquad k! \approx \sqrt{2\pi k}\, e^{k\ln(k)-k}.$$

Stirling's formula is commonly used to approximate a binomial distribution by a Gaussian one.

For $r = 2$, we obtain what we call the improved Stirling's formula (a2) $$k! \approx \sqrt{2\pi k}\, e^{k\ln(k)-k}\left(1+\frac{1}{12k}\right).$$

Eq. (1) for the mean majority deficit contains factorials of (m – 1), (m – 1)/2, (n/m – 1) and (n/m – 1)/2 because of Eqs. (3) and (4). We obtain two approximations for the mean majority deficit by consistently replacing the factorials by the approximations (a1) and (a2), respectively:

(A1) $$E[\mathrm{MD}] \approx .5 * \left(S_n - \frac{n}{\pi\sqrt{(m-1)\left(n/m-1\right)}} \right)$$

and

(A2) $$E[\mathrm{MD}] \approx .5 * \left(S_n - \frac{n}{\pi}\sqrt{\frac{1}{(m-1)(n/m-1)}\frac{(1+1/(12(m-1)))}{(1+1/(6(m-1)))^2}\times\frac{(1+1/(12(n/m-1)))}{(1+1/(6(n/m-1)))^2}} \right).$$

As is well-known (see Felsenthal and Machover 1998, p. 56, e.g.), approximating a binomial with a single-case probability of .5 and with l trials (where trials correspond to votes in our case) by a Gaussian distribution yields a relative error that is proportional to $1/l$. Thus, the relative error of A1 may be estimated by

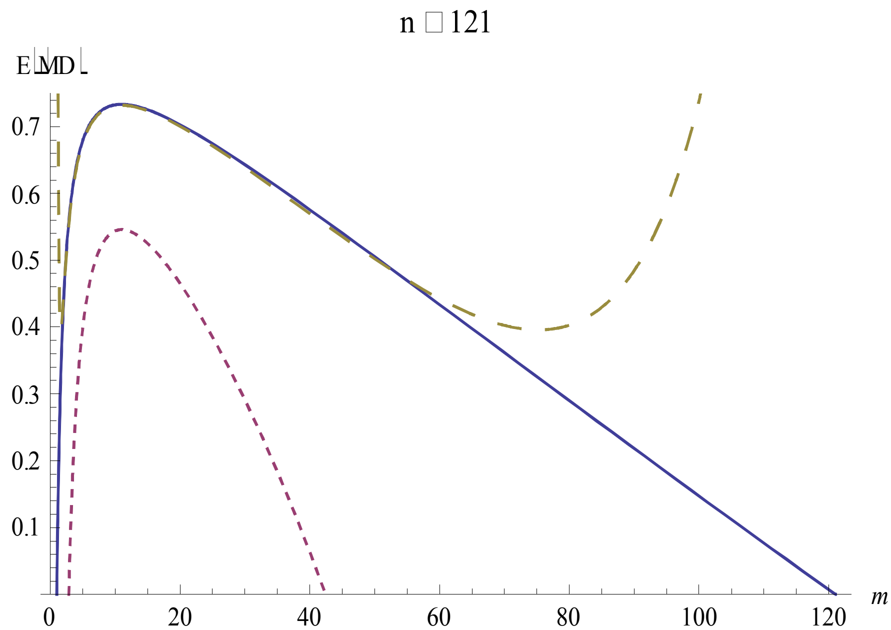
$$\left(1 \pm \frac{1}{m-1}\right)\times\left(1 \pm \frac{1}{n/m-1}\right).$$

This means that the approximation A1 is only good for $m >> 1$ and $n/m >> 1$,which is the middle range of our graphs.

The results for both approximations for n = 121 are shown in Fig. 2, where the dashed line represents the approximation for r = 1, whereas the dotted line represents the approximation for r = 2. Note, first that both curves have a maximum at $m = \sqrt{n}$. It can in fact be shown that every approximation from the Stirling series has a zero derivative at this place; this follows from the fact that all approximations fulfil the conditions of Theorem 1.

As is also plain from Fig. 2, neither approximation is very good for very small m, or for very large m (i.e. for small n/m). This was expected from our earlier considerations. But from the figure we may also conclude that the second approximation is very close to the analytic result for a sufficiently broad range of *m*-values, whereas the first approximation is not.

Fig. 2: The exact result (solid blue line), the approximation A1 (short-dashed) and A2 (long-dashed).



Both approximations help us to constrain the shape of the curve. The derivative of A1 with respect to *m* reads

(9) $$5n\frac{n-m^2}{\pi\sqrt{m}(n-m)^{3/2}(m-1)^{3/2}} \quad .$$

In the interval $[1,n]$, there is thus only one zero of the derivative, and we can see immediately that the derivative changes its sign from positive to negative, as m increases and crosses $\sqrt{n}$. As a consequence, A1 predicts exactly one maximum of the mean majority deficit in $(1, n)$, with a location at $m = \sqrt{n}$. This is the result we wish to understand. Still, arguing from A1 alone is problematic, since A1 does not match the analytic curve very well.

A2 is a bit more difficult to investigate. In the interval $[1,n]$, its derivative has three zeroes, one of them at $m = \sqrt{n}$. One can prove that, for $n \geq 6$, the second derivative at $m = \sqrt{n}$ is negative such that we have a maximum. The other zeroes of the derivative cannot be local maxima. It can be shown that, in the limits $m \to 1$ and $m \to n$, A2 approximates S_n instead of 0. That is artificial, and so are the two other zeroes of A2.

We conclude this section with three remarks.

First, in the literature, the probability that a voter is pivotal in a simple majority game with k voters is sometimes approximated as $P_k \approx \sqrt{2/(\pi k)}$ (cf. Felsenthal and Machover 1998, p. 56). For an odd number of voters the approximation arises as follows: In Eq. (6), the factorials in the binomials are approximated using Stirling's formula (a1), and in the final result, (k-1) is replaced by k. If this approximation is used for evaluating $P(i \to const)$ and $P(const \to fa)$ in the probability of double pivotality, we obtain:

$$(10) \quad P(i \to const)P(const \to fa) = \frac{2}{\pi}\sqrt{\frac{1}{m}}\sqrt{\frac{m}{n}} = \frac{2}{\pi}\sqrt{\frac{1}{n}} \ .$$

Morriss (1987/2002, pp. 189-91) uses this approximation in order to provide an argument against indirect democracy. But the approximation is very rough, and it suppresses the m-dependence of E[MD]. This is why we cannot work with this approximation in our study.

Second, an interesting question is how large the maximum mean majority deficit at $m = \sqrt{n}$ is. Using approximation A1, we find that the probability of double pivotality at this point is

$$(11) \quad \frac{1}{\pi\left(\sqrt{n}-1\right)} \ .$$

Accordingly, the maximal E[MD] takes approximately the following value

$$(12) \quad \frac{1}{2} \times \left(S_n - \frac{n}{\pi\left(\sqrt{n}-1\right)} \right).$$

Consider now Banzhaf voting under simple majority voting and under a two-tier voting system with the worst possible E[MD]. In the limit of large numbers of voters n, their ratio approaches $\sqrt{2/\pi}$. This is the ratio that Morriss (1987/2002, p. 190) obtains by using the approximation described in our first point, independently from m.

Third, care is required in interpreting our results. In our figures and arguments, we take E[MD] to be a continuous function of m. But ultimately, m can only take integer values, and the continuous curves are only extrapolations. Moreover, m has to be chosen in such a way that n/m is an integer. Suppose, for instance, that somebody wants to maximize the mean majority deficit (which is indeed not a recommendable thing to do). In the general case, $m = \sqrt{n}$ will not be an integer. The question what to do in this case will be taken up in Sec. 3.

2.2 Second case: even *n*

Let us now consider the case of an even population size. If we partition the population in exactly m units, then m may be even or odd. In more detail, we may discriminate between the following three subcases:

a. n even, m even and n/m odd. This covers the special case of $m = n$.

b. n even, m odd and n/m even. This covers the special case of $m = 1$.

c. n even, m even and n/m even. This case requires n to be a multiple of 4.

A subcase where both m and n/m are odd, cannot occur, of course, because it would follow that n is odd. As the list makes also plain, subcases 2a and 2b will occur for any even n. On the contrary, subcase 2c will only be possible for n-values that are multiples of 4. Let us now specify the probabilities of a single citizen being doubly pivotal for all of the three subcases.

a. n even, m even and n/m odd.

$$P(DP) = Binomial(m-1, m/2) \times Binomial\left(n/m-1, \tfrac{1}{2}(n/m-1)\right) \times \frac{1}{2^{m-1}2^{n/m-1}}$$

b. n even, m odd and n/m even.

$$P(DP) = Binomial(m-1, (m-1)/2) \times Binomial(n/m-1, (n/2m)) \times \frac{1}{2^{m-1}2^{n/m-1}}$$

c. n even, m even and n/m even.

$$P(DP) = Binomial(m-1, m/2) \times Binomial(n/m-1, (n/2m)) \times \frac{1}{2^{m-1}2^{n/m-1}}$$

Note that the formulae are slightly different.

From the probability of double pivotality, E[MD] can easily be calculated for each of the cases using Eqs. (1) and (3) – (6). Accordingly, for each of the cases we get a slightly different expression for

E[MD]. Each of these formulae can be real-extended. Call the real-extensions of E[MD] for the three subcases $f_a(m)$, $f_b(m)$ and $f_c(m)$, respectively. Strictly speaking, these functions also depend on a value of n, but for simplicity of notation, we drop the n. We plot the real-extensions in the top panel of Fig. 3.

Figure 3: The real-extended E[MD] for even n and the subcases 2a (m even and n/m odd): pink, long-dashed line; 2b (m odd and n/m even): green, short-dashed line; 2c (m even and n/m even): solid blue line. In the lower panel, the green dots denote the divisors of n = 100 and the corresponding values of E[MD].

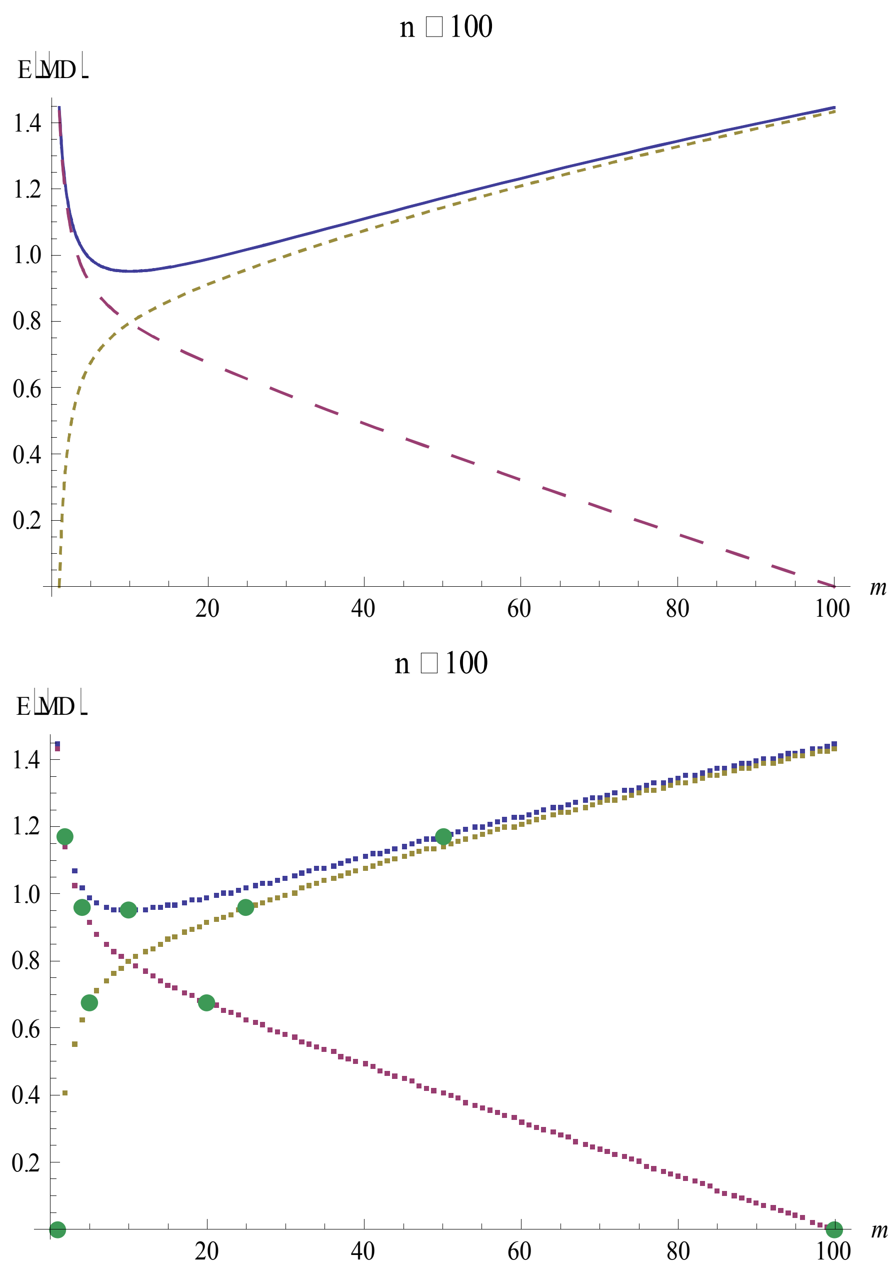
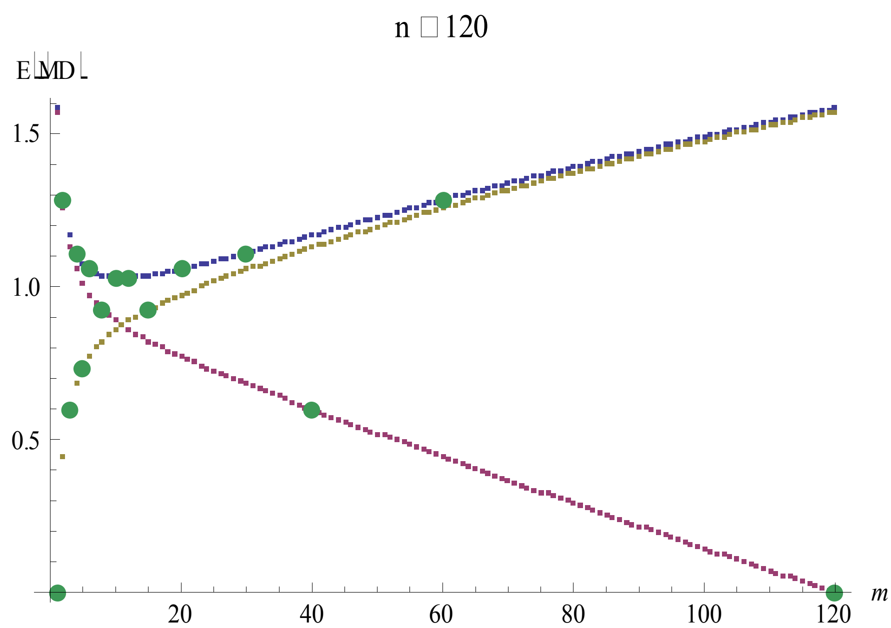


Figure 4. As Figure 3, but for n = 120.



The lower panel of Fig. 3 shows how one has to apply viewgraphs such as the one in the top panel: For a particular partition into *m* units, one has to check whether *m* is odd and whether *n/m* is odd. That specifies the case one is in, and, depending on the case, one has to select a curve. The *E*[MD] is the value of the curve at the particular *m*.

There are a few general things to be noted here. First, the curve for the third case obeys the relation that we have also found for the case of odd *n*:

(13) $$f_c(m) = f_c(n/m)\ .$$

Following the proof of Theorem 1, we can infer that $f_c(m)$ has a zero derivative at $m = \sqrt{n}$, but this time, it is apparently a minimum – at least for the examples that we have so far considered.

The curves for the first and the second case are related via:

(14) $$f_a(m) = f_b(n/m)$$

This means that it does not matter whether we partition the population in *m* or in *n/m* units – the *E*[MD] is the same for both cases. An immediate consequence is that $f_a(m)$ and $f_b(m)$ intersect with each other at $m = \sqrt{n}$. So that point is again an important one, but not because it has a maximum.

The next question is whether there is a general conclusion regarding that value of m for which the E[MD] is worst (i.e. maximal). We note two observations that are relevant at this point. The observations are not to be understood as mathematically waterproof results.

First, for each n and each m, $f_c(m)$ is always larger than the other curves.

Second, as n increases, $f_a(m)$ approaches $f_c(m)$ for $m < \sqrt{n}$; at the same time, $f_b(m)$ approaches $f_c(m)$ for $m > \sqrt{n}$.

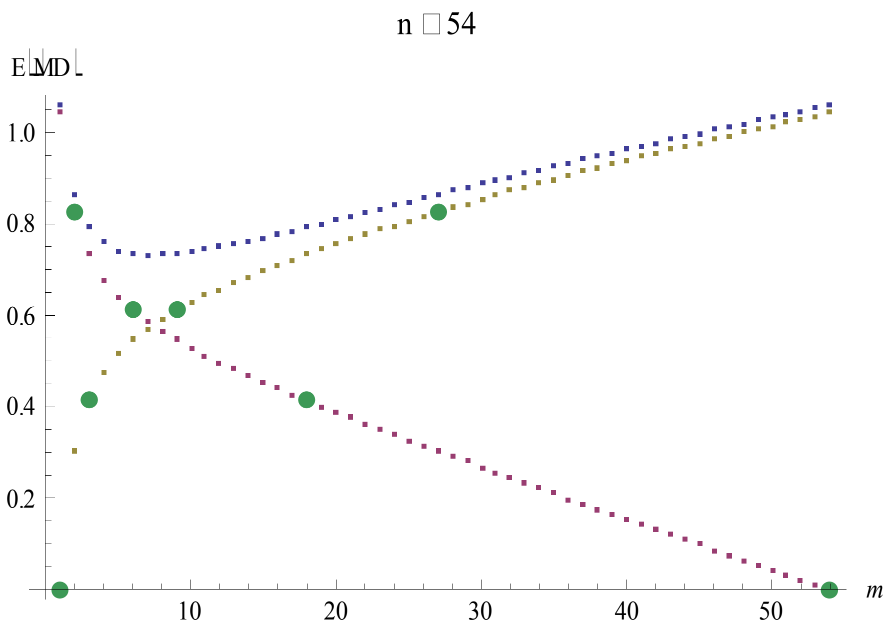
Taking for granted these observations, we infer the following:

a. If n is a multiple of 4, the third curve $f_c(m)$ is relevant. Because of our first observation, the worst maximum E[MD] must lie on this curve or below – the curve is a bound for the maximum E[MD]. This curve has a minimum at $m = \sqrt{n}$. Therefore, the further we get away from $m = \sqrt{n}$, the larger is $f_c(m)$. Now the furthest we can get from that point with a permitted m-value (m a divisor of n), is $m = 2$ or $m = n/2$. Since n is a multiple of 4, we know that E[MD] coincides with the curve at this point. Thus, $m = 2$ or $m = n/2$ provide the maximal E[MD].

b. If n is not a multiple of 4, $f_c(m)$ is not relevant, and we have to work with the other two curves. Obviously, you reach the maximum E[MD], if you start with $m = 1$ and move to the next larger divisor of n until you reach the first even one. But that m equals 2, because n is even and must have 2 as its divisor. At that point, you will have the maximum E[MD]. Of course, n/m at this value of m will give another maximum E[MD]. That means that, rather than starting at $m = 1$, you can also start at $m = n$ and then go down until you find the first odd divisor, and this will be $n/2$. For $n = 54$ that would mean that there is a maximum E[MD] at $m = 2$ and $m = 27$. Compare Fig. 5 for this.

Altogether, we find a universal rule: If n is even, then the maximum E[MD] is to be found at $m = 2$ and $m = n/2$.

Figure 5. As Fig. 3, but with n = 54.



3. Partitions into units that do not have exactly the same size

So far, we have only considered partitions with units that are exactly equipopulous. Now if the population size n is a prime number, then will be no way to obtain such a partition. And even in case n is not prime, there will not be many ways to partition the population into exactly equipopulous units. We therefore consider partitions with units that are not exactly equipopulous.

[will be completed soon]

4. Discussion

We have shown that in two-tier voting systems with equal-sized units, the mean majority deficit is maximal when the number of units equals the square root of the population size and the population size is odd. This may clearly be named ‘a square-root rule’. In the literature, a number of square root rules have been found so far. We clarify the significance of our result by comparing it to three other square root rules in voting theory.

The so-called first square root rule concerns Banzhaf voting power in a two-tier voting system with units that are large in population, but not necessarily equipopulous. According to the first square root rule, the voting powers of the people – their probabilities of being doubly pivotal – are equalized, if the weights of the units in the board of representatives are set proportional to the square roots of the population sizes (see Felsenthal and Machover 1998, pp. 66–8 for a statement of the rule and further references, e.g. Penrose 1946 and Banzhaf 1966). Since equal shares of voting power for the different people constitute a desirable characteristic of a voting system (cf. Felsenthal and Machover 1998, pp. 63–4), the first square root rule is often appealed to in normative assessments of voting rules (see Felsenthal and Machover 2000 and Życzkowski and Słomczyński 2004 for a recent example concerning the Council of Ministers in the European Union). Clearly, our result differs from the first square root rule since we are concerned with minimizing the mean majority deficit rather than with equalizing voting power.

The second square root rule states that the mean majority deficit is minimal in a two-tier voting system with sufficiently large units, if and only if the weights for units in the board of representatives are proportional to the square roots of the population sizes (see Felsenthal and Machover 1998, pp. 74–5 for a discussion). Although both the second square root rule and our result concern the mean majority deficit, there are significant differences. First, whereas the second square root rule aims to determine the weights for a company with fixed units of variable sizes in order to minimize the mean majority deficit, we aim to determine a partition of the company into equipopulous units that minimises the mean majority deficit.

Recently, a third square root rule has been proven by Edelman (2004, see also Edelman 2005 for further discussion). Edelman considers two-tier voting systems under which a voter has several votes that may be cast independently. Since these voting systems are markedly different from the voting systems that we investigate, there is no connection between our results and Edelman’s square root rule.

What might be the political significance of this result?

Suppose that we are designing a decision-making procedure for an institution—to stay close to home, say, a university. How should votes on matters that affect the complete institution be taken? There are multiple concerns that need to be taken into account.

First, a positive majority deficit, and in particular a large positive majority deficit, reflects badly on the legitimacy of the decision. It suffices to remember the public reaction to the US presidential election in 2000. Let call this the *majoritarian concern*—that is, the voting rule should "come as close as possible to produce outcomes that conform to the wishes of the majority of the entire electorate" (see Felsenthal and Machover 2000, p. 18, 23–6).

Second, there is a *motivational concern*. In university-wide polling, voters often feel alienated from the process and turnout is typically lower than in a federalised structure. Voters feel more engaged in the process of decision-making when they can vote within a unit with which they feel a sense of allegiance and when there is a public record of these votes.

Third, there is a *deliberational concern*. Democracy is not only about voting—it is also about the deliberative process that informs decision-making. Now deliberations in university-wide forums are notoriously frustrating. A federal structure in which deliberations can take place within the separate units and within the board of representatives is preferable to ensure that all voices have been heard. And even if majority votes about the issue at hand in the lower tiers are simply entered into a majority vote at the higher tier, it is often the case that representatives are asked to convey the reasoning process that underlies the majority vote of their unit or they may take decisions on other issues through group deliberation.

These concerns are *pro tanto* reasons. There may even be additional pro tanto reasons.[2] All of these reasons may pull in different directions in institutional design. The majoritarian concern clearly favours a one-tier voting system. And furthermore, if we do adopt two-tiered voting systems, for an odd population size, it steers us away from a partition with equal-sized units whose sizes equal the number of units. In case the population size is even, it steers us away from a partition with 2 units or with units with 2 persons each. The motivational concern favours a two-tier voting in which the units are determined by the allegiances of the faculty. The deliberational concern pulls precisely in the opposite direction from the majoritarian concern for an odd population size. It is easy to show that if we wish to

[2] It is sometimes required that the sensitivity be high in order to have a voting system that is sensitive to the votes of the citizens (see Felsenthal and Machover 1998, Subsec. 3.3). But we need not consider sensitivity here, because maximizing sensitivity is the same as minimizing the mean majority deficit. Furthermore, in the voting power literature, the voting powers of the different people are often equalized in order to make the voting system fair. Once more, we need not consider this desideratum here – in the settings considered in Sec. 2, each person has the same voting power.

get as close as possible to some ideal size for group deliberation both at the lower and upper tier, then we should let unit size be as close as possible to the number of units.[3] So there is a tension between these three *pro tanto* reasons for the institutional design of a voting system. The relative weight of these reasons will depend on contextual factors.

There is a particular context in which the majoritarian concern does have ample weight. Consider a pre-existing federation with millions of people and with roughly equal-sized states. There is a two-tier decision making structure for certain decisions in which the majority votes of the states are amalgamated through a majority vote at the level of the federation. Think of the Electoral College in a US-style federation but with equal-sized states. Now states may express some dissatisfaction with such a 'Winner-Takes-All' system—a dissatisfaction that is sometimes heard about the US Electoral College. In particular, narrow but persistent minorities at the lower tier may feel that their voice is simply not heard. When there is a majority deficit for an important decision, it is inviting to blame the 'Winner-Takes-All' system. This poses a threat to the legitimacy of the democratic decision-making structure. One might then argue that if only the votes could be recorded in smaller units, say in the districts, then the decision would not have suffered from a majority deficit. Of course a majority deficit could occur within any federal structure, but one might be inclined to think that the smaller the units are, the closer we are to 'the people' and hence we will be more likely to avoid outcomes of votes that suffer from a majority deficit.

Our result shows that this reasoning is incorrect at least for the typical range of unit-size that is under consideration. In the US Electoral College, the winner-takes-all system operates with 50 states and the District of Columbia, whereas a system with congressional districts operates with 435 districts. *Under the assumption of equal-sized states*, the mean majority deficit is increasing as we move from a division into 1 to $\sqrt{n}$ smaller units.[4] For a total voter turnout of n = 122,295,345, as we witnessed in the US presidential elections in 2004, the mean majority deficit increases in m over the range from m = 1,..., 11,059 units.[5] Splitting up the federation into reasonably-sized smaller units to avoid the misgivings about winner-takes-all states in large states may have certain advantages, but it worsens the expectation that votes will be marred by majority deficits and hence should not be defended as a safeguard for the legitimacy of democratic institutions in this respect.

---

[3] Here is a precise argument: Let b be the optimal size for deliberation. Assuming an m-value in the interval $(1, n)$, the size of the units will be $|n/m - b|$ off the optimal size, and the size of the board will be $|m - b|$ off the optimal size. In order to ensure that both groups are not too far off the optimal size, we look for the particular m under which the maximum of $|n/m - b|$ and $|m - b|$ is minimal. That is, we look for the smallest possible epsilon-bound around b such that both the size of the units and the size of the board are within that bound. A few calculations show that this problem is solved for $m = \sqrt{n}$. Partitioning a population into $m = \sqrt{n}$ units is thus in some sense optimal for deliberation at both levels.

[4] Here our argument is based upon the results in Sec. 2.

[5] Voter turnout taken from www.fec.gov/pubrec/fe2004/tables.pdf .

But what if we start from 50+1 unequal-sized states and would move to a system with roughly equal-sized districts? This reflects the political reality in the US. How does this affect the mean majority deficit? We have not been able to derive any analytical generalisations about shifts from unequal-sized units to equal-sized subunits. Since we wish to restrict ourselves in this paper to analytical results, we refer the reader to our future work.

Acknowledgement

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

References

Beisbart, C. and Bovens, L. (2008), A Power Measure Analysis of Amendment 36 in Colorado, *Public Choice* 134: 231–246

Edelman, P. H. (2004). Voting power and at-large representation. *Mathematical Social Sciences*, *47*, 219–232.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power*: *theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Felsenthal, D. S., & Machover, M. (2000). Enlargement of the EU and weighted voting in its Council of Ministers. Voting Power report 01/00, London School of Economics and Political Science, Centre for Philosophy of Natural and Social Science, London; downloadable from http://eprints.lse.ac.uk/archive/00000407.

Grofman, B., & Feld, S. (2005). Thinking about the Political Impacts of the Electoral College. *Public Choice 123*, 1–18.

Hinich, M. J., Mickelsen, R. and Ordeshook, P.C. (1975). The Electoral College vs. a Direct Vote: Policy Bias, Reversals, and Indeterminate Outcomes. *Journal of Mathematical Sociology*, *4*, 3–35.

Morriss, P. (1987/2002), Power. A Philosophical Analysis, Manchester: Manchester University Press second edition 2002.

Życzkowski K, Słomczyński W (2004) Voting in the European Union: The square root system of Penrose and a critical point, Mimeo. http://arxiv.org/ftp/cond-mat/papers/0405/0405396.pdf (checked July 2008).

Appendix 1. Proof of Theorem 1

Let n be any real number larger than 1. The real-extended E[MD] reads as follows:

$$E[\mathrm{MD}] = .5 \times \left(S_n - n \times h(m-1) \times h(n/m-1)\right)$$

with

$$h(l) = \frac{1}{2^l}\frac{\Gamma(l)}{\Gamma(l/2)^2} \ .$$

Consider now E[MD] as a function of m: E[MD] = $f(m)$. We consider that function on the interval $[1,m]$. f has the following property:

$$f(m) = f(n/m) \ .$$

It is useful to define another function g via

$$g(a) = f(a\sqrt{n}) \, .$$

g originates, if m is measured in units of $\sqrt{n}$. g is therefore here considered on the interval $\left[1/\sqrt{n}, \sqrt{n}\right]$. On that interval g is differentiable, as is f on the other interval. We have:

$$g(a) = f(a\sqrt{n}) = f\left(n/\left(a\sqrt{n}\right)\right) = f\left(\sqrt{n}/a\right) = g(1/a) \, .$$

We show that g has a zero derivative at $a = 1$, from which we can follow that f has zero derivative at $m = \sqrt{n}$. For this, let a_k be a series of reals with $a_k \in \left(1, \sqrt{n}\right)$ for every k and assume that the series converges to 1. It follows that, for any k,

$$g(a_k) = g(1/a_k).$$

From the mean value theorem (or Rolle's theorem) it follows that, for each natural number k, there is a $c_k \in [1/a_k, a_k]$, such that $g'(c_k) = 0$. Since the intervals can be made arbitrarily small, it follows that $g'(1) = 0$.

We can also show that g has an extremum at this location, if it is not a straight line in some interval around 1. For this assume that $g'(a)$ is non-negative on some interval $(1,e)$. Consider the interval $(1/e, 1)$. For any $b \in (1/e,1)$ we can calculate $g(b)$ by taking $g(1/b)$. We can therefore obtain the derivative of g at b by deriving $g(1/b)$. This yields:

$$\frac{d}{db} g(1/b) = -\frac{1}{b^2} g'(1/b).$$

As a consequence, in the interval $(1/e,1)$, g has the opposite sign from that on $(1,e)$. Thus, the first derivative changes its sign at 1, which means that there is either a maximum or a minimum at this point. QED.