TI 2014-019/I
Tinbergen Institute Discussion Paper

Solving the Inverse Power Problem in Two-Tier Voting Settings

Matthias Weber

Faculty of Economics and Business, University of Amsterdam, CREED, Tinbergen Institute, the Netherlands.

Solving the Inverse Power Problem in Two-Tier Voting Settings*

Matthias Weber

University of Amsterdam (CREED), Tinbergen Institute

February 6, 2014

Abstract

There are many situations in which different groups make collective decisions by committee voting, where each group is represented by a single person. Theoretical concepts suggest how the voting systems in such committees should be designed, but these abstract rules can usually not be implemented perfectly. To find voting systems that approximate these rules the so called inverse power problem needs to be solved. I introduce a new method to address this problem in two-tier voting settings using the coefficient of variation. This method can easily be applied to a wide variety of settings and rules. After deriving the new method, I illustrate why it is to be preferred over more traditional methods.

Keywords: inverse power problem; indirect voting power; two-tier voting; Penrose's Square Root Rule

*Financial support by grant number 406-11-022 of The Netherlands Organisation for Scientific Research (NWO) is gratefully acknowledged. Thanks for suggestions and comments go to Aaron Kamm, Boris van Leeuwen, Nicola Maaser, Rei Sayag, and Arthur Schram.

1 Introduction

Two-tier voting refers to situations where different groups have to make a collective decision and do so by voting in an assembly of representatives with one representative per group. Many decisions are taken daily through such voting by all kinds of institutions. The best studied case of such two-tier voting is the Council of the European Union,[1] but it is by far not the only institution making use of some sort of two-tier voting – other institutions are for example the UN General Assembly, WTO, OPEC, African Union, German Bundesrat, ECB, and thousands of boards of directors and professional and non-professional associations. While the fiercest debates so far have arisen during the reform of voting rules in the EU Council, the importance of two-tier voting is likely to increase further in the future.[2]

The question of how such two-tier voting systems should be designed is unsolved and can in full generality certainly not be solved. Nevertheless, there are theoretical concepts that provide guidelines, such as for example Penrose's Square Root Rule (Penrose, 1946). Many of these concepts are based on equalizing some concept of (indirect) voting power. These abstract rules can usually not be perfectly implemented. The problem of finding voting systems that approximate these theoretical rules, is called the inverse power problem ('inverse' refers to mapping from a distribution of power to a voting system in contrast to mapping from a voting system to a distribution of power). In this paper, I introduce a method that can be widely applied in many different two-tier voting settings.[3]

I introduce a new method that measures the extent of inequality in indirect voting power of voting systems based on the statistical coefficient of variation. This new method yields indirect voting power that is as equal as possible directly, contrary to more traditional methods that take a detour by first deriving a desired distribution of voting power in the assembly of representatives. It yields better results and it is also more intuitive for beginning researchers and more salient when talking to policy makers than other rules. I derive the new method and alongside a 'classic method' in a

[1] The literature on two-tier voting within the EU includes, among many others, Baldwin and Widgrén (2004), Beisbart et al. (2005), Felsenthal and Machover (2004), Laruelle and Valenciano (2002), Le Breton et al. (2012), Napel and Widgrén (2006), and Sutter (2000).

[2] Globalizaion and the emergence of democracy in many parts of the world make collaboration in supra-national organizations more necessary and easier. Furthermore, modern communication technologies facilitate the organization in interest-groups, clubs, and associations, even when the members are not geographically close.

[3] Solving the inverse power problem is far from trivial, see for example Alon and Edelman (2010), De et al. (2012), Fatima et al. (2008), Kurz (2012), Kurz and Napel (2012), Leech (2003), and De Nijs and Wilmer (2012). It is not the aim of this paper to develop a computationally efficient algorithm to solve the inverse power problem for weighted voting in one particular setting.

setting where equal indirect Banzhaf power is desired, but the new method can also be applied in many different settings – in Appendix A, I show how it can be applied in a setting where equal indirect Shapley-Shubik power is desired; the new method can also be applied to some rules that are based on utility and not on voting power, such as for example an egalitarian rule as in Laruelle and Valenciano, 2010.

This paper is organized as follows. In Section 2, I describe a setting where equal indirect Banzhaf power is desired (Penrose's Square Root Rule). In Section 3, the inverse power problem is described and the new method as well as a classic method to address this problem are derived. Furthermore, I discuss the main advantages of the new method and illustrate the differences with two examples. Section 4 concludes (and Appendix A shows an application of the method in a setting where equal indirect Shapley-Shubik power is desired).

2 Equalizing Indirect Voting Power

In this section, I first introduce some notation. Then, I introduce Penrose's Square Root Rule, which is concerned with equal indirect Banhaf power, and briefly derive its motivation.

2.1 Preliminaries

There are N different groups, each group i consists of n_i individuals. Each group elects one representative. The representatives then come together in an assembly to vote. The representatives are elected through majority voting and they act in the best interest of their group. The voting system that should govern the voting in this assembly of representatives is the focus of most of the two-tier voting literature.

The following notation and definition will be used in the remainder. Coalitions are sets of voters voting in favor of adopting a proposal (yes-voters) or against it (no-voters) and denoted by capital letters (except for N, which represents the number of voters here or in general the number of groups – it is thus best to think of the voting systems as voting systems in an assembly of representatives and of the voters as the representatives now). Thus a coalition is always a subset of $\{1,...,N\}$. Note that a voting system is fully characterized by the set of winning coalitions. Voting systems, i.e. sets of winning coalitions, are denoted by calligraphic letters. A voting system $\mathcal{W}$ is admissible if it satisfies the following conditions ($\mathcal{W}$ is thus a set of (winning) coalitions, which are subsets of 2^N): (i) $\{1,...,N\} \in \mathcal{W}$, (ii) $\emptyset \notin \mathcal{W}$, (iii) if $S \in \mathcal{W}$ then $S^C \notin \mathcal{W}$, (iv) if $S \in \mathcal{W}$ and $S \subseteq T$ then $T \in \mathcal{W}$.[4] The set of admissible voting system contains any 'reasonable' voting system. This set is larger than the set of all weighted voting systems – double majority systems, for example, as used in the EU Council of Ministers can in general not be represented by weighted voting.

2.2 One Rule and its Motivation: Penrose's Square Root Rule

The probably most prominent normative concept of how two-tier voting systems should be designed is the following.[5]

Penrose's Square Root Rule. *The voting power of (the representative of) a group as measured by the Banzhaf index should be proportional to the square root of its population size.*

The main idea of this rule is to make it equally likely for every individual to influence the overall outcome of the two-tier voting procedure, independently of the group she belongs to. The standard motivation of this rule derives from a particular setting, which is described briefly below.[6]

First I give a few very brief definitions in accordance with the literature. If a winning coalition turns into a losing coalition without voter j we say that voter j has a swing.[7] The absolute Banzhaf index of a voter j is defined as the number of possible winning coalitions that turn into losing coalitions without voter j, divided by the total number of possible coalitions. The normalized or relative Banzhaf index is the absolute Banzhaf index normalized so that the sum of the indices of all voters equals one.

The particular set-up that is usually used to motivate Penrose's Square Root Rule is the following. Voting is binary, i.e. a proposal can be either accepted or rejected – we are thus in a take-it-or-leave-it setting (see Laruelle and Valenciano, 2008). Every individual, no matter which group she belongs to, favors the adoption of a proposal

[4] In words this means that the grand coalition (everyone voting for something) is always winning, the empty coalition (nobody voting for something) is always losing, if a coalition is winning the complement is not winning (those not in a winning coalition cannot also form a winning coalition) and if a winning coalition gains additional support it is still winning.

[5] Note that this is just one rule that is used to describe the new method. For a critical discussion of the rule, see Laruelle and Valenciano (2008); see also Laruelle and Valenciano (2005).

[6] The derivation of this rule in the same setting and style in a bit more detail can be found in Turnovec (2009); see also Turnovec et al. (2008).

[7] It does not matter whether one considers only winning coalitions that turn losing if the respective player is removed or also all losing coalitions that turn winning if the respective player is added.

with probability one half, independently of all other individuals.[8] Majority voting takes place within each group and the outcome determines the vote of the representative. The representatives of all groups come together and it is determined according to their votes and the voting system in the assembly of representatives whether a proposal is adopted or rejected.

Denote by Ψ_i^B the absolute Banzhaf power index of an individual in group i arising from majority voting in this group and by Φ_i^B the absolute Banzhaf power index of group i in the assembly of representatives, depending on the voting system in place. Then the probability that an individual in group i has a swing with respect to the overall outcome of the voting procedure (i.e. that she influences with her vote within the group the overall outcome in the assembly of representatives) is Ψ_i^B times Φ_i^B. Thus the probability of influencing the overall outcome is equal for all individuals if $\Psi_i^B \Phi_i^B$ is equal for all individuals or equivalently if

$$\Psi_i^B \Phi_i^B = \alpha \tag{1}$$

for some constant α. Numbering groups from 1 to N and individuals in group i from 1 to n_i, it can easily be shown that equation (1) holds for all i if the normalized Banzhaf index of each group i is equal to

$$\frac{\frac{1}{\Psi_i^B}}{\sum_{j=1}^{N} \frac{1}{\Psi_j^B}}.$$

The normative rule how to design voting systems described here states that the indirect voting power $\Psi_i^B \Phi_i^B$ should be equal for all individuals independently of which group they are in, i.e. that equation (1) should hold. The reason why this is often referred to as square root rule is the following. Ψ_i^B in equation (1) can be approximated by $\sqrt{\frac{2}{\pi n_i}}$, thus equation (1) holds if the Banzhaf indices of the groups are proportional to the square root of population size.[9]

[8]If every voter favors the adoption of a proposal with probability one half independently of everyone else, the absolute Banzhaf index of a voter is the probability that this voter has a swing.

[9]See for example Felsenthal and Machover (1998) or Laruelle and Widgrén (1998). The exact value is $\Psi_i^B = \frac{n_i!}{2^{n_i}((n_i/2)!)^2}$. Usually, researchers use the approximation, which is not a problem for applications where the groups are countries that are easily large enough to make the approximation very good. As the theory can also be applied to small groups, e.g. in companies, boards, clubs, etc., it can sometimes be better to use the exact values. I will not use the approximation in this paper, but still talk about the Banzhaf index being proportional to the square root of group size (working with the exact value or the approximation makes no conceptual difference).

3 Solving the Inverse Power Problem

In this section, I derive a 'classic' and a new method to solve the inverse power problem. I do this in a setting where indirect Banzhaf voting power is of interest, but the method is by no means restricted to such settings, it can similarly be derived and applied in many other settings. I show the motivation and application of the new method in settings where equal indirect Shapley-Shubik power is desired in the appendix.

3.1 A Classic Method

The system of equations (1) usually does not hold exactly for any one voting system. It is thus necessary to find an approximation, i.e. to find a voting system that corresponds as closely as possible to an equal distribution of indirect voting power across all individuals. One way to do this is to take a voting system that minimizes the deviation of the normalized Banzhaf index from the vector that would yield equal voting weights. Taking the euclidean distance as error term, this yields minimizing a term of the form

$$\sum_{i=1}^{N} \left(\frac{\Phi_i^B}{\sum_{j=1}^{N} \Phi_j^B} - \frac{\frac{1}{\Psi_i^B}}{\sum_{j=1}^{N} \frac{1}{\Psi_j^B}} \right)^2 .$$

Remembering that the groups have different sizes and that the idea is to equalize voting power at the individual level, it seems natural to also weigh the error terms by group size. This means that the weight placed on the error term is equal for each individual rather than for each group. This leads to choosing a voting system that minimizes

$$\sum_{i=1}^{N} n_i \left(\frac{\Phi_i^B}{\sum_{j=1}^{N} \Phi_j^B} - \frac{\frac{1}{\Psi_i^B}}{\sum_{j=1}^{N} \frac{1}{\Psi_j^B}} \right)^2 .$$

In order for this error term not to increase with group size or number of groups, one can divide by the total number of individuals. Furthermore, as the deviation is in squared terms, one can take the square root, such that the error term is measured in the 'unit' of indirect voting power rather than in its square. This does not change the outcome and leads to the minimization of[10]

$$err(\Psi^B,\Phi^B) := \sqrt{\frac{1}{\sum_{i=1}^N n_i}\sum_{i=1}^N n_i \left(\frac{\Phi_i^B}{\sum_{j=1}^N \Phi_j^B} - \frac{\frac{1}{\Psi_i^B}}{\sum_{j=1}^N \frac{1}{\Psi_j^B}}\right)^2}. \tag{2}$$

I will refer to minimizing expression (2) over different voting systems (or more generally to using $err(\Psi,\Phi)$ as measure of inequality of a voting system) as the classic method.

3.2 Using the Coefficient of Variation to Solve the Inverse Power Problem

At first sight the classic method seems like a natural choice. Thinking back to the actual goal of this exercise – finding a voting system such that individuals have voting powers as equal as possible – one might prefer a different solution, however. If there is a method that aims at low variation of indirect voting power directly, this method is more natural (using the 'desired' distribution of normalized Banzhaf indices of the voting system in the assembly that results from equal indirect power is then an unnecessary detour). Indeed, such a method exists. The coefficient of variation is a well established statistical concept that measures variation. After introducing this concept briefly, I show how it can be derived in a meaningful way in this voting power setting.

The coefficient of variation in statistics is defined as the ratio of the population standard deviation σ to the population mean μ,

$$cv = \frac{\sigma}{\mu}.$$

It is thus the inverse of the signal-to-noise ratio. The advantage of using the coefficient of variation over the standard deviation is that the standard deviation always has to be understood in the context of the mean (e.g. multiplying all data points by two leads to a higher standard deviation but to the same coefficient of variation). The coefficient of variation is independent of the unit of measurement.

If the system of equations (1) holds, all individuals have equal (indirect) voting power.

[10]The outcome of the minimization procedure does not change, because the number of individuals is fixed in such a comparison. Aside from the advantages of keeping the order of magnitude similar across different numbers of groups and group sizes and keeping the 'unit' of calculation the same, it is also easier to see some similarities and differences to the new procedure I introduce below. Therefore, I choose to mainly use this slightly more complicated expression.

Thus, keeping in mind that the error at the individual level is what one should be interested in, it seems to be straightforward to minimize

$$\sum_{i=1}^{N}\sum_{j=1}^{n_i}\left(\Psi_i^B\Phi_i^B-\alpha\right)^2=\sum_{i=1}^{N}n_i\left(\Psi_i^B\Phi_i^B-\alpha\right)^2. \tag{3}$$

Remember that equal indirect voting power corresponds to equation (1) holding for *any* α. Thus, it is natural to give each voting system its 'best shot', i.e. to let α depend on the voting system (remember that the Φ_i^B also depend on the voting system):

$$\alpha=\arg\min_{\gamma}\sum_{i=1}^{N}n_i\left(\Psi_i^B\Phi_i^B-\gamma\right)^2.$$

It can easily be shown that then

$$\alpha=\frac{1}{\sum_{i=1}^{N}n_i}\sum_{i=1}^{N}n_i\Psi_i^B\Phi_i^B=:\overline{\Psi^B\Phi^B}. \tag{4}$$

Note that $\overline{\Psi^B\Phi^B}$ is the mean of $\Psi^B\Phi^B$ (taken at the individual level).

Minimizing expression (3) with α as in (4) still has some shortcomings.[11] Also here, one would like that the error term does not necessarily increase with group size or number of groups, which can again be solved by dividing by the number of individuals. Furthermore, it again seems desirable to measure the variation of indirect voting power in the same unit as indirect voting power per se rather than in its square, so again, one can take the square root. Finally, it is desirable that the scale used does not change the relevant expressions, i.e. that merely multiplying the indices Φ^B of all groups with a constant or the unit of measurement do not change the outcome. This then also means that it does not matter whether one uses the absolute Banzhaf index Φ^B or the normalized index. This can be achieved through dividing by $\overline{\Psi^B\Phi^B}$. One then arrives at the coefficient of variation (at the individual level), which can thus be seen as a naturally extended and normalized version of expression (3).[12] The new method minimizes the following expression over different voting systems (or more generally,

[11]The trivial 'solution' of all Φ_i equal to zero is not a problem, because all admissible voting systems adopt a proposal with the support of the grand coalition.

[12]Expression (3) with α from (4) is a multiple of the population variance of $\Psi^B\Phi^B$. After dividing by the number of individuals and taking the square root one arrives at the population standard deviation.

uses this expression as a measure of inequality of voting power):

$$cv(\Psi^B\Phi^B) := \frac{\sqrt{\frac{1}{\sum_{i=1}^N n_i}\sum_{i=1}^N n_i\left(\Psi_i^B\Phi_i^B - \overline{\Psi^B\Phi^B}\right)^2}}{\overline{\Psi^B\Phi^B}}. \tag{5}$$

In many cases, there is a trade-off between equality and efficiency, because the voting systems that are the most equal are not necessarily the most efficient ones. If such a trade-off exists, the coefficient of variation as in expression (5) can be readily used as input for a policy maker in its decision function, where this measure of inequality is combined with a measure of efficiency. Such a trade-off does not always exist or is not always of importance where it exists. For example in bargaining committee settings (see e.g. Laruelle and Valenciano, 2008), which are probably the most important settings motivating equal indirect Shapley-Shubik power, efficiency concerns are absent. If one wants to select a voting system from a set of voting systems exclusively from a point of view of equalizing indirect voting power, the formulas for the two methods can be written down as in the next paragraph.

It is of course possible to restrict the set of voting systems from which one selects. Instead of minimizing expressions (2) or (5) over all admissible voting systems, one can also minimize over a subset of these. Such subsets could for example be all weighted voting systems, all weighted voting systems satisfying some additional conditions, all double majority voting systems, or all voting system with a certain number of winning coalitions.[13] Denoting by $\mathbf{V}$ the set of all admissible voting systems and by $\mathbf{W} \subseteq \mathbf{V}$ a subset from which we want to choose, we arrive at the following formulas that select a voting system approximating equal indirect Banzhaf voting power (now explicitly writing down the dependence of Φ^B on the voting system $\mathcal{W}$). Using the classic method of minimizing the squared deviation from the desired vector of normalized Banzhaf indices per group (weighted by group size), the recommended voting system is

$$\begin{aligned}\mathcal{V}_{classic} &= \underset{\mathcal{W}\in\mathbf{W}}{\arg\min}\ err(\Psi^B, \Phi^B(\mathcal{W})) \\ &= \underset{\mathcal{W}\in\mathbf{W}}{\arg\min}\sqrt{\frac{1}{\sum_{i=1}^N n_i}\sum_{i=1}^N n_i\left(\frac{\Phi_i^B(\mathcal{W})}{\sum_{j=1}^N \Phi_j^B(\mathcal{W})} - \frac{\frac{1}{\Psi_i^B}}{\sum_{j=1}^N \frac{1}{\Psi_j^B}}\right)^2}.\end{aligned} \tag{6}$$

Using the new method of minimizing the coefficient of variation of indirect voting

[13]Minimizing expression (2) over all weighted voting systems where the voting weights are proportional to the square root of the population size ('choosing the quota') is referred to as Jagiellonian compromise, see Słomczyński and Życzkowski (2006) and Słomczyński and Życzkowski (2010).

power (as measured by the Banzhaf index), the recommended voting system is

$$\begin{aligned}\mathcal{V}_{new} &= \underset{\mathcal{W}\in\mathbf{W}}{\arg\min}\; cv(\Psi^B, \Phi^B(\mathcal{W})) \\ &= \underset{\mathcal{W}\in\mathbf{W}}{\arg\min} \frac{\sqrt{\frac{1}{\sum_{i=1}^N n_i}\sum_{i=1}^N n_i\left(\Psi_i^B\Phi_i^B(\mathcal{W}) - \overline{\Psi^B\Phi^B(\mathcal{W})}\right)^2}}{\overline{\Psi^B\Phi^B(\mathcal{W})}}.\end{aligned} \tag{7}$$

3.3 Differences between the Methods and Advantages of the New Method

The two methods can lead to different outcomes and the differences can be non-negligible. A first application of the new method to address the inverse power problem is in Weber (2014). There, the two methods do not give the same outcome in a considerable number of cases.[14]

The new method is to be preferred over the classic one. It appropriately measures variation of voting power at the individual level directly. Taking the detour via the 'power equalizing distribution' as in the classic method can lead to inaccuracies. This is best illustrated by examples, in particular by the example in Section 3.3.1.

On top of measuring power at the individual level, the coefficient of variation is a relative measure. This means for example that the coefficient of variation judges the indirect voting powers of two equally sized groups of individuals of 0.01 and 0.02 to exhibit more variation than 0.08 and 0.09. This relative difference is the quantity that is of interest. The absolute numbers of voting power are often not very telling and usually arguments involving indirect voting power involve relative numbers. For example in the discussions about reforms of the voting system in the EU, arguments were made that a voting system is unfair, because country *X*' citizens' indirect voting power is four times as large as the one of country *Y*' citizens. Arguments were usually not about one country's citizens having indirect voting power that is 0.000002 higher than that of other countries' citizens. If this is how one argues the quantities involved when solving the inverse power problem should be chosen accordingly. The example in Section 3.3.2 illustrates this difference in a bit more detail.

This is the major advantage of the new method: It measures the correct quantity. Furthermore, there are two other advantages. First, this new method is more intuitive. If one wants a system where voting power at the individual level is as equal as possible,

[14]The environment in which the inverse power problem is addressed in Weber (2014) is quite specific as the number of groups and the sets of voting systems from which always one is selected are very small.

the most intuitive solution for beginning researchers or practitioners that want to become familiar with the theory is to take the voting system where this indirect voting power has low variation across individuals (rather than making the detour of minimizing errors with respect to an 'equalizing' distribution). Second, the new method is more salient. When someone who uses the classic method talks to a policy maker, the outcome of the minimization procedure could be understood to represent more than equal indirect voting power (e.g. it could be mistaken to also include efficiency considerations or other trade-offs, because many people might talk about an 'optimal distribution' rather than about an 'indirect Banzhaf power equality approximating distribution').

In the remainder of this section I will illustrate how the new method is more suitable than the classic method with two (hypothetical) examples. The first example shows differences between the methods when the means of indirect Banzhaf voting power are equal between the voting systems, but standard deviations are different. In the second example standard deviations are equal, but means are different. A further example, where indirect Shapley-Shubik power is of interest, can be found in Appendix A.2.

3.3.1 First Example

There are six groups, numbered from 1 to 6. Groups 1 and 2 have 10 members each, the other groups have 5 members. This means that in the first stage (the election of the representatives) individuals have voting power $\Psi^B_{1,2} = 0.2460938$ and $\Psi^B_{3,4,5,6} = 0.375$, respectively. Indirect voting power would be equal across all individuals if the voting systems were such that

$$\frac{\Phi^B_{1,2}}{\sum_{i=1}^6 \Phi^B_i} = 0.2162162 \qquad \text{and} \qquad \frac{\Phi^B_{3,4,5,6}}{\sum_{i=1}^6 \Phi^B_i} = 0.1418919.$$

Now we compare two (hypothetical) voting systems $\mathcal{W}_1$ and $\mathcal{W}_2$. The voting systems are such that the normalized Banzhaf indices are as follows:

$$\begin{aligned}
\frac{\Phi^B_1(\mathcal{W}_1)}{\sum_{i=1}^6 \Phi^B_i(\mathcal{W}_1)} &= 0.2162162 + 0.05, & \frac{\Phi^B_2(\mathcal{W}_1)}{\sum_{i=1}^6 \Phi^B_i(\mathcal{W}_1)} &= 0.2162162 - 0.05,\\
\frac{\Phi^B_{3,4,5,6}(\mathcal{W}_1)}{\sum_{i=1}^6 \Phi^B_i(\mathcal{W}_1)} &= 0.1418919, \text{ and} & \frac{\Phi^B_{1,2}(\mathcal{W}_2)}{\sum_{i=1}^6 \Phi^B_i(\mathcal{W}_2)} &= 0.2162162\\
\frac{\Phi^B_{3,4}(\mathcal{W}_2)}{\sum_{i=1}^6 \Phi^B_i(\mathcal{W}_2)} &= 0.1418919 + 0.05, & \frac{\Phi^B_{5,6}(\mathcal{W}_2)}{\sum_{i=1}^6 \Phi^B_i(\mathcal{W}_2)} &= 0.1418919 - 0.05.
\end{aligned}$$

Assume for simplicity and to have a nice illustration that normalized and absolute Banzhaf indices are equal. Now we can calculate the indirect voting power of each individual, depending on the group she is in. This yields

$$\begin{aligned} \Psi_1^B \Phi_1^B(\mathcal{W}_1) &= 0.06551414, & \Psi_2^B \Phi_2^B(\mathcal{W}_1) &= 0.04090477, \\ \Psi_{3,4,5,6}^B \Phi_{3,4,5,6}^B(\mathcal{W}_1) &= 0.05320946, \quad \text{and} & \Psi_{1,2}^B \Phi_{1,2}^B(\mathcal{W}_2) &= 0.05320946, \\ \Psi_{3,4}^B \Phi_{3,4}^B(\mathcal{W}_2) &= 0.07195946, & \Psi_{5,6}^B \Phi_{5,6}^B(\mathcal{W}_2) &= 0.03445946. \end{aligned}$$

One can easily see that using the classic method, both voting systems would be judged to be equally equal (the error term equals $\frac{1}{20}$ for both voting systems). If one looks carefully at the indirect voting power, this does not seem justified, though. Under both voting systems, there are 20 individuals with indirect voting power 0.05320946, which is also the mean of indirect voting power under both voting systems. Under both voting systems, there are 10 individuals with higher voting power and 10 with lower power. The absolute difference between the higher value of voting power and the middle value is always equal to the difference between the middle value and the lower value; just that these differences are higher under the second voting system than under the first one. It is $cv(\Psi^B \Phi^B(\mathcal{W}_1)) = 0.1635184$ and $cv(\Psi^B \Phi^B(\mathcal{W}_2)) = 0.249171$. The new method thus selects $\mathcal{W}_1$.

3.3.2 Second Example

In the last example, the mean of indirect voting power was equal under both voting systems. Considering the standard deviation would in these examples thus lead to the same outcome as considering the coefficient of variation. In the following example, the standard deviation of indirect voting power under two systems is equal, while the coefficient of variation is different.

There are four groups. The first group has 9 members while the other groups have 3 members each. The Banzhaf power indices in the first stage (when electing the representative) are $\Psi_1^B = 0.2734375$ and $\Psi_{2,3,4}^B = 0.5$, respectively. Indirect voting power would be equal for all individuals if the normalized Banzhaf indices in the assembly of representatives were

$$\frac{\Phi_1^B}{\sum_{i=1}^4 \Phi_i^B} = 0.3786982 \qquad \text{and} \qquad \frac{\Phi_{2,3,4}^B}{\sum_{i=1}^4 \Phi_i^B} = 0.2071006.$$

Now we compare again two voting systems, $\mathcal{W}_1$ and $\mathcal{W}_2$. Assume again for simplicity that absolute and normalized Banzhaf indices are equal and assume that the two voting systems are such that

$$\frac{\Phi_1^B(\mathcal{W}_1)}{\sum_{i=1}^4 \Phi_i^B(\mathcal{W}_1)} = 0.3786982 - 0.09, \qquad \frac{\Phi_{2,3,4}^B(\mathcal{W}_1)}{\sum_{i=1}^4 \Phi_i^B(\mathcal{W}_1)} = 0.2071006 + 0.03,$$
$$\frac{\Phi_1^B(\mathcal{W}_2)}{\sum_{i=1}^4 \Phi_i^B(\mathcal{W}_2)} = 0.3786982 + 0.09, \quad \text{and} \quad \frac{\Phi_{2,3,4}^B(\mathcal{W}_2)}{\sum_{i=1}^4 \Phi_i^B(\mathcal{W}_2)} = 0.2071006 - 0.03.$$

Which of these two voting systems do the two methods select? For the classic method, the two systems approximate equal indirect voting power equally well. This can easily be seen as follows. The terms in parentheses in expression (6) are for both voting systems always either -0.09 or $+0.09$ for the parts referring to the large group and either $+0.03$ or -0.03 for the parts referring to the other groups. As only the squares of these values enter expression (6), these two voting systems are 'equally equal' for the classic method. The new method, in contrast, does make a difference between these two voting systems. The indirect voting power of each individual is under the first voting system

$$\Psi_1^B \Phi_1^B(\mathcal{W}_1) = 0.07894092 \qquad \text{and} \qquad \Psi_{2,3,4}^B \Phi_{2,3,4}^B(\mathcal{W}_1) = 0.1185503,$$

and under the second voting system

$$\Psi_1^B \Phi_1^B(\mathcal{W}_2) = 0.1281597 \qquad \text{and} \qquad \Psi_{2,3,4}^B \Phi_{2,3,4}^B(\mathcal{W}_2) = 0.0885503.$$

Remember that exactly half of the individuals are in the large group. While under the first voting system an individual in the half of the population with more power holds 1.50176 times as much indirect voting power as an individual in the other half of the population, this ratio is only 1.447309 under the second voting system.[15] The coefficient of variation is 0.2005627 for $\mathcal{W}_1$ and 0.182776 for $\mathcal{W}_2$. Thus, the new method selects $\mathcal{W}_2$.

4 Concluding Remarks

I have introduced a method to address the inverse power problem in two-tier voting settings that is based on the statistical coefficient of variation. After deriving it in a setting where equal indirect Banzhaf voting power is desired, I have shown why this

[15]The absolute values of the differences $\Psi_1^B \Phi_1^B(\mathcal{W}_1) - \Psi_{2,3,4}^B \Phi_{2,3,4}^B(\mathcal{W}_1)$ and $\Psi_1^B \Phi_1^B(\mathcal{W}_2) - \Psi_{2,3,4}^B \Phi_{2,3,4}^B(\mathcal{W}_2)$ are equal.

method is to be preferred over a more classic method. This new method can be applied in many different settings, including all settings where indirect voting power is to be equalized. It can also be applied to rules involving utility instead of voting power like egalitarianism as in Laruelle and Valenciano (2010). In Appendix A, I show how the method can be applied in settings where indirect Shapley-Shubik power is the quantity of interest and illustrate the advantages of the new method with an example in such a setting.

The main advantage of the new method is that it measures the correct quantity, which is the variation of indirect voting power. Further advantages of the method are that it is more intuitive for researchers and practitioners that start getting acquainted with the theory and that it is more salient when talking to policy makers. The method can be used to select voting systems from a predetermined set to approximate equal indirect voting power as well as possible – this seems, among other cases, reasonable in bargaining committee settings, where efficiency concerns are absent. The coefficient of variation could also be used as input into a decision function or social welfare function of a policy maker when some trade-offs exist, such as for example an equity-efficiency trade-off.

Furthermore, in many cases, researchers need to address the inverse power problem while it is not the focus of their research project. I hope that the new method I have presented will be of help to them in their future research.

References

Alon, N. and Edelman, P. H. (2010). The inverse Banzhaf problem. *Social Choice and Welfare*, 34(3):371–377.

Baldwin, R. and Widgrén, M. (2004). Winners and losers under various dual majority rules for the EU Council of Ministers.

Beisbart, C., Bovens, L., and Hartmann, S. (2005). A utilitarian assessment of alternative decision rules in the Council of Ministers. *European Union Politics*, 6(4):395–418.

De, A., Diakonikolas, I., and Servedio, R. (2012). The inverse Shapley value problem. In *Automata, Languages, and Programming*, pages 266–277. Springer.

De Nijs, F. and Wilmer, D. (2012). Evaluation and improvement of Laruelle-Widgrén inverse Banzhaf approximation. *arXiv preprint arXiv:1206.1145*.

Fatima, S., Wooldridge, M., and Jennings, N. R. (2008). An anytime approximation method for the inverse Shapley value problem. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, pages 935–942. International Foundation for Autonomous Agents and Multiagent Systems.

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

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

Kurz, S. (2012). On the inverse power index problem. *Optimization*, 61(8):989–1011.

Kurz, S. and Napel, S. (2012). Heuristic and exact solutions to the inverse power index problem for small voting bodies. *arXiv preprint arXiv:1202.6245*.

Laruelle, A. and Valenciano, F. (2002). Inequality among EU citizens in the EU's Council decision procedure. *European Journal of Political Economy*, 18(3):475–498.

Laruelle, A. and Valenciano, F. (2005). Assessing success and decisiveness in voting situations. *Social Choice and Welfare*, 24(1):171–197.

Laruelle, A. and Valenciano, F. (2008). *Voting and collective decision-making: Bargaining and power*. Cambridge University Press.

Laruelle, A. and Valenciano, F. (2010). Egalitarianism and utilitarianism in committees of representatives. *Social Choice and Welfare*, 35(2):221–243.

Laruelle, A. and Widgrén, M. (1998). Is the allocation of voting power among EU states fair? *Public Choice*, 94(3):317–339.

Le Breton, M., Montero, M., and Zaporozhets, V. (2012). Voting power in the EU Council of Ministers and fair decision making in distributive politics. *Mathematical Social Sciences*, 63(2):159–173.

Leech, D. (2003). Power indices as an aid to institutional design: the generalised apportionment problem. *Jahrbuch für Neue Politische Okonomie*, 22:107–121.

Maaser, N. and Napel, S. (2007). Equal representation in two-tier voting systems. *Social Choice and Welfare*, 28(3):401–420.

Napel, S. and Widgrén, M. (2006). The inter-institutional distribution of power in EU codecision. *Social Choice and Welfare*, 27(1):129–154.

Penrose, L. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, 109(1):53–57.

Shapley, L. and Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, 48(03):787–792.

Shapley, L. S. (1953). A value for n-person games. *The Annals of Mathematical Statistics*, (28):307–317.

Słomczyński, W. and Życzkowski, K. (2006). Penrose voting system and optimal quota. Technical report, arXiv preprint physics/0610271.

Słomczyński, W. and Życzkowski, K. (2010). Jagiellonian Compromise: An alternative voting system for the Council of the European Union. *Institutional Design and Voting Power in the European Union. Ashgate, London*.

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(4):433–449.

Turnovec, F. (2009). Fairness and squareness: Fair decision making rules in the EU council? Technical report, Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies.

Turnovec, F., Mercik, J., and Mazurkiewicz, M. (2008). Power indices methodology: decisiveness, pivots and swings. *Power, Freedom and Voting, Essays in Honor of Manfred J. Holler (Braham, M. and F. Steffen, eds.)*, pages 23–37.

Weber, M. (2014). Choosing voting systems behind the veil of ignorance: A two-tier voting experiment. *CREED Working Paper, University of Amsterdam*.

A Appendix

A.1 Application to a Different Problem – Equal Indirect Shapley-Shubik Power

As mentioned, the new method can be applied to different problems. Here, I show how it can be used if equal indirect voting power as measured by the Shapley-Shubik index is desired. This rule can be motivated in a bargaining setting or in a probabilistic setting.[16] While the bargaining motivation has some advantages, the probabilistic motivation is much closer to the motivation of Penrose's Square Root Rule in Section 2. Therefore, I only show the probabilistic motivation here. The new method of addressing the inverse power problem can be used independently of which of these motivations is used. Much of the derivation of the probabilistic motivation and also of the derivation of the methods can be done similarly to Section 2.2. It is thus kept very brief.

A prominent normative concept of how two-tier voting systems should be designed that is different from the one described above is the following:[17]

Proportional Shapley-Shubik Power. *The voting power of (the representative of) a group as measured by the Shapley-Shubik index should be proportional to its population size.*

In contrast to the derivation of Penrose's Square Root Rule, we now assume that all voters differ in the strength of their feelings over the issue at stake. One can then order all voters from strong preference to strong dislike. In general, voter j is in a pivotal position if the coalition of voters that would like the adaption of a proposal more strongly than voter j does not have the power to pass it, while the coalition of voters that would like the adoption of the proposal less (dislike it more) does not have the power to block it. A voter in a pivotal position is thought to have decisive influence over the outcome of the voting process.

Let me state the relevant definitions, in accordance with the literature. Let $(i_1, ..., i_M)$ be a permutation of voters (voters are numbered from 1 to M, the voting system – i.e. the set of winning coalitions – is denoted by $\mathcal{W}$). If voter j's position in the permutation is i_k, then voter j is pivotal if $\{i_1, ..., i_{k-1}\} \notin \mathcal{W}$ and $\{i_1, ..., i_k\} \in \mathcal{W}$. The Shapley-Shubik power index of voter j is the number of permutations in which j is pivotal divided by the total number of permutations $n!$. Note that the sum of the Shapely-Shubik indices of all voters equals one and that this index represents the probability of being pivotal if all permutations (that can be seen as preference orderings) are equally likely.

Denote by Ψ_i^S the Shapley-Shubik power index of an individual in group i and by Φ_i^S the Shapley-Shubik index of group i in the assembly of representatives, depending on the voting system in the assembly. Assuming that all permutations are equally likely in both stages of the voting procedure (and assuming again majority voting at the group level), the probability that an individual in group i is pivotal in the first stage while the

[16] The Shapley-Shubik index originates from cooperative game theory (Shapley and Shubik, 1954, Shapley, 1953).

[17] For a more detailed derivation of the concept in a very similar style, see Turnovec (2009).

representative of group i is pivotal in the second stage is $\Psi_i^S \Phi_i^S$. Thus the probability of influencing the overall outcome in this sense is equal for all individuals if

$$\Psi_i^S \Phi_i^S = \alpha \tag{8}$$

for some constant α. As the Shapley-Shubik indices of all voters sum up to one, it is $\Psi_i^S = \frac{1}{n_i}$. Thus, equation (8) holds for all i if the Shapley-Shubik index of each group i is equal to $\frac{n_i}{\sum_{j=1}^N n_j}$, i.e. if the Shapley-Shubik indices of the groups are proportional to their sizes.

This equality is in general not feasible. Proceeding as in Section 3, it is again possible to calculate deviations of the prescribed weights or to look for a voting system with small variation of indirect voting power (as now measured by $\Psi_i^S \Phi_i^S$) directly. Then we arrive at the same formulas as in Section 3, except for the superscripts. With the error term of the classic method $err(\Psi^S, \Phi^S)$ and the coefficient of variation $cv(\Psi^S, \Phi^S)$ defined exactly as in Section 3, the formulas to select voting systems from a predetermined set of voting systems are as follows.[18] Using the classic method of minimizing the squared deviation from the desired vector of Shapley-Shubik indices per group (weighted by group size), the recommended voting system is

$$\begin{aligned} \mathcal{V}_{classic} &= \underset{\mathcal{W} \in \mathbf{W}}{\arg\min}\ err(\Psi^S, \Phi^S(\mathcal{W})) \\ &= \underset{\mathcal{W} \in \mathbf{W}}{\arg\min} \sqrt{\frac{1}{\sum_{i=1}^N n_i} \sum_{i=1}^N n_i \left(\frac{\Phi_i^S(\mathcal{W})}{\sum_{j=1}^N \Phi_j^S(\mathcal{W})} - \frac{\frac{1}{\Psi_i^S}}{\sum_{j=1}^N \frac{1}{\Psi_j^S}} \right)^2}. \end{aligned} \tag{9}$$

Using the new method of minimizing the coefficient of variation of indirect voting power (as measured by the Shapley-Shubik index), the recommended voting system

[18] For the particular case where indirect equal Shapley-Shubik power is desired, one could simplify the formulas, using that $\Psi_i^S = \frac{1}{n_i}$ and, where possible, leaving out multiplication factors and square roots. This makes the classic method equivalent to minimizing $\sum_{i=1}^N n_i \left(\Phi_i^S(\mathcal{W}) - \frac{n_i}{\sum_{j=1}^N n_j} \right)^2$ and the new method equivalent to minimizing

$$\sum_{i=1}^N n_i \left(\frac{\Phi_i^S(\mathcal{W})}{n_i} - \frac{1}{\sum_{j=1}^N n_j} \right)^2 \text{ or } \sum_{i=1}^N \frac{1}{n_i} \left(\Phi_i^S(\mathcal{W}) - \frac{n_i}{\sum_{j=1}^N n_j} \right)^2 \text{ or } \sum_{i=1}^N n_i \left(\Psi_i^S \Phi_i^S(\mathcal{W}) - \frac{1}{\sum_{j=1}^N n_j} \right)^2,$$

where the last expression is similar to the expression Maaser and Napel (2007) use in a closely related setting. In this setting, using the coefficient of variation is also always equal to using the standard deviation, because the indirect voting power of all individulas always sums up to one. As one of the aims of this paper is to show that the new method can be applied generally in different settings, I do not use these rearrangements.

is

$$
\begin{aligned}
\mathcal{V}_{new} &= \operatorname*{arg\,min}_{\mathcal{W}\in\mathbf{W}} \; cv(\Psi^S, \Phi^S(\mathcal{W})) \\
&= \operatorname*{arg\,min}_{\mathcal{W}\in\mathbf{W}} \frac{\sqrt{\frac{1}{\sum_{i=1}^N n_i}\sum_{i=1}^N n_i \left(\Psi_i^S\Phi_i^S(\mathcal{W}) - \overline{\Psi^S\Phi^S(\mathcal{W})}\right)^2}}{\overline{\Psi^S\Phi^S(\mathcal{W})}}.
\end{aligned} \tag{10}
$$

The same arguments concerning the advantages of the new method as in Section 3.3 are valid, i.e. it is based on measuring the right quantity and it is more intuitive and more salient. That it can make a difference which method is used in this setting and that the new method is to be preferred is illustrated in the following example.

A.2 Example in a 'Shapley-Shubik Setting'

Assume that we are in a setting where equal indirect Shapley-Shubik power is desired, independently of whether this desire is motivated probabilistically or from a bargaining point of view. There are five groups, the first group has 10 members, the other four groups have 5 members each. The Shapley-Shubik power in the first stage is for each individual $\Psi_1^S = 0.1$ or $\Psi_{2,3,4,5}^S = 0.2$, respectively. This means that indirect Shapley-Shubik power is equal across all individuals for $\Phi_1^S(\mathcal{W}) = \frac{1}{3}$ and $\Phi_{2,3,4,5}^S(\mathcal{W}) = \frac{1}{6}$. Now we want to compare again two (hypothetical) voting systems. The first voting system is such that

$$
\Phi_1^S(\mathcal{W}_1) = \frac{1}{3} - \frac{1}{6} = \frac{1}{6}, \qquad \Phi_{2,3}^S(\mathcal{W}_1) = \frac{1}{6}, \text{ and} \qquad \Phi_{4,5}^S(\mathcal{W}_1) = \frac{1}{6} + \frac{1}{12} = \frac{1}{4},
$$

while the second voting system is such that

$$
\Phi_1^S(\mathcal{W}_2) = \frac{1}{3}, \qquad \Phi_{2,3}^S(\mathcal{W}_2) = \frac{1}{6} - \frac{1}{12} = \frac{1}{12}, \text{ and} \qquad \Phi_{4,5}^S(\mathcal{W}_2) = \frac{1}{6} + \frac{1}{12} = \frac{1}{4}.
$$

With these voting systems, indirect Shapley-Shubik powers are as follows:

$$
\Psi_1^S\Phi_1^S(\mathcal{W}_1) = \frac{1}{60}, \qquad \Psi_{2,3}^S\Phi_{2,3}^S(\mathcal{W}_1) = \frac{1}{30}, \text{ and} \qquad \Psi_{4,5}^S\Phi_{4,5}^S(\mathcal{W}_1) = \frac{1}{20},
$$

$$
\Psi_1^S\Phi_1^S(\mathcal{W}_2) = \frac{1}{30}, \qquad \Psi_{2,3}^S\Phi_{2,3}^S(\mathcal{W}_2) = \frac{1}{60}, \text{ and} \qquad \Psi_{4,5}^S\Phi_{4,5}^S(\mathcal{W}_2) = \frac{1}{20}.
$$

Thus, under both voting systems, there are 10 individuals with indirect Shapley-Shubik power $\frac{1}{60}$, 10 with $\frac{1}{30}$, and 10 with $\frac{1}{20}$. Thus, there is no difference in terms of how equal they are and accordingly the new method judges them to approximate equal indirect power equally well. The classic method does distinguish between them, though, with a higher error term for the first voting system.[19]

[19] The error terms are $err(\Psi^S, \Phi^S(\mathcal{W}_1)) = 10(\frac{1}{6} - \frac{1}{3})^2 + 5(\frac{1}{4} - \frac{1}{6})^2 + 5(\frac{1}{4} - \frac{1}{6})^2 = \frac{25}{72}$ for the first and $err(\Psi^S, \Phi^S(\mathcal{W}_2)) = 5(\frac{1}{12} - \frac{1}{6})^2 + 5(\frac{1}{12} - \frac{1}{6})^2 + 5(\frac{1}{4} - \frac{1}{6})^2 + 5(\frac{1}{4} - \frac{1}{6})^2 = \frac{10}{72}$ for the second voting system.