



# JENA ECONOMIC RESEARCH PAPERS



# 2012 – 045

## **Heuristic and exact solutions to the inverse power index problem for small voting bodies**

by

**Sascha Kurz  
Stefan Napel**

[www.jenecon.de](http://www.jenecon.de)

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact [markus.pasche@uni-jena.de](mailto:markus.pasche@uni-jena.de).

Impressum:

Friedrich Schiller University Jena  
Carl-Zeiss-Str. 3  
D-07743 Jena  
[www.uni-jena.de](http://www.uni-jena.de)

Max Planck Institute of Economics  
Kahlaische Str. 10  
D-07745 Jena  
[www.econ.mpg.de](http://www.econ.mpg.de)

© by the author.

# Heuristic and exact solutions to the inverse power index problem for small voting bodies\*

Sascha Kurz and Stefan Napel  
University of Bayreuth, 95440 Bayreuth, Germany  
{sascha.kurz,stefan.napel}@uni-bayreuth.de

July 17, 2012

## Abstract

Power indices are mappings that quantify the influence of the members of a voting body on collective decisions a priori. Their nonlinearity and discontinuity makes it difficult to compute inverse images, i.e., to determine a voting system which induces a power distribution as close as possible to a desired one. The paper considers approximations to this inverse problem for the Penrose-Banzhaf index by hill-climbing algorithms and exact solutions which are obtained by enumeration and integer linear programming techniques. They are compared to the results of three simple solution heuristics. The heuristics perform well in absolute terms but can be improved upon very considerably in relative terms. The findings complement known asymptotic results for large voting bodies and may improve termination criteria for local search algorithms.

**Keywords** electoral systems; simple games; weighted voting games; square root rule; Penrose limit theorem; Penrose-Banzhaf index; institutional design

**JEL Classification** C61; C71; D02

## 1 Introduction

Collective decision rules and, in particular, heterogeneous voting weights for members of a committee, council, or shareholder meeting translate into *influence* on collective decisions in a nonlinear and even discontinuous fashion. This can be seen, for instance, by considering a decision quota of  $q = 50\%$  and players  $i \in N = \{1, 2, 3\}$  whose voting weights are given by either the vector (i)  $w = (w_1, w_2, w_3) = (33.\bar{3}, 33.\bar{3}, 33.\bar{3})$ , (ii)  $w' = (50 - \varepsilon, 48 + \varepsilon, 2)$ , or (iii)  $w'' = (50 + \varepsilon, 48 - \varepsilon, 2)$  for small  $\varepsilon > 0$ . The major weight change from  $w$  to  $w'$  does not affect possibilities to form a *winning coalition* at all, where coalition  $S \subseteq N$  is called winning if the cumulative weight of its members exceeds the quota. Namely,  $S$  is winning if and only if  $|S| \geq 2$ . By symmetry, the distribution of influence can a priori be expected to equal  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  for either of the voting systems described by  $(q; w)$  and  $(q; w')$ . The minor change from  $w'$  to  $w''$ , in contrast, renders voter 1 a dictator with associated power distribution  $(1, 0, 0)$ .

Social scientists, philosophers and mathematicians have investigated various *voting power indices* which try to quantify the a priori distribution of influence on committee decisions. The Shapley-Shubik index (Shapley and Shubik 1954) and the Penrose-Banzhaf index (PBI) (Penrose 1946; Banzhaf 1965) are most prominent, but by far not the only ones.<sup>1</sup> They help researchers clarify the non-trivial a priori

---

\*We thank Nicola Maaser, Friedrich Pukelsheim and two anonymous referees for constructive comments on an earlier version. Napel acknowledges the generous hospitality of the Max Planck Institute of Economics, Jena, where parts of this paper were written.

<sup>1</sup>See Felsenthal and Machover (1998) or Laruelle and Valenciano (2008) for overviews.

power implications of different voting weight assignments to a wider audience. The combinatorial nature of weighted voting systems can easily mislead the general public's intuition and even that of political practitioners. For instance, it was apparently not noted that the voting weights of the original six members of the European Economic Community, in use from 1958 to 1973, rendered Luxembourg a *null player* whenever the EEC Council applied its qualified majority rule, i.e., the country's vote was irrelevant for the collective decision. The public discussion – very heated in, but not restricted to, Poland and Germany – in the wake of the 2007 EU summit at which new voting rules for the EU Council were agreed reflected persistent confusion between voting weights and power.

Even to specialists, the discrete nature of voting rules still poses challenges. This is true, in particular, for the optimal design of a voting system. Certain normative desiderata, such as the equal representation of bottom-tier voters in a two-tier voting system, often call for a specific distribution of voting power. It is then a non-trivial exercise to find a deterministic voting rule that comes as close as possible to inducing the desired a priori power distribution.<sup>2</sup> Simple hill-climbing algorithms, such as the ones considered by Laruelle and Widgrén (1998), Leech (2002a, 2002b, 2003), Aziz et al. (2007), or Fatima et al. (2008), deliver excellent results for many instances of this so-called *inverse power index problem* but have rarely been evaluated in a systematic fashion.<sup>3</sup> One can neither rule out that only a *local minimum* of the distance between the desired and the induced power vector has been identified. Nor are *bounds* known on the possible gap to a globally optimal voting rule. The latter might involve the intersection of several one-dimensional  $(q; w)$ -rules and, therefore, need not even be a feasible result of the applied search algorithm.

For very big  $n$  the distinctions between voting weight and voting power tend to become negligible under simple majority rule if each voter's relative weight is small and vanishing in the limit (and one stays outside a class of somewhat pathological examples).<sup>4</sup> Corresponding asymptotic results for  $n \rightarrow \infty$  which render the inverse problem trivial have first been suggested in the work of Lionel S. Penrose (1946, 1952). Rigorous investigations of asymptotic proportionality of voting weight and power, which is often referred to as the *Penrose limit theorem*, have only more recently been provided by Neyman (1982), Lindner and Machover (2004), Chang et al. (2006), and Lindner and Owen (2007).

Motivated by qualified majority voting in the EU, Słomczyński and Życzkowski (2006, 2007) have identified an elegant way to approximately solve the inverse problem for  $n$  still very distant from infinity if the decision quota  $q$  is a free parameter and all individual voting power targets are small compared to  $\sqrt{n}$  times their quadratic mean. Their heuristic suggestion is particularly appealing because it avoids discrepancies between voting weight and power. Namely, approximate proportionality between the normalized weight vector  $w = (w_1, \dots, w_n)$  and the induced PBI  $B(q; w)$  is achieved even for  $n \ll \infty$  when the quota is set to  $q^* = \frac{1}{2}(1 + \sqrt{\sum w_i^2})$ . A desired power vector  $\beta$  can hence approximately be induced simply by choosing  $w = \beta$  and then calculating  $q^*$ . Because the rule  $(q^*; w)$  is simple and minimizes unwanted confusion between voting weight and power, it has motivated the prominent “Jagiellonian Compromise” (also known as *double square root voting system*) in the discussion of future voting rules for the EU Council (see, e.g., Kirsch et al. 2007).

Whether the decision quota  $q$  is a free parameter, so that Słomczyński and Życzkowski's optimal  $q^*$  indeed can be chosen, or not, depends on the application at hand. Even if it can, the lack of bounds on how well the  $(q^*; \beta)$ -heuristic performs relative to the respective globally optimal solution to the inverse problem provides motivation for further research. Knowing more about the quality

<sup>2</sup>Non-deterministic rules such as *random dictatorship* or also *random quota rules* (Dubey and Shapley 1979, sec. 5) can easily solve the problem, but are generally not regarded as satisfactory.

<sup>3</sup>We thank an anonymous referee for pointing out to us that Laruelle and Widgrén (1998, sec. 6) contains the earliest published specification of an iterative algorithm for solving the inverse problem. This algorithm has recently been evaluated and improved by de Nijs and Wilmer (2012). Lucas (1992, p. 44) is one of the first to describe the basic approach to implementing a target PBI vector by a suitable voting rule; while Nurmi (1982, p. 206f) seems to be the first to state this institutional design problem and to discuss its relevance.

<sup>4</sup>The case in which the relative weights of some voters do *not* vanish as  $n \rightarrow \infty$  – so that a few large voters or *atomic players* stick out in an “ocean” of infinitesimal voters (defining a so-called *oceanic game*) – reduces to the analysis of a modified voting body  $v'$  for the PBI. It involves only the atomic voters and a quota  $q'$  which is obtained from  $q$  by subtracting half of the aggregate weight of the infinitesimal voters (see Dubey and Shapley 1979, sec. 9). If the number of atomic voters is small then the corresponding inverse problem is exactly of the type investigated here.

of the  $(q^*; \beta)$ -heuristic is especially important for situations in which the heuristic can be expected to perform rather badly. Specifically, the derivation of  $q^*$  is based on a continuous approximation of the fundamentally discrete distribution of the cumulative weight of a random coalition. Its use is problematic when this approximation is inaccurate. This pertains particularly – but not exclusively – to “small” voting bodies with few members.

For a given number  $n$  of players, the set of different binary voting systems or *simple games* is finite. This finiteness entails that many desired power distributions cannot be approximated too well. Nontrivially, this remains true even for large  $n$ : recent results by Alon and Edelman (2010) imply the existence of sequences of desired power distributions  $\{\beta^n\}_{n=1,2,\dots}$  which stay at least a constant positive distance away (in the  $\|\cdot\|_1$ -norm, which adds up deviations from target for all voters  $i = 1, \dots, n$ ) from *any* Penrose-Banzhaf power distribution.

The finiteness of the set of simple games at the same time suggests a trivial algorithm for solving the inverse problem: enumerate all systems  $v$  with  $n$  players, compute the respective power distribution – say, the PBI  $B(v)$  – and then pick a game  $v^*$  that induces the smallest achievable difference between ideal vector  $\beta$  and  $B(v)$  according to a suitable measure of distance (e.g., the metric induced by a particular vector norm).

To this end, a growing literature has investigated methods for the efficient enumeration of voting systems (see, e.g., Keijzer 2009; Keijzer et al. 2010; Kurz 2012a). But, up to now, even the *number* of complete simple games (and also of weighted voting games) is unknown for  $n > 9$ . So enumeration works only for voting bodies with very few members. Exact solutions to the inverse problem can, fortunately, also be obtained for somewhat larger  $n$  by *integer linear programming* (ILP) techniques. Such an approach was recently presented in Kurz (2012b). It stands in the tradition of earlier applications of ILP to electoral systems, as discussed in Grilli di Cortona et al. (1999), Pennisi et al. (2007) or Ricca et al. (2012).

The key idea is to translate the problem of finding a game  $v$  whose PBI vector has a distance no greater than a given value  $\alpha \geq 0$  from the target vector  $\beta$  into a set of linear inequality constraints plus the integer requirement that each coalition is either winning or losing, and then to use efficient ILP solver software in order to check if a solution exists. A bisection process on  $\alpha$ , which consecutively halves the range of tentative optimal values of  $\alpha$  that have not yet been ruled out by the computations, can be stopped after finitely many iterations (namely, once  $|\alpha_{t+1} - \alpha_t|$  has fallen below the theoretical minimum distance between any two distinct  $n$ -voter PBI vectors).

This paper draws on complete enumeration, the indicated ILP bisection method, and standard hill-climbing algorithms in order to evaluate the accuracy of three heuristic solutions to the inverse problem for the PBI. The first heuristic simply combines  $w = \beta$  with  $q^\circ = 50\%$ ; the second combines it with the “optimal quota”  $q^*$  derived by Słomczyński and Życzkowski (2007); the third uses  $\bar{q} = \frac{1}{2} + 1/\sqrt{\pi n}$ . The latter quota is the average of  $q^*$  computed over a set of  $\beta$ -vectors which is of particular interest for the egalitarian design of two-tier voting systems (Słomczyński and Życzkowski 2011).

We compute differences between the respective heuristic and exact solutions for three different metrics and a comprehensive grid of conceivable target vectors with up to  $n = 7$  voters. We study rather time-consuming approximations of the exact solutions for a large sample of grid points for  $8 \leq n \leq 20$  as well as selected real-world examples based on the so-called *Penrose square root rule* and EU population figures. The results allow the estimation of bounds for the accuracy of a heuristic solution and can provide informed termination criteria for conventional local search algorithms. This may be useful in applications where a specific voting power distribution is sought for a moderate number of council delegates, committee members, or business shareholders. We also analyze the significant magnitude of relative errors that the mentioned heuristics can produce even for large  $n$  in pathological cases.

In the following Section 2 we first introduce binary voting systems and their basic properties. The inverse power index problem is then formalized in Section 3, along with a brief discussion of the worst-case bounds which follow from Alon and Edelman’s (2010) results. In Section 4 we present the design of our comparative investigation. The corresponding results are the topic of Section 5. We conclude in Section 6.

## 2 Binary voting systems

We consider *binary* voting systems, i.e., each voter  $i \in N = \{1, \dots, n\}$  casts a binary vote (e.g., “yes” or “no”) and this determines a binary collective decision. Such a situation can mathematically be described by a Boolean function  $v: 2^N \rightarrow \{0, 1\}$ , where  $2^N$  denotes the set of subsets of  $N$ . A coalition  $S \subseteq N$  can, e.g., be interpreted as the set of “yes”-voters for a particular (unspecified) proposal.

**Definition 1** A simple game is a monotone Boolean function, i.e., a mapping  $v: 2^N \rightarrow \{0, 1\}$  with  $v(S) \leq v(T)$  for all  $S \subseteq T$ , which additionally satisfies  $v(\emptyset) = 0$ ,  $v(N) = 1$ .

Coalition  $S \subseteq N$  is called *winning* if  $v(S) = 1$ , and otherwise losing.  $S \subseteq N$  is called a *minimal winning coalition* if it is winning and all proper subsets are losing. A simple game is uniquely characterized by its set of minimal winning coalitions. We refer the reader to Taylor and Zwicker (1999) for a detailed introduction to simple games.

As illustration consider the simple game  $v$  which is characterized by the set  $\{\{1\}, \{2, 3\}\}$  of its minimal winning coalitions. Taking all supersets of the minimal winning coalitions yields  $\{\{1\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}, \{2, 3\}\}$  as the set of all winning coalitions.<sup>5</sup> The remaining subsets are losing.

The monotonicity imposed in Definition 1 is a very weak requirement. By introducing Isbell’s *desirability relation*, i.e.,  $i \sqsupseteq j$  for two voters  $i, j \in N$  if and only if  $v(\{i\} \cup S \setminus \{j\}) \geq v(S)$  is satisfied for all  $\{j\} \subseteq S \subseteq N \setminus \{i\}$  (see, e.g., Isbell 1956), one can define a particularly relevant subclass of the set of all simple games  $\mathcal{S}$ :

**Definition 2** A simple game  $v$  is called *complete* (also called *directed*) if the binary relation  $\sqsupseteq$  is a total preorder, i.e.,

- (1)  $i \sqsupseteq i$  for all  $i \in N$ ,
- (2)  $i \sqsupseteq j$  or  $j \sqsupseteq i$  (including “ $i \sqsupseteq j$  and  $j \sqsupseteq i$ ”) for all  $i, j \in N$ , and
- (3)  $i \sqsupseteq j$ ,  $j \sqsupseteq h$  implies  $i \sqsupseteq h$  for all  $i, j, h \in N$ .

In our small example one can easily check that  $1 \sqsupseteq 2 \sqsupseteq 3$ . So  $v \in \mathcal{C}$  where  $\mathcal{C} \subset \mathcal{S}$  denotes the set of all complete simple games. Note that also  $3 \sqsupseteq 2$ , i.e., voters 2 and 3 are equally desirable.

Many binary voting systems which are used in practice belong to a further refinement of  $\mathcal{S}$ :

**Definition 3** A simple game  $v$  is *weighted* if there exist non-negative weights  $w_i \in \mathbb{R}_{\geq 0}$  and a positive quota  $q \in \mathbb{R}_{> 0}$  such that  $v(S) = 1$  if and only if  $\sum_{i \in S} w_i \geq q$ .

A weighted representation of our small illustrative example is given by  $(q; w) = [2; 2, 1, 1]$ . We call weighted simple games *weighted voting games* and denote their collection by  $\mathcal{W}$ . Every weighted voting game is complete while not every complete simple game is weighted, i.e.,  $\mathcal{W} \subset \mathcal{C} \subset \mathcal{S}$ .<sup>6</sup> The presently known enumeration results for the three considered classes of binary voting systems are summarized in Table 1 (up to isomorphisms). See, e.g., Kurz (2012a) for details.

There are several equivalent representations of binary voting structures besides Boolean functions and lists of minimum winning coalitions. Simple games can, for instance, be described as independent sets in a graph, and Carreras and Freixas (1996) have introduced a very efficient matrix parameterization of  $\mathcal{C}$ . Our computation algorithms will exploit yet another possibility. Namely, we use that voting systems can be represented as points of a polyhedron which have integer coordinates only: for each

<sup>5</sup>All minimal winning and hence all winning coalitions in a *proper* simple game have non-empty intersection. So the considered example  $v$  is not proper. We do not rule out improper games. They can play a meaningful role even in majoritarian democratic systems, e.g., if a sufficiently big parliamentary minority can install a special investigation committee or call for a referendum.

<sup>6</sup>But each complete simple game and even each simple game can be represented as the *intersection* of  $1 \leq k < \infty$  weighted voting games. The minimal number  $k$  of weighted voting games is called the *dimension* of the simple game in question (see, e.g., Taylor and Zwicker 1999, Deineko and Woeginger 2006).

$n$	1	2	3	4	5	6	7	8	9	10
$\#\mathcal{S}$	1	3	8	28	208	16351	$>4.7 \cdot 10^8$	$>1.3 \cdot 10^{18}$	$>2.7 \cdot 10^{36}$	unknown
$\#\mathcal{C}$	1	3	8	25	117	1171	44313	16175188	284432730174	unknown
$\#\mathcal{W}$	1	3	8	25	117	1111	29373	2730164	989913344	unknown

Table 1: Number of distinct simple games, complete simple games, and weighted voting games

$S \subseteq N$  define  $0 \leq x_S \leq 1$  and add the constraints  $x_\emptyset = 0$ ,  $x_N = 1$ , and  $x_S \leq x_T$  for all  $\emptyset \subseteq S \subset T \subseteq N$ . Each integer solution  $(x_\emptyset, x_{\{1\}}, \dots, x_N) \in \{0, 1\}^{2^n}$  of this system of linear inequalities corresponds one-to-one to a simple game (with  $v(S) = x_S$ ). Complete simple games and weighted voting games are described by additional constraints and auxiliary variables  $w_i \in \mathbb{R}_{\geq 0}$  for the weights. For instance, the inequality

$$w(S) = \sum_{i \in S} w_i \leq q - 1 + M \cdot x_S \tag{1}$$

implies that the sum of weights of a coalition  $S$  is at most  $q - 1$  if coalition  $S$  is losing, i.e.,  $x_S = 0$ . For a sufficiently large  $M$ , which can be computed explicitly, inequality (1) is automatically satisfied for  $x_S = 1$ . Similarly the constraint

$$\sum_{i \in S} w_i \geq q + M(1 - x_S) \tag{2}$$

forces winning coalitions to have a weight sum exceeding or meeting the quota. Because each weighted voting game admits a representation  $(q; w_1, \dots, w_n)$  where  $w(S) \leq w(T) - 1$  for all losing coalitions  $S$  and all winning coalitions  $T$ , inequalities (1) and (2) capture the weightedness requirement for a simple game described by  $(x_\emptyset, x_{\{1\}}, \dots, x_N)$ .

### 3 The inverse power index problem

Power indices are mappings from a set of feasible voting structures, such as  $\mathcal{S}$  or  $\mathcal{W}$ , to non-negative real vectors which are meant to quantify the influence of the members of a voting body on collective decisions. The inverse power index problem consists in finding a voting system, e.g.,  $(q; w) \in \mathcal{W}$ , which induces a power distribution as close as possible to a desired one. More formally, for a given number  $n$  of voters, the general *inverse power index problem* involves a set  $\Gamma$  of feasible voting structures for  $n$  players, a power index  $\phi: \Gamma \rightarrow \mathbb{R}_{\geq 0}^n$ , a desired power distribution  $\beta \in \mathbb{R}_{\geq 0}^n$ , and a metric  $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  which measures the deviation between two power vectors. Of course,  $d(x, y) = \|x - y\|$  is a suitable choice for any vector norm  $\|\cdot\|$ . Given these ingredients the inverse power index problem amounts to finding a solution to the minimization problem

$$\min_{v \in \Gamma} d(\phi(v), \beta). \tag{3}$$

In this paper, we consider the special instances of this problem where  $\Gamma \in \{\mathcal{S}, \mathcal{C}, \mathcal{W}\}$ . We include  $\mathcal{S}$  and  $\mathcal{C}$  because they are significantly larger domains for  $n \geq 5$  (see Table 1) and some prominent real-world electoral systems fail to correspond to weighted voting games. Examples include the current voting rules (Treaty of Nice) and the future ones (Treaty of Lisbon) for the EU Council, which require majorities in more than one dimension (e.g., the Nice rules call for 255 out of 345 votes, 14 out of 27 member states, and 62% of EU population). We take the (normalized) Penrose-Banzhaf index  $B(v)$  as the voting power index of interest.

**Definition 4** For a given  $n$ -player simple game  $v$  the absolute Penrose-Banzhaf index  $B'_i(v)$  for player

$i$  is defined as

$$B'_i(v) = \frac{1}{2^{n-1}} \cdot \sum_{\emptyset \subseteq S \subseteq N \setminus \{i\}} v(S \cup \{i\}) - v(S).$$

The (normalized) Penrose-Banzhaf index (PBI)  $B_i(v)$  for player  $i$  is defined as

$$B_i(v) = \frac{B'_i(v)}{\sum_{j=1}^n B'_j(v)}.$$

Our distance computations will be based on the  $\|\cdot\|_1$ -norm (i.e., the sum of deviations between  $B_i(v)$  and  $\beta_i$  for all players  $i$ ), the  $\|\cdot\|_\infty$ -norm (i.e., the maximum deviation), and a weighted version of the former. Section 4 will provide more details.

To the best of our knowledge, there exists only one (non-trivial) non-approximative result on how well the inverse problem can be solved for the PBI in the worst case. For completeness and later reference we include this rather recent finding by Alon and Edelman (2010) here. It considers a given game  $v$  with  $n$  players in which  $1 - \varepsilon$  of the total (normalized) PBI is concentrated amongst  $k < n$  “major” players. Alon and Edelman then provide a construction for a game  $\tilde{v}$  such that the worth  $\tilde{v}(S)$  of a coalition  $S$  depends only on  $T = S \cap \{1, \dots, k\}$ , i.e., the subset of major players in  $S$  fully determines whether  $S$  is winning in game  $\tilde{v}$  or not.<sup>7</sup> One can easily observe that  $\tilde{v}$  is a simple game and  $B_i(\tilde{v}) = 0$  for all  $i > k$ .<sup>8</sup> The essential finding of Alon and Edelman then is that the deviation  $\|B(v) - B(\tilde{v})\|_1$  is bounded from above by a function which depends on  $\varepsilon$  and the number  $k$  of major players only. Considering the  $k$ -player simple game  $v'$ , which arises from  $\tilde{v}$  by removing the  $n - k$  null players, their result can be stated as follows:

**Theorem 1 (Alon-Edelman)** Consider the simple game  $v$  with players  $N = \{1, \dots, k, \dots, n\}$  and  $0 < \varepsilon < \frac{1}{k+1}$ . If  $\sum_{i=k+1}^n B_i(v) \leq \varepsilon$ , then there exists a simple game  $v'$  with  $k$  voters such that

$$\sum_{i=1}^k |B_i(v) - B_i(v')| + \sum_{i=k+1}^n B_i(v) \leq \frac{(2k+1)\varepsilon}{1 - (k+1)\varepsilon} + \varepsilon.$$

This result is very useful for obtaining lower bounds on distances in the context of the inverse problem because one may suitably reduce the problem from  $n$  to  $k$  players and only make an error with the indicated bound. Specifically, let  $\beta = (\beta_1, \dots, \beta_n)$  be a desired power distribution and  $k, \varepsilon$  be constants satisfying  $0 < \varepsilon < \frac{1}{k+1}$ . Let us denote the unknown  $n$ -player simple game whose PBI has smallest distance to  $\beta$  by  $v^*$ . One can then bound  $\|B(v^*) - \beta\|_1$  by distinguishing two cases.

First, suppose that  $\sum_{i=k+1}^n B_i(v^*) \geq \varepsilon$ . In this case, we can only apply Theorem 1 if the inequality happens to be tight. However, we know that the deviations between  $B_i(v^*)$  and  $\beta_i$  for players  $k+1, \dots, n$  are at least as big as

$$\left| \sum_{i=k+1}^n B_i(v^*) - \sum_{i=k+1}^n \beta_i \right|.$$

This can range from  $1 - \sum_{i=k+1}^n \beta_i$  when  $\sum_{i=k+1}^n B_i(v^*) = 1$  to  $|\varepsilon - \sum_{i=k+1}^n \beta_i|$  when  $\sum_{i=k+1}^n B_i(v^*) = \varepsilon$  in the considered case. Analogous reasoning applies to the additional deviations for players  $1, \dots, k$ , and we thus have the bound

$$\|B(v^*) - \beta\|_1 \geq \min_{1 \geq x \geq \varepsilon} \left| 1 - x - \sum_{i=1}^k \beta_i \right| + \left| x - \sum_{i=k+1}^n \beta_i \right| =: l_1.$$

<sup>7</sup>Specifically, one sets  $\tilde{v}(S) = 1$  if and only if  $\sum_{U \subseteq \{k+1, \dots, n\}} v(T \cup U) \geq \frac{2^{n-k}}{2}$ . So the coalition  $T$  of major players – and hence all supersets  $S' = T \cup U$  that are obtained by adding different coalitions  $U$  of “minor” players – is winning in  $\tilde{v}$  if and only if a majority of the latter coalitions  $S'$  are winning in  $v$ .

<sup>8</sup>Moreover, we have  $\tilde{v} \in \mathcal{C}$  if  $v \in \mathcal{C}$  and  $\tilde{v} \in \mathcal{W}$  if  $v \in \mathcal{W}$ , i.e., the construction respects completeness or weightedness of the simple game in question.

In the second case, i.e., when  $\sum_{i=k+1}^n B_i(v^*) < \varepsilon$ , Theorem 1 applies. It tells us that there exists some  $k$ -player simple game  $v'$  with

$$\sum_{i=1}^k |B_i(v^*) - B_i(v')| + \sum_{i=k+1}^n |B_i(v^*) - 0| \leq \frac{(2k+1)\varepsilon}{1-(k+1)\varepsilon} + \varepsilon.$$

If we use  $\tilde{v}$  to denote the  $n$ -player game that extends  $v'$  by adding  $k+1, \dots, n$  as null players (so that  $B(\tilde{v}) = (B_1(v'), \dots, B_k(v'), 0, \dots, 0)$ ), this can also be written as

$$\|B(v^*) - B(\tilde{v})\|_1 \leq \frac{(2k+1)\varepsilon}{1-(k+1)\varepsilon} + \varepsilon. \quad (4)$$

Now if we solve the  $k$ -player inverse problem for the (typically non-normalized)  $k$ -vector  $\beta' = (\beta_1, \dots, \beta_k)$  which coincides with the first  $k$  components of  $\beta$ , then the resulting minimal distance  $\varepsilon'$  is a lower bound for the distance between  $B(\tilde{v})$  and  $\beta$ , i.e.,

$$\|B(\tilde{v}) - \beta\|_1 \geq \varepsilon'. \quad (5)$$

We can then appeal to the triangle inequality for metric  $d_1(x, y) = \|x - y\|_1$  and conclude

$$\|B(v^*) - \beta\|_1 \geq \|B(\tilde{v}) - \beta\|_1 - \|B(\tilde{v}) - B(v^*)\|_1 \geq \varepsilon' - \frac{(2k+1)\varepsilon}{1-(k+1)\varepsilon} + \varepsilon := l_2$$

from (4) and (5). Thus, in either of the two cases we have  $\|B(v^*) - \beta\|_1 \geq \min(l_1, l_2)$ .

Let us illustrate this by an example and suppose that one seeks to find a voting game with a power distribution as close as possible to  $\beta^n = (0.75, 0.25, 0, \dots, 0) \in \mathbb{R}_{\geq 0}^n$  for  $n \geq 2$ . We will show  $\|B(v) - \beta^n\|_1 \geq \frac{1}{9}$  for all  $n$ -player simple games  $v$ .<sup>9</sup> To this end we choose  $k = 2$  and  $\varepsilon = \frac{1}{18}$ . If  $\sum_{i=3}^n B_i(v^*) \geq \varepsilon$  for the unknown distance-minimizing game  $v^*$  then we have

$$\min_{1 \geq x \geq \varepsilon} \left| 1 - x - \sum_{i=1}^k \beta_i \right| + \left| x - \sum_{i=k+1}^n \beta_i \right| = \min_{1 \geq x \geq \varepsilon} 2x = 2\varepsilon = \frac{1}{9}.$$

So in this case we have  $\|B(v^*) - \beta^n\|_1 \geq \frac{1}{9}$ . In the other case of  $\sum_{i=3}^n B_i(v^*) < \varepsilon$ , we solve the inverse power index problem for  $k = 2$  players and  $\beta' = \beta^2 = (0.75, 0.25)$ . Since the only possible 2-player PBI vectors are given by  $\{(1, 0), (\frac{1}{2}, \frac{1}{2}), (0, 1)\}$  we have a minimal deviation of  $\varepsilon' = \frac{1}{2}$ . Because  $\frac{(2k+1)\varepsilon}{1-(k+1)\varepsilon} + \varepsilon = \frac{7}{18}$  we have  $l_2 = \frac{1}{9}$  and again conclude  $\|B(v^*) - \beta^n\|_1 \geq \frac{1}{9}$ . Hence,  $\beta^n$  cannot be approximated by the PBI of a simple game with an  $\|\cdot\|_1$ -error less than  $\frac{1}{9}$ . The latter is the sharpest possible bound obtainable from Theorem 1. It can be improved computationally to slightly more than  $\frac{14}{37}$  for  $n \leq 11$  on  $\mathcal{S}$  and for  $n \leq 16$  on  $\mathcal{C}$  and  $\mathcal{W}$  (see Kurz 2012b).

## 4 Design of the computational investigation

When the inverse problem arises in political applications of constitutional design, PBI vectors  $\beta$  which are proportional to the square root of a population size vector  $p$  play an elevated role. The reason is that – under the probabilistic assumptions which underlie the PBI – a binary voting system  $v$  with  $B(v) = \beta$  and

$$\beta_i = \frac{\sqrt{p_i}}{\sum_{j=1}^n \sqrt{p_j}} \quad (6)$$

<sup>9</sup>Such artificially constructed target vectors  $\beta^n$ , for which Alon and Edelman's results have bite, may not be of much practical relevance. But they indicate the problems of requiring any fixed level of accuracy in the stopping rule of a local search algorithm. Moreover, one can conceive of real-world enterprises in which the majority and minority partners indeed seek to split voting power 3 : 1 and want to render all  $n - 2$  other stakeholders null players.



would equalize the voting power of citizens in a *two-tier system* in which  $n$  delegates adopt the bottom-tier majority opinion of their respective constituency  $i \in \{1, \dots, n\}$  and then cast a  $w_i$ -weighted vote in a top-tier assembly (e.g., the EU Council). See Penrose (1946), Felsenthal and Machover (1998), Kaniovski (2008) or Kurz et al. (2012) for details. In our computations we will consider this *Penrose square root rule* for varying  $n$  and some historical population figures in order to select target vector examples which have a specific political motivation.

In principle, however, *any* vector in  $\mathbb{R}_{\geq 0}^n$  whose entries sum up to 1 might be a desired power distribution  $\beta$ . For instance, partners of a non-profit R&D joint venture might have made relative financial contributions of  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9})$  and possibly want to align a priori voting power in the directorate to this vector as well as possible. Ideally, for a given number  $n$  of voters, one would like to compare the exact and heuristic solutions to the inverse problem for *all* possible normalized target vectors  $\beta \in \Delta(n-1)$ , where  $\Delta(n-1)$  denotes the  $n-1$ -dimensional unit simplex. This is computationally infeasible. We, however, complement our analysis of politically motivated square root vectors by vectors  $\beta$  from a discrete subset of  $\Delta(n-1)$ , namely a finite grid on  $\Delta(n-1)$  with step size 0.01. We also resort to approximations of the exact solution when  $n$  is too large.

We will compare the (approximated) exact solution of the inverse problem on domain  $\mathcal{S}$ ,  $\mathcal{C}$ , or  $\mathcal{W}$  for a given desired PBI  $\beta$  with three different heuristics. These stay in the class  $\mathcal{W}$  of weighted voting games and have in common that voting weights are set equal to the desired voting power, i.e.,  $w = \beta$ . They pick a distinct quota, and hence typically a different voting system  $v \in \mathcal{W}$ .

The first heuristic – referred to as the *50%-heuristic* – just chooses  $q^\circ = \frac{1}{2}$ . Simple majority is arguably the most common majority rule in practice. The 50%-heuristic simply picks it and ignores the potentially large discrepancies between voting weight and voting power that can arise. This can be motivated by the Penrose limit theorem when  $n$  is at least moderately big (see fn. 4 however).

The second, more sophisticated heuristic has been suggested by Słomczyński and Życzkowski (2006, 2007). Their motivation was to implement PBI vectors proportional to the square root of population sizes in the European Union, but the heuristic applies to arbitrary target vectors. Namely, the *q\*-heuristic* selects the quota

$$q^* = \frac{1}{2} \cdot \left(1 + \sqrt{\sum_i w_i^2}\right)$$

for an arbitrary  $w = \beta \in \Delta(n-1)$ . Słomczyński and Życzkowski (2007) derive this quota by considering the random weight  $W$  which is accumulated if all coalitions  $S \subseteq N$  are *equiprobable*, as the PBI's probabilistic justifications suppose. Equiprobability at the level of coalitions is equivalent to assuming that each voter  $i \in \{1, \dots, n\}$  joins the formed coalition independently of the others with probability  $\frac{1}{2}$ . The mean of  $W$  hence is  $\mu = \sum_{i=1}^n \frac{1}{2} w_i = \frac{1}{2}$  and its variance is  $\sigma^2 = \frac{1}{4} \sum_{i=1}^n w_i^2$ . Being the sum of independent bounded random variables,  $W$  is approximately normally distributed *if*  $n$  is sufficiently large and each of the weights is sufficiently small.<sup>10</sup> Assuming that this is the case and, therefore, that the discrete random variable  $W$  can be replaced by the continuous one  $\tilde{W}$ , the inflection point of the corresponding normal density of  $\tilde{W}$  is located at  $q^* = \mu + \sigma$ . Since the second derivative of  $\tilde{W}$ 's density vanishes at  $q^*$ , one can approximate the density by a linear function with reasonably high accuracy. This linear approximation then allows to establish approximate proportionality of  $B(q^*; w)$  and  $w$ . We refer to Słomczyński and Życzkowski (2007) for details.

Our final heuristic, which we will refer to as the *q̄-heuristic*, replaces  $q^*$  by

$$\bar{q} = \frac{1}{2} + \frac{1}{\sqrt{\pi n}}.$$

This quota approximates the expected value of  $q^*$  when  $\beta$  is proportional to the component-wise square root of a population size vector  $p = (p_1, \dots, p_n)$  which is drawn from a flat Dirichlet distribution (see Słomczyński and Życzkowski 2011). The motivation for computing such an average is the following:

---

<sup>10</sup>A key technical requirement is that  $w_j \ll \sqrt{\sum w_i^2}$  for all  $j \in N$ , i.e.,  $w_j \sqrt{n}$  is sufficiently smaller than the quadratic mean of the weights  $\sqrt{\frac{1}{n} \sum w_i^2}$ .

even though the  $q^*$ -heuristic can approximate the Penrose square root rule (6) very transparently for a given population distribution  $p$ , frequent changes in the population would call not only for frequent changes of the prescribed voting weights  $w$  but also of the quota  $q^*$ . That current voting weights in the EU already recur to population figures, which are updated on an annual basis, suggests that weight changes may be regarded as unproblematic. A varying decision threshold – perhaps  $q = 65\%$  in one year,  $q' = 61\%$  in the next, then  $q'' = 67\%$ , etc. – however seems politically less palatable. It may then make sense to average  $q^*$  over a wide range of values for  $w = \beta \propto \sqrt{p}$ , and the  $\bar{q}$ -heuristic simply assumes that all population distributions  $p \in \Delta(n-1)$  are equally likely.<sup>11</sup> Because  $\bar{q} \rightarrow \frac{1}{2}$  as  $n \rightarrow \infty$ , the 50%-heuristic is the limit of the  $\bar{q}$ -heuristic and can be viewed as an approximation of it for not too small  $n$ .

Let us remark that investigations by Kurth (2008) have called attention to numerical problems when heuristics which involve irrational voting weights and quotas, as the  $q^*$  or  $\bar{q}$ -heuristics commonly do, are implemented. Rounding after, e.g., 4 decimal places can result in voting systems which differ significantly from what was intended. Because it is impractical to deal with weights of a hundred decimal places or more, it is attractive to work with the underlying Boolean functions or integer points of a suitable polyhedron as long as possible, and to determine minimal integer weights  $w$  and a quota  $q$  which efficiently represent a given  $v \in \mathcal{W}$  when needed.<sup>12</sup> We use this approach here whenever possible, and refer the interested reader to Freixas and Molinero (2009, 2010), Freixas and Kurz (2011), or Kurz (2012a).

We calculate the globally optimal solution to the inverse problem for a given target PBI  $\beta$  by complete enumeration of the elements in the respective class of binary voting systems for  $n \leq 7$  (see Table 1). For larger  $n$ , we mostly focus on approximations of the exact solution. These are obtained either by a hill-climbing algorithm or, preferably, by ILP techniques. How the latter are used is explained in the Appendix in detail. The implemented ILP-based bisection algorithm would yield globally optimal solutions when given enough running time and memory. We interrupted it for efficiency reasons whenever a desired precision had been reached. The key idea of the ILP-based approach is to consider the integer polyhedron which contains all simple games whose PBI is less than a given factor  $\alpha > 0$  away from the desired vector  $\beta$ . It can be checked by using readily available ILP solver software if this polyhedron is empty. Then, no such game exists and  $\alpha$  needs to be raised. If not,  $\alpha$  can be lowered. The minimal level of  $\alpha$  (or an approximation with desired precision) together with the corresponding voting systems, can thus be found by the bisection method: namely, by iteratively halving the interval defined by the best lower and upper bounds that have been computed so far. Pseudo-code of the algorithm is provided in Appendix A.

In evaluating the quality of the mentioned heuristics, we consider distances to the desired power vector,  $\beta$ , and to the globally optimal one,  $B(v^*)$ , in three different metrics. The first one is the metric  $d_1(x, y) = \|x - y\|_1 = \sum_{i=1}^n |x_i - y_i|$  induced by the  $\|\cdot\|_1$ -norm, which is also considered in Theorem 1. The second is the metric induced by the  $\|\cdot\|_\infty$ -norm, i.e.,  $d_\infty(x, y) = \|x - y\|_\infty = \max_{i \in \{1, \dots, n\}} |x_i - y_i|$ . We refrain from also considering the Euclidean metric induced by the  $\|\cdot\|_2$ -norm, which has been considered, e.g., by Słomczyński and Życzkowski (2007). The reason is that this would turn the ILP formulation of the inverse power index problem into a *binary non-linear programming* one. This would be considerably harder to solve and add relatively little information because  $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$  for all  $x \in \mathbb{R}^n$ .

More interesting, in our view, is a variation of  $d_1$  which takes the Bernoulli model that underlies the PBI and Penrose's square root rule seriously. This model assumes that all bottom-tier voters in constituency  $i \in \{1, \dots, n\}$  cast a “yes” or “no” vote equiprobably and independently of all others. The probability for one of  $p_i$  individual voters in constituency  $i$  to be *pivotal* for the constituency's aggregate decision – i.e., to induce the  $i$ -delegate at the top-tier council to cast voting weight  $w_i$  in

<sup>11</sup>The expected value of the  $p$ -specific optimal quotas  $q^*(p)$  for a particular (e.g., Dirichlet) distribution of  $p$ , of course, need not coincide with the quota that is optimal when  $p$  is treated as a random variable. Stochastic optimization techniques are likely to yield a somewhat better  $q$ -heuristic than the one suggested by Słomczyński and Życzkowski (2011).

<sup>12</sup>A *minimal* integer representation of a weighted voting game has the advantage that the PBI and other power indices can be computed particularly quickly.

favor of “yes” by individually voting “yes”, and “no” by voting “no” – is approximately  $\sqrt{2/(\pi p_i)}$ . The joint probability of a given voter being pivotal in his or her constituency  $i$  and of this constituency being pivotal at the top tier is hence  $B_i(v) \cdot \sqrt{2/(\pi p_i)}$ . This is why the square root PBI vector in equation (6) equalizes the indirect influence of citizens on collective decisions across constituencies. If one now weights any deviation between (i) the probability for a given voter in constituency  $i$  to be doubly pivotal and (ii) the egalitarian ideal of  $\beta_i \cdot \sqrt{2/(\pi p_i)}$  with  $\beta_i = \sqrt{p_i} / \sum_{j=1}^n \sqrt{p_j}$  equally, then the total misrepresentation associated with the top-tier voting system  $v$  amounts to

$$\sum_{i=1}^n p_i \cdot |\beta_i - B_i(v)| \cdot \sqrt{2/(\pi p_i)} = c \cdot \sum_{i=1}^n \sqrt{p_i} \cdot |\beta_i - B_i(v)|$$

for  $c > 0$ . Whenever the desired vector  $\beta$  is derived from Penrose’s square root rule and a vector  $p$  which represents EU population data, we will, therefore, also consider the variation of metric  $d_1$  which weights absolute deviations by the square root of relative population, i.e., study the metric<sup>13</sup>

$$d_1'(x, y) = \sum_{i=1}^n \sqrt{\frac{p_i}{\sum_{j=1}^n p_j}} \cdot |x_i - y_i|.$$

## 5 Computational results

In this section we present our numerical results. Subsection 5.1 considers the EU Council of Ministers as a prototype of a real-world weighted voting system. We then look at the entire discretized space of possible power distributions for  $n \leq 7$  and random samples thereof in Subsection 5.2. In order to study analytically how deviations between simple heuristics and actual optimization depend on  $n$ , we investigate a particular parametric example in Subsection 5.3.

Exact solutions to the inverse power index problem that are reported in Tables 2–4 have been obtained using our ILP-based bisection algorithm, as described above and in Appendix A. For Table 5 we have used exhaustive enumeration of all possible Banzhaf vectors for  $n \leq 7$ , and resorted to approximations obtained by hill-climbing algorithms for  $8 \leq n \leq 20$ . The bisection part of the ILP approach was implemented in C++ while we used the ILOG CPLEX Interactive Optimizer 12.4.0.0 in order to solve individual integer linear programming problems. The hill-climbing algorithms used in Subsection 5.2 were implemented in C++. We employed a Quad-Core AMD Opteron processor with 2700 Mhz, 132 GB RAM, and a cache size of 512 KB on a 64-bit Linux system as our hardware.

### 5.1 Examples of real-world weighted voting systems

We first consider the (EEC or EC or) EU Council of Ministers in the years 1958, 1973, 1981, 1986, 1995, 2006, and 2011 with respectively  $n \in \{6, 9, 10, 12, 15, 25, 27\}$  members as examples. The historical population data for  $n \in \{6, \dots, 15\}$  are taken from Felsenthal and Machover (1998, sec. 5.3), the data for  $n \in \{25, 27\}$  are official Eurostat figures downloaded on 19.01.2012. The target power distributions  $\beta$  are the respective “fair” ones computed by Penrose’s square root rule (see equation (6)).

In Tables 2–4 we compare the three considered heuristics under different metrics with the (approximated) optimal solution of the inverse power index problem. We distinguish between  $\mathcal{S}$ ,  $\mathcal{C}$ , and  $\mathcal{W}$  as the sets of admissible voting structures. Besides the absolute deviations (measured in the respective metric) we also report an indicator of relative quality: if the distance between  $\beta$  and the PBI  $B(v^*)$  of the optimal solution  $v^* \in \mathcal{S}$  is  $\alpha$ , then this is the *unavoidable* absolute “error” associated with the given instance of the inverse problem. Now if a certain heuristic delivers a distance of  $\delta$  then  $(\delta - \alpha)/\alpha$  can be regarded as the *avoidable error* relative to global optimization in  $\mathcal{S}$ . It is labeled *S-error* in the tables. A value of 1 (or 100%) means that the heuristic’s approximation error is twice the unavoidable one.

<sup>13</sup>Consideration of a similar variation of  $d_\infty$  broadly confirms the comparisons based on  $d_1$ ,  $d_1'$ , and  $d_\infty$ .

The “†”-symbol indicates that the stated value in Tables 2–4 has not been computationally proven to be optimal: for simple games and  $n = 9$ , for instance, we stopped the ILP solution process after memory usage of 31 GB and 18461700 branch-and-bound nodes; for  $n = 10$ , we interrupted after 301 GB and 16735508 nodes. The “††”-symbol indicates that a lower bound for the minimal distance in  $\mathcal{S}$  or  $\mathcal{C}$  was inferred from  $\mathcal{W}$ . The “††”-marked numbers need not be optimal a fortiori. The “ $\infty$ ” entries indicate avoidable errors greater than factor 500. Finally,  $0.000000^\dagger$  or  $0.000000^{\dagger\dagger}$  represent positive numbers  $< 0.5 \cdot 10^{-6}$ .

The computation times for obtaining the numbers in the  $\mathcal{C}$ -column in Table 2, using the hardware and software described above, ranged from less than a second for  $n = 6$  to 5 days for  $n = 12$ . The exact solution for  $n = 10$  in  $\mathcal{W}$  took 2 days; the approximate one for  $n = 27$  was obtained in 3 hours.<sup>14</sup>

$n$	$v^* \in \mathcal{S}$	$v^{**} \in \mathcal{C}$	$v^{***} \in \mathcal{W}$	50%-heuristic		$q^*$ -heuristic		$\bar{q}$ -heuristic	
	$d_1$	$d_1$	$d_1$	$d_1$	$\mathcal{S}$ -error	$d_1$	$\mathcal{S}$ -error	$d_1$	$\mathcal{S}$ -error
6	0.051857	0.051857	0.051857	0.300398	4.79	0.091100	0.76	0.091100	0.76
9	0.005294 <sup>†</sup>	0.008641	0.010359	0.065528	11.38	0.060195	10.37	0.069792	12.18
10	0.002639 <sup>†</sup>	0.004840	0.007219	0.038751	13.68	0.033229	11.59	0.026466	9.03
12	0.001033 <sup>†</sup>	0.001033 <sup>†</sup>	0.005170 <sup>†</sup>	0.028700	26.78	0.019827	18.19	0.019827	18.19
15	0.000476 <sup>††</sup>	0.000476 <sup>††</sup>	0.000476 <sup>†</sup>	0.026742	55.18	0.006820	13.33	0.006361	12.36
25	0.000000 <sup>††</sup>	0.000000 <sup>††</sup>	0.000000 <sup>†</sup>	0.019422	“ $\infty$ ”	0.000744	“ $\infty$ ”	0.003096	“ $\infty$ ”
27	0.000000 <sup>††</sup>	0.000000 <sup>††</sup>	0.000000 <sup>†</sup>	0.018003	“ $\infty$ ”	0.000633	“ $\infty$ ”	0.002457	“ $\infty$ ”

Table 2: Performance for Penrose square root targets in the  $d_1$ -metric (1958–2011 EU data)

The reported numbers give rise to several observations that are independent of the chosen metric:<sup>15</sup> (i) The approximation errors of the heuristics and the optimal solutions in  $\mathcal{W}$  (and a fortiori in  $\mathcal{C}$  and  $\mathcal{S}$ ) tend to zero as  $n$  increases. (ii) Except for  $n = 9$ , the  $q^*$ - and the  $\bar{q}$ -heuristics perform noticeably better than the 50%-heuristic. (iii) The  $q^*$  and  $\bar{q}$ -heuristics produce comparable absolute deviations from the ideal for  $n \leq 15$  but differ by a factor of 2 or more for  $n > 15$ . We conjecture that this has to do with the normal density approximation, which is underlying Słomczyński and Życzkowski’s (2007) derivation of  $q^*$ , becoming noticeably more accurate for the population distribution in the enlarged EU. This allows  $q^*$ ’s performance to improve by almost an order of magnitude between  $n = 15$  and 25, while performance of the  $\bar{q}$ -heuristic (which ignores the specific population distribution at hand and picks the optimal quota averaged over many possible distributions) improves only by about the same factor as the 50%-heuristic and thus in line with the asymptotic proportionality results captured by the Penrose limit theorem.

We also find that (iv) the respective optimal weighted games  $v^{***} \in \mathcal{W}$  yield deviations that are only moderately higher than those of  $v^* \in \mathcal{S}$  in absolute terms. This might be interpreted as indicating that relatively little is lost by restricting attention to weighted voting games in conventional hill-climbing algorithms. It should be noted, however, that we could not prove optimality in  $\mathcal{S}$  for  $n \geq 9$ , and for  $n \geq 15$  we only have upper bounds obtained from  $\mathcal{W}$ . So the observation might not be very robust. And, in relative terms, the errors in  $\mathcal{S}$  or  $\mathcal{C}$  are several times smaller than those in  $\mathcal{W}$  for  $n = 10$  or 12.

Finally, (v) the *relative* errors of the heuristics compared to either  $v^* \in \mathcal{S}$  or  $v^{***} \in \mathcal{W}$  are sizeable even for small  $n \leq 15$ ; and they become huge for  $n > 25$ . For small  $n$  like  $n = 6$  or 9, the unavoidable error, i.e., the distance between  $B(v^*)$  and the target vector  $\beta$ , is still big because comparatively few distinct PBI values exist. At the same time, such numbers  $n$  are far too small for the normal approximation which underlies the  $q^*$  and  $\bar{q}$ -heuristics or for the asymptotics which motivate the 50%-heuristic to have leverage. So the heuristics do not perform well in absolute terms, but they are not that

<sup>14</sup>The additional constraints (1) and (2) which ensure  $v$ ’s weightedness in our ILP formulation considerably slow down the computations (because the so-called *integrality gap* increases). In contrast, the analogous constraints which ensure completeness impose useful structure on the problem compared to unrestrained optimization in  $\mathcal{S}$ . This explains, e.g., why an exact solution can be reported for  $n = 10$  in Table 4 in the  $\mathcal{C}$ -column but not in the  $\mathcal{S}$  and  $\mathcal{W}$ -columns.

<sup>15</sup>Note that the three metrics behave differently when, e.g., distance between  $(1, 0, \dots, 0)$  and  $(\frac{1}{n}, \dots, \frac{1}{n}) \in \Delta(n - 1)$  is considered for increasing  $n$ . Deviations should, therefore, be compared only within and not across tables.

$n$	$v^* \in \mathcal{S}$	$v^{**} \in \mathcal{C}$	$v^{***} \in \mathcal{W}$	50%-heuristic		$q^*$ -heuristic		$\bar{q}$ -heuristic	
	$d'_1$	$d'_1$	$d'_1$	$d'_1$	$\mathcal{S}$ -error	$d'_1$	$\mathcal{S}$ -error	$d'_1$	$\mathcal{S}$ -error
6	0.018967	0.021487	0.021487	0.110284	4.81	0.027465	0.45	0.027465	0.45
9	0.001902 <sup>†</sup>	0.002752	0.003513	0.019015	9.00	0.018935	8.96	0.017643	8.28
10	0.000803 <sup>†</sup>	0.001442	0.001909 <sup>†</sup>	0.008893	10.07	0.007325	8.12	0.005489	5.84
12	0.000309 <sup>†</sup>	0.000447 <sup>†</sup>	0.000810 <sup>†</sup>	0.007840	24.37	0.004005	11.96	0.004005	11.96
15	0.000152 <sup>††</sup>	0.000152 <sup>††</sup>	0.000152 <sup>†</sup>	0.007790	50.26	0.001230	7.09	0.001554	9.23
25	0.000000 <sup>††</sup>	0.000000 <sup>††</sup>	0.000000 <sup>†</sup>	0.004874	“∞”	0.000213	“∞”	0.000751	“∞”
27	0.000000 <sup>††</sup>	0.000000 <sup>††</sup>	0.000000 <sup>†</sup>	0.004411	“∞”	0.000176	“∞”	0.000578	“∞”

Table 3: Performance for Penrose square root targets in the  $d'_1$ -metric (1958–2011 EU data)

$n$	$v^* \in \mathcal{S}$	$v^{**} \in \mathcal{C}$	$v^{***} \in \mathcal{W}$	50%-heuristic		$q^*$ -heuristic		$\bar{q}$ -heuristic	
	$d_\infty$	$d_\infty$	$d_\infty$	$d_\infty$	$\mathcal{S}$ -error	$d_\infty$	$\mathcal{S}$ -error	$d_\infty$	$\mathcal{S}$ -error
6	0.014948	0.014948	0.014948	0.082758	4.54	0.032728	1.19	0.032728	1.19
9	0.001498 <sup>†</sup>	0.001840	0.002240	0.019238	11.84	0.015909	9.62	0.023179	14.47
10	0.000575 <sup>†</sup>	0.001211	0.001960 <sup>†</sup>	0.011574	19.13	0.006316	9.98	0.009721	15.91
12	0.000229 <sup>†</sup>	0.000138 <sup>†</sup>	0.000865 <sup>†</sup>	0.007940	33.67	0.005756	24.13	0.005756	24.13
15	0.000066 <sup>††</sup>	0.000066 <sup>††</sup>	0.000066 <sup>†</sup>	0.005923	88.74	0.001798	26.24	0.001202	17.21
25	0.000000 <sup>††</sup>	0.000000 <sup>††</sup>	0.000000 <sup>†</sup>	0.003834	“∞”	0.000173	“∞”	0.000384	“∞”
27	0.000000 <sup>††</sup>	0.000000 <sup>††</sup>	0.000000 <sup>†</sup>	0.003434	“∞”	0.000156	“∞”	0.000277	“∞”

Table 4: Performance for Penrose square root targets in the  $d_\infty$ -metric (1958–2011 EU data)

bad in relative terms because of high unavoidable errors. Now as  $n$  increases, the heuristics perform significantly better in absolute terms. However, the unavoidable error vanishes even more quickly as the number of distinct simple voting games and, hence, of feasible PBI vectors increases very fast in  $n$  (see Table 1).

Observation (v) is probably the most interesting: whenever one seeks an optimal solution of the inverse power index problem, all three heuristics are unsatisfactory from a pure operations research perspective. The heuristic solutions can be improved by very large factors, and this becomes more rather than less pronounced as  $n$  grows. Of course, from an applied point of view the absolute approximation errors get so small for large  $n$  that they may be regarded as negligible. They might still be relevant, however. To get a sense for what a deviation at the 5th decimal place means consider, e.g., the ideal Penrose square root power distribution  $\beta^{27}$  for the EU Council from 2011 and compute the analogous vector  $\beta^{27'}$  which would result if 50000 people moved from Germany to France or were mis-counted in the statistics. Then  $\|\beta^{27} - \beta^{27'}\|_1 \approx 0.0000634$ .

## 5.2 Finite grid of objective vectors

Every vector in  $\mathbb{R}_{\geq 0}^n$  whose entries sum to 1, i.e., each element of  $\Delta(n-1)$ , can in principle be a desired power distribution in a specific context. We approximate this infinite space by a finite set  $G^n$ . We impose  $\beta_1 \geq \beta_2 \geq \dots \geq \beta_n$  and let the desired power of the first  $n-1$  voters be an integral multiple of  $s = 0.01$ ; the desired power of the  $n$ -th voter follows from the sum condition. We refer to  $G^n$  as our *grid* of target vectors and to each  $\beta \in G^n$  as a *grid point*.<sup>16</sup> Table 5 reports key statistics

<sup>16</sup>Step size  $s$  has to be chosen with care: the number of grid points can be intractably great already for small  $n$  if  $s$  is too small. But a larger  $s$  induces a coarser grid of feasible target vectors. This becomes more and more problematic as  $n$  increases because of the corresponding natural decrease of an individual voter’s relative power (on average equal to  $1/n$ ). Choosing  $s = 0.25$ , for instance, would result in the four grid points  $(0.5, 0.25, 0.25)$ ,  $(0.5, 0.5, 0)$ ,  $(0.75, 0.25, 0)$ , and  $(1, 0, 0)$  for  $n = 3$ . And  $G^n$  would contain merely five grid points for *any*  $n \geq 4$ :  $(0.25, 0.25, 0.25, 0.25, 0, \dots, 0)$ ,  $(0.5, 0.25, 0.25, 0, \dots, 0)$ ,  $(0.5, 0.5, 0, \dots, 0)$ ,  $(0.75, 0.25, 0, \dots, 0)$ , and  $(1, 0, \dots, 0)$ .

for the distribution of unavoidable deviations from the ideal vectors in the  $d_1$  and  $d_\infty$ -metrics: its median, average, 10%, 5%, and 1%-percentile. The deviation figures are based on the enumerated exact solutions in  $\mathcal{W}$  for  $n \leq 7$  and approximations thereof for larger  $n$ . For instance, the number 0.01077 for  $n = 5$  in the right-most column indicates that for 1% of the considered 46262 different target vectors  $\beta$  one can obtain a distance  $d_\infty(\beta, B(v^{***})) \leq 0.011$ , and the remaining 99% target vectors can only be approximated less well within  $\mathcal{W}$ . A number of grid points in parentheses indicates the size of the considered random sample whenever only a subset of all grid points could be dealt with. The deviation statistics in the corresponding rows (in light color) involve a sample error in addition to the small error of using a conventional local hill-climbing algorithm instead of global optimization in  $\mathcal{W}$ . For example, the number 0.0011 for  $n = 15$  in the third column indicates that *half* of the 10000 target vectors  $\beta$  which were sampled at random (with replacement) could, by some weighted voting game, be achieved with a  $d_1$ -distance of 0.0011 or less; the remaining draws resulted in target vectors for which our search algorithm terminated with a best achievable  $B(v^{***})$  further away.

$n$	#grid points	$d_1$ -metric					$d_\infty$ -metric				
		med.	av.	10%	5%	1%	med.	av.	10%	5%	1%
2	51	0.2400	0.2451	0.0400	0.0200	0.0000	0.12000	0.12255	0.02000	0.01000	0.00000
3	884	0.2400	0.2278	0.1000	0.0667	0.0200	0.12000	0.11391	0.05000	0.03333	0.01000
4	8037	0.1600	0.1622	0.0800	0.0600	0.0400	0.07000	0.07131	0.03667	0.03000	0.01500
5	46262	0.1010	0.1135	0.0600	0.0509	0.0324	0.04000	0.04292	0.02273	0.02000	0.01077
6	189509	0.0667	0.0790	0.0400	0.0356	0.0200	0.02222	0.02630	0.01333	0.01069	0.00815
7	596763	0.0422	0.0543	0.0257	0.0213	0.0165	0.01255	0.01629	0.00762	0.00667	0.00495
8	(10000)	0.0226	0.0248	0.0154	0.0137	0.0108	0.00601	0.00661	0.00404	0.00358	0.00281
9	(10000)	0.0148	0.0161	0.0100	0.0089	0.0070	0.00357	0.00393	0.00241	0.00216	0.00169
10	(10000)	0.0097	0.0107	0.0065	0.0059	0.0046	0.00216	0.00239	0.00145	0.00129	0.00103
11	(10000)	0.0064	0.0070	0.0043	0.0038	0.0031	0.00131	0.00146	0.00088	0.00079	0.00064
12	(10000)	0.0041	0.0045	0.0028	0.0024	0.0019	0.00079	0.00088	0.00052	0.00047	0.00037
13	(10000)	0.0026	0.0029	0.0017	0.0016	0.0013	0.00047	0.00053	0.00032	0.00028	0.00023
14	(10000)	0.0016	0.0018	0.0011	0.0010	0.0008	0.00028	0.00032	0.00019	0.00017	0.00014
15	(10000)	0.0011	0.0012	0.0008	0.0007	0.0006	0.00017	0.00019	0.00012	0.00011	0.00009
16	(10000)	0.0007	0.0008	0.0005	0.0005	0.0004	0.00011	0.00012	0.00009	0.00008	0.00007
17	(10000)	0.0006	0.0006	0.0004	0.0004	0.0003	0.00009	0.00009	0.00007	0.00007	0.00006
18	(10000)	0.0005	0.0005	0.0004	0.0004	0.0003	0.00008	0.00008	0.00006	0.00006	0.00005
19	(10000)	0.0005	0.0005	0.0004	0.0004	0.0003	0.00008	0.00008	0.00006	0.00006	0.00005
20	(10000)	0.0005	0.0005	0.0004	0.0004	0.0003	0.00007	0.00007	0.00006	0.00005	0.00004

Table 5: Distribution of unavoidable absolute deviations  $d_1(\beta, B(v^{***}))$  and  $d_\infty(\beta, B(v^{***}))$

Tables 6–8 report analogous statistics for the distribution of absolute distances for the three heuristics (considering each grid point for up to  $n = 20$ ).<sup>17</sup> A comparison of the respective deviation statistics with those in Table 5 broadly confirm the observations that were made for the very specific target vectors derived from Penrose’s square root rule in Section 5.1: the average and each reported percentile of the avoidable deviations decrease in  $n$ . They can be regarded as small in absolute terms, but they are sizeable in relative terms. Again the 50%-heuristic is clearly outperformed (in the sense of first order stochastic dominance) by the  $q^*$  and  $\bar{q}$ -heuristics for  $n \geq 3$ .

<sup>17</sup>The computation times behind the unavoidable errors in Table 5 ranged from less than 1s or 40m for the  $n = 5$  and  $n = 7$  rows, respectively, to 14h for the  $n = 8$  row and 31h for  $n = 20$ . Quite some time is spent on approximating the exact solution of the inverse problem. Times for the heuristic in, e.g., Table 7 were only 1s, 15s, 50m, and 3h, respectively. A sample of 10000 grid points represents a reasonable compromise between precision and computational effort. Raising the sample size to 100000 would, e.g, have produced the median, average and quantile entries (0.0041, 0.0045, 0.0027, 0.0025, 0.0020) for the  $d_1$ -metric and  $n = 12$ ; lowering it to 1000 would have resulted in (0.0042, 0.0047, 0.0028, 0.0026, 0.0018).

$n$	#grid points	$d_1$ -metric					$d_\infty$ -metric				
		med.	av.	10%	5%	1%	med.	av.	10%	5%	1%
2	51	0.480	0.480	0.080	0.020	0.000	0.240	0.240	0.040	0.010	0.000
3	884	0.560	0.555	0.200	0.133	0.047	0.280	0.278	0.100	0.067	0.023
4	8037	0.440	0.509	0.200	0.153	0.080	0.210	0.249	0.083	0.063	0.033
5	46262	0.347	0.448	0.160	0.127	0.075	0.153	0.209	0.061	0.049	0.029
6	189509	0.297	0.389	0.129	0.103	0.066	0.120	0.177	0.045	0.035	0.023
7	596763	0.247	0.338	0.101	0.080	0.052	0.097	0.151	0.033	0.025	0.016
8	1527675	0.206	0.297	0.080	0.063	0.041	0.080	0.132	0.025	0.019	0.012
9	3314203	0.176	0.265	0.064	0.051	0.034	0.068	0.118	0.020	0.015	0.009
10	6292069	0.153	0.240	0.053	0.043	0.029	0.059	0.107	0.016	0.012	0.007
11	10718685	0.136	0.220	0.046	0.037	0.025	0.052	0.099	0.014	0.010	0.006
12	16713148	0.123	0.205	0.041	0.033	0.023	0.047	0.092	0.012	0.009	0.005
13	24234058	0.112	0.193	0.038	0.030	0.021	0.044	0.087	0.011	0.008	0.005
14	33097743	0.104	0.183	0.035	0.028	0.020	0.041	0.083	0.010	0.008	0.004
15	43018955	0.097	0.175	0.033	0.027	0.019	0.038	0.079	0.010	0.007	0.004
16	53662038	0.092	0.169	0.032	0.026	0.018	0.037	0.076	0.009	0.007	0.004
17	64684584	0.087	0.164	0.031	0.025	0.017	0.035	0.074	0.009	0.006	0.004
18	75772412	0.084	0.159	0.030	0.024	0.017	0.034	0.072	0.009	0.006	0.003
19	86658411	0.081	0.156	0.029	0.024	0.016	0.033	0.071	0.008	0.006	0.003
20	97132873	0.078	0.153	0.028	0.023	0.016	0.032	0.070	0.008	0.006	0.003

Table 6: Distribution of absolute deviations for the 50%-heuristic

### 5.3 Analytical example

A statement which would be analogous to observation (v) in Section 5.1 cannot be deduced on the basis of statistical information as provided by Tables 6–8. We, therefore, close our computational investigation by studying a particularly simple analytical example. It shows transparently that – as indicated by observation (v) – *relative* deviations between the considered heuristics and globally optimal solutions need not disappear for  $n \rightarrow \infty$ .

Consider the desired power distribution

$$\beta^n = \frac{1}{2n-1} \underbrace{(2, \dots, 2)}_{n-1 \text{ twos}}, 1$$

for  $n \geq 2$  and choose  $w^n = \beta^n$  as all three heuristics do.<sup>18</sup> For any quota  $q \in I_1^j = \frac{1}{2n-1} \cdot (2j-1, 2j]$ , where  $1 \leq j \leq n-1$  and  $j \in \mathbb{N}$ , the PBI of the smallest constituency is exactly zero and, by symmetry, the (normalized) PBI of each of the other constituencies equals  $\frac{1}{n-1}$ . For the remaining possibilities  $q \in I_2^j = \frac{1}{2n-1} \cdot (2j, 2j+1]$  where  $0 \leq j \leq n-1$ , all constituencies have a PBI of  $\frac{1}{n}$ . Denoting the corresponding weighted games by  $v_{1,j}^n$  and  $v_{2,j}^n$  one obtains

$$\begin{aligned} d_1(v_{1,j}^n, \beta^n) &= \frac{2}{2n-1}, \\ d_1(v_{2,j}^n, \beta^n) &= \frac{2}{2n-1} \cdot \frac{n-1}{n}, \\ d_\infty(v_{1,j}^n, \beta^n) &= \frac{1}{2n-1}, \text{ and} \\ d_\infty(v_{2,j}^n, \beta^n) &= \frac{1}{2n-1} \cdot \frac{n-1}{n}. \end{aligned}$$

<sup>18</sup>The construction is inspired by a sequence of weighted voting games to which the Penrose limit theorem does *not* apply even though every voter's relative weight vanishes as  $n \rightarrow \infty$ . Namely, the sequence  $\{(\frac{1}{2}; w^n)\}_{n \in \mathbb{N}}$  belongs to the class of somewhat pathological examples alluded to on p. 2 (cf. Lindner and Owen 2007).

n	#grid points	$d_1$ -metric					$d_\infty$ -metric				
		med.	av.	10%	5%	1%	med.	av.	10%	5%	1%
2	51	0.480	0.480	0.080	0.020	0.000	0.240	0.240	0.040	0.010	0.000
3	884	0.400	0.434	0.160	0.107	0.040	0.200	0.217	0.080	0.053	0.020
4	8037	0.340	0.370	0.160	0.120	0.060	0.147	0.172	0.065	0.050	0.025
5	46262	0.280	0.312	0.133	0.107	0.062	0.113	0.138	0.052	0.040	0.023
6	189509	0.227	0.263	0.109	0.088	0.058	0.088	0.112	0.038	0.030	0.020
7	596763	0.189	0.224	0.087	0.070	0.047	0.071	0.093	0.027	0.021	0.014
8	1527675	0.158	0.192	0.066	0.053	0.035	0.056	0.079	0.020	0.015	0.010
9	3314203	0.133	0.168	0.051	0.040	0.026	0.047	0.068	0.014	0.011	0.007
10	6292069	0.114	0.148	0.039	0.030	0.019	0.039	0.060	0.011	0.008	0.005
11	10718685	0.098	0.132	0.030	0.023	0.014	0.033	0.054	0.008	0.006	0.003
12	16713148	0.086	0.120	0.024	0.017	0.010	0.029	0.049	0.006	0.004	0.002
13	24234058	0.075	0.110	0.019	0.013	0.007	0.026	0.045	0.005	0.003	0.002
14	33097743	0.068	0.102	0.016	0.011	0.005	0.023	0.042	0.004	0.003	0.001
15	43018955	0.061	0.096	0.013	0.008	0.004	0.021	0.040	0.003	0.002	0.001
16	53662038	0.056	0.091	0.011	0.007	0.003	0.019	0.038	0.003	0.002	0.001
17	64684584	0.052	0.087	0.009	0.006	0.003	0.018	0.036	0.002	0.001	0.001
18	75772412	0.049	0.083	0.008	0.005	0.002	0.017	0.035	0.002	0.001	0.000
19	86658411	0.046	0.081	0.007	0.004	0.002	0.016	0.034	0.002	0.001	0.000
20	97132873	0.044	0.078	0.006	0.004	0.002	0.015	0.033	0.002	0.001	0.000

Table 7: Distribution of absolute deviations for the  $q^*$ -heuristic

So, independently of the quota, the  $\|\cdot\|_1$ -error is  $\frac{2}{2n-1} + O(n^{-2})$  and the  $\|\cdot\|_\infty$ -error is  $\frac{1}{2n-1} + O(n^{-2})$ . The  $\bar{q}$  and  $q^*$ -heuristics prescribe quotas of

$$\bar{q} = \frac{1}{2} + \frac{1}{\sqrt{\pi n}}, \text{ and}$$

$$q^* = \frac{1}{2} + \frac{\sqrt{4n-3}}{4n-2},$$

respectively. They and  $q^\circ = 50\%$  fall into  $I_1^j$  and  $I_2^j$  for infinitely many  $n$ . Thus, all three rules render the smallest constituency a null player infinitely many times as  $n \rightarrow \infty$ , just as it happened to Luxembourg in the EEC Council between 1958 and 1973, and yield the indicated distances.

In contrast, there always exists a simple game  $v^* \in \mathcal{S}$  whose PBI attains  $\beta^n$  *exactly* for  $6 \leq n \leq 18$ .<sup>19</sup> And we conjecture that this remains true for  $n \geq 19$ . So the corresponding distance between  $\beta$  and  $B(v^*)$  equals 0 independently of the considered metric. Approximation results for complete simple games and distances for the heuristic choice of  $w = \beta^n$  with an “optimal” quota  $q$  that leads to  $v_{2,j}^n$  (abbreviated as  $q$ -heuristic) are summarized in Tables 9 and 10.<sup>20</sup> Since the unavoidable error in the class of simple games  $\mathcal{S}$  (and hence of finite intersections of weighted games  $v \in \mathcal{W}$ ) is zero for  $6 \leq n \leq 18$  and presumably beyond, we consider the  $\mathcal{C}$ -error in order to evaluate the relative performance of the  $q$ -heuristic.

Tables 9 and 10 suggest that (i) the  $\mathcal{C}$ -error converges to a positive constant in case of the  $d_1$ -metric and (ii) this error even seems to grow without bound when the  $d_\infty$ -metric is used. The key finding that *relative errors fail to disappear as  $n$  grows large* can be made more rigorous. To this end, consider

<sup>19</sup>An example for  $n = 6$  is given by the following set of minimal winning coalitions:  $\{2, 4, 5, 6\}$ ,  $\{2, 3, 4, 5\}$ ,  $\{1, 3, 5, 6\}$ ,  $\{1, 3, 4, 5\}$ ,  $\{1, 2, 4, 6\}$ , and  $\{1, 2, 3, 5\}$ , which attains the PBI vector  $\frac{1}{44}(8, 8, 8, 8, 8, 4) = \frac{1}{11}(2, 2, 2, 2, 2, 1)$ .

<sup>20</sup>We remark that we have imposed suitable equivalence classes of voters in the optimizations for  $n \geq 15$  in order to reduce the computational burden. Voters  $i$  and  $j$  are in the same equivalence class if  $v(U) = v(U \cup \{j\} \setminus \{i\})$  for all coalitions  $U$  with  $i \in U$  and  $j \notin U$ . This is more restrictive than requiring identical PBI values of  $i$  and  $j$  only.



$n$	#grid points	$d_1$ -metric					$d_\infty$ -metric				
		med.	av.	10%	5%	1%	med.	av.	10%	5%	1%
2	51	0.280	0.327	0.040	0.020	0.000	0.140	0.164	0.020	0.010	0.000
3	884	0.320	0.332	0.140	0.100	0.040	0.160	0.166	0.070	0.050	0.020
4	8037	0.300	0.304	0.147	0.110	0.050	0.130	0.138	0.063	0.045	0.020
5	46262	0.250	0.263	0.132	0.100	0.060	0.101	0.111	0.050	0.040	0.023
6	189509	0.204	0.224	0.104	0.085	0.056	0.077	0.088	0.036	0.029	0.019
7	596763	0.159	0.187	0.079	0.064	0.043	0.057	0.069	0.024	0.020	0.013
8	1527675	0.120	0.155	0.058	0.047	0.033	0.041	0.055	0.017	0.013	0.009
9	3314203	0.098	0.133	0.044	0.036	0.024	0.033	0.046	0.012	0.010	0.006
10	6292069	0.079	0.115	0.033	0.026	0.018	0.026	0.039	0.009	0.007	0.004
11	10718685	0.071	0.103	0.027	0.021	0.014	0.023	0.035	0.007	0.005	0.003
12	16713148	0.058	0.092	0.021	0.016	0.010	0.018	0.030	0.005	0.004	0.002
13	24234058	0.050	0.084	0.016	0.012	0.007	0.015	0.027	0.004	0.003	0.002
14	33097743	0.045	0.078	0.014	0.010	0.006	0.014	0.025	0.003	0.002	0.001
15	43018955	0.042	0.074	0.012	0.009	0.005	0.012	0.024	0.003	0.002	0.001
16	53662038	0.039	0.070	0.011	0.007	0.004	0.011	0.022	0.002	0.002	0.001
17	64684584	0.040	0.070	0.011	0.008	0.004	0.012	0.023	0.002	0.002	0.001
18	75772412	0.037	0.067	0.009	0.007	0.003	0.011	0.022	0.002	0.001	0.001
19	86658411	0.041	0.069	0.011	0.008	0.004	0.013	0.024	0.002	0.002	0.001
20	97132873	0.038	0.067	0.010	0.007	0.003	0.012	0.023	0.002	0.001	0.001

Table 8: Distribution of absolute deviations for the  $\bar{q}$ -heuristic

the sequence of weighted voting games  $\{(q^n; w^n)\}_{n \in \mathbb{N}}$  with

$$(q^n; w^n) = (2n - a - 4; \underbrace{3, \dots, 3}_a, \underbrace{2, \dots, 2}_{n-a-1}, 1)$$

for a suitable parameter  $a$ . If  $a$  is chosen equal to about  $\frac{6n}{7}$  (the exact number is provided in Lemma 2) then the  $d_1$ -distance between  $\beta^n$  and  $B(q^n; w^n)$  asymptotically tends to  $\frac{1}{2n}$ .<sup>21</sup> The distance achieved by these specific weighted voting games provides an upper bound for the distance achieved by the respective optimal weighted game  $v^{***} \in \mathcal{W}$ . The  $\mathcal{W}$ -error of the heuristic under the  $d_1$ -metric is, therefore, asymptotically bounded from below by

$$\lim_{n \rightarrow \infty} \frac{\left| \frac{2}{2n-1} \cdot \frac{n-1}{n} - \frac{1}{2n} \right|}{\frac{1}{2n}} = 1.$$

So the  $\mathcal{W}$ -error and a fortiori also the  $\mathcal{C}$ -error stay at around 100% even as  $n \rightarrow \infty$  when discrepancies between  $\beta^n$  and  $B(v^{**})$  (or  $B(v^{***})$ ) are evaluated by the  $d_1$ -metric.

Analogously, one can choose  $a$  equal to about  $\frac{2n}{3}$  and then check that the  $d_\infty$ -distance between the game's PBI and  $\beta^n$  asymptotically tends to  $\frac{1}{n^2}$  as  $n \rightarrow \infty$ . This again translates into an asymptotic lower bound for the  $\mathcal{W}$ -error, and a fortiori the  $\mathcal{C}$ -error; these relative errors go to infinity at an approximately linear speed.

## 6 Conclusion

The computations which we have reported in Section 5.1 confirm that if one wants to implement the Penrose square root rule for population data from today's European Union, the  $q^*$ -heuristic of

<sup>21</sup>Details on the simple but tedious computations are provided in Appendix B.

$n$	$v^* \in \mathcal{S}$	$v^{**} \in \mathcal{C}$	$v^{***} \in \mathcal{W}$	$q$ -heuristic	
	$d_1$	$d_1$	$d_1$	$d_1$	$\mathcal{C}$ -error
2	0.333333	0.333333	0.333333	0.333333	0.000000
3	0.266667	0.266667	0.266667	0.266667	0.000000
4	0.214286	0.214286	0.214286	0.214286	0.000000
5	0.038647	0.158730	0.158730	0.177778	0.120000
6	0.000000	0.113636	0.113636	0.151515	0.333333
7	0.000000	0.085470	0.085470	0.131868	0.542857
8	0.000000	0.066667	0.066667	0.116667	0.750000
9	0.000000	0.064171	0.064171	0.104575	0.629630
10	0.000000	0.061042	0.061042	0.094737	0.552000
11	0.000000	0.052158	0.052158	0.086580	0.659944
12	0.000000	0.047254	0.047254	0.079710	0.686856
13	0.000000	0.042353	0.042353	0.073846	0.743590
14	0.000000	0.037037	0.037037	0.068783	0.857143
15	0.000000	0.034483 <sup>††</sup>	0.034483 <sup>†</sup>	0.064368	0.866667
16	0.000000	0.033845 <sup>††</sup>	0.033845 <sup>†</sup>	0.060484	0.780576
17	0.000000	0.032221 <sup>††</sup>	0.032221 <sup>†</sup>	0.057041	0.770270
18	0.000000	0.030866 <sup>††</sup>	0.030866 <sup>†</sup>	0.053968	0.748490
19		0.028108 <sup>††</sup>	0.028108 <sup>†</sup>	0.051209	0.821862
20		0.025641 <sup>††</sup>	0.025641 <sup>†</sup>	0.048718	0.900000

Table 9: Deviations from  $\beta^n$  in the  $d_1$ -metric (analytical example)

Ślomiczyński and Życzkowski and, to a lesser extent, the even simpler  $\bar{q}$ -heuristic perform very well in absolute terms. That is, the distance between a (normalized) square root target distribution  $\beta$  and the PBI  $B(q^*, \beta)$  is close to zero. However, the considered heuristics can still be very far from the globally optimal solution to the inverse problem in relative terms. This finding applies even when only weighted voting games are allowed as feasible solutions. And it is not restricted to small voting bodies, but holds for the current number of EU members  $n = 27$ .

The extensive computations reported in Section 5.2 confirm this observation. They provide the first systematic evaluation of the unavoidable deviations between arbitrary target PBI power vectors and those that are actually implementable for voting bodies with up to  $n = 20$  members. Numbers such as the ones reported in Table 5 can potentially be useful in order to improve termination criteria for local search algorithms (e.g., Leech 2002a, 2003), which have been used in applied studies. If, say, a locally optimal candidate solution for an inverse problem with  $n = 11$  voters has a  $d_1$ -deviation from the desired vector  $\beta$  greater than 0.0064, then Table 5 indicates that the odds of further improvements in the class of weighted voting games are 50:50 and search presumably should continue in a different part of the game space. If, however, the deviation is smaller than 0.0031, then the odds are rather 1:99; termination might then make sense.

That desired PBI distributions which concentrate a major share of relative power amongst a few voters pose problems for the considered heuristics is not surprising. After all, the derivation of  $q^*$  by Ślomiczyński and Życzkowski (2007) involves a technical condition (see fn. 10) from which one can conclude that the target PBI of a single voter should approach zero at least as fast as  $1/\sqrt{n}$ . It is much less obvious, however, that, first, it is not sufficient to have a target vector  $\beta$  without “outliers” in order to obtain a heuristic solution that is good relative to the exact one and, second, the relative errors may get larger rather than smaller as  $n$  increases. This emerged from the extensive numerical computations reported in Sections 5.1–5.2 and has been formally demonstrated for a specific analytical example in Section 5.3. One might, therefore, summarize our findings as justifying and potentially even calling for case-specific optimization rather than the application of a generally rather good heuristic – not only for small but even for large voting bodies.

$n$	$v^* \in \mathcal{S}$	$v^{**} \in \mathcal{C}$	$v^{***} \in \mathcal{W}$	$q$ -heuristic	
	$d_\infty$	$d_\infty$	$d_\infty$	$d_\infty$	$\mathcal{C}$ -error
2	0.166667	0.166667	0.166667	0.166667	0.000000
3	0.133333	0.133333	0.133333	0.133333	0.000000
4	0.107143	0.107143	0.107143	0.107143	0.000000
5	0.019324	0.050505	0.050505	0.088889	0.760000
6	0.000000	0.034759	0.034759	0.075758	1.179487
7	0.000000	0.022624	0.022624	0.065934	1.914286
8	0.000000	0.015686	0.015686	0.058333	2.718750
9	0.000000	0.014199	0.014199	0.052288	2.682540
10	0.000000	0.008772	0.008772	0.047368	4.400000
11	0.000000	0.008282	0.008282	0.043290	4.227273
12	0.000000	0.007688	0.007688	0.039855	4.183908
13	0.000000	0.005373	0.005373	0.036923	5.871795
14	0.000000	0.005109	0.005109	0.034392	5.732143
15	0.000000	0.004628 <sup>†</sup>	0.004815 <sup>†</sup>	0.032184	5.954839
16	0.000000	0.003619 <sup>††</sup>	0.003619 <sup>†</sup>	0.030242	7.357143
17	0.000000	0.003463 <sup>††</sup>	0.003463 <sup>†</sup>	0.028520	7.235294
18	0.000000	0.003297 <sup>††</sup>	0.003297 <sup>†</sup>	0.026984	7.185185
19		0.002600 <sup>††</sup>	0.002600 <sup>†</sup>	0.025605	8.848225
20		0.002502 <sup>††</sup>	0.002502 <sup>†</sup>	0.024359	8.737500

Table 10: Deviations from  $\beta^n$  in the  $d_\infty$ -metric (analytical example)

## Appendix A: ILP formulation for the inverse Penrose-Banzhaf index problem

Even though stating the inverse power index problem as an optimization problem is trivial (see (3)), coming up with an implementable formulation for finding an exact solution is not. Voting systems are discrete objects, and so some kind of discrete optimization is needed. Exhaustive enumeration (see Keijzer et al. 2010) is limited at best to  $n \leq 9$  (see Table 1). A much more tractable alternative is to describe the set of feasible binary voting systems by integer variables and to use some of the available optimization software packages. These allow significantly larger numbers of variables when dealing with *linear* rather than *non-linear* (mixed) integer optimization problems. So it is unfortunate that problem (3) cannot directly be translated into a linear problem for the Penrose-Banzhaf index.<sup>22</sup> The “work-around”, which has first been suggested by Kurz (2012b) and is adopted here, is to use ILP techniques in order to merely find out whether some binary voting system  $v$  exists whose PBI vector  $B(v)$  is at most a specified distance  $\alpha$  apart from the target  $\beta$ . This *feasibility problem* can be solved much more easily than the underlying minimization problem. Still, one can iteratively determine the exact solution of (3) by varying  $\alpha$ .

We will mostly confine our description to the case of measuring distance by the  $d_1$ -metric. Adaptations to the  $d'_1$  or  $d_\infty$ -metric are straightforward. They involve heterogeneous coefficients in inequality (19) below for  $d'_1$ , and neither  $i$ -subscripts in (16)–(18) nor a summation in (19) for  $d_\infty$ .

The PBI vector  $(1, 0, \dots, 0)$  of a dictator has at most a  $d_1$ -distance of 2 from any normalized power distribution (summing to 1). This is in fact the worst case, and the minimal achievable deviation  $\alpha^*$  must lie inside the interval  $[l_1, u_1]$ , where  $l_1 = 0$  and  $u_1 = 2$ . In each iteration  $t = 1, \dots, T$  of the algorithm we will check whether  $\alpha = (u_t - l_t)/2$  is a feasible distance between target  $\beta$  and the PBI values generated by the considered class of voting systems. If so, we set  $u_{t+1} = \alpha$  and leave  $l_{t+1} = l_t$

<sup>22</sup>Interestingly, one can easily linearize the analogous inverse problem for the Shapley-Shubik power index (SSI). So even though the PBI is easier to compute than the SSI, the corresponding inverse problem is more difficult.

unchanged; otherwise we update  $l_{t+1} = \alpha$  and leave  $u_{t+1} = u_t$  unchanged. In each iteration the length of the interval  $[l_t, u_t]$  shrinks by a factor of 2. Since the total number of swings in an  $n$ -player voting game lies between  $n$  and  $m \binom{n}{m} < n2^n$  where  $m = \lfloor \frac{n}{2} \rfloor + 1$  (see, e.g., Felsenthal and Machover 1998, sec. 3.3), two distinct PBI vectors differ, both in the  $d_1$  and the  $d_\infty$ -metric, by at least  $(\frac{1}{n2^n})^2$ . A finite number  $T$  of iterations are, therefore, sufficient for obtaining a solution. More specifically,  $O(n)$  bisections on  $\alpha$  are needed before  $u_t - l_t \leq (\frac{1}{n2^n})^2$  and further improvements become theoretically impossible.

A pseudo-code description of this bisection approach reads as follows:<sup>23</sup>

**Input:** desired power index vector  $\beta$ , class of binary voting systems  $\Gamma$ , metric  $d(\cdot)$

**Output:** minimum  $d$ -distance  $\alpha^*$  between  $\beta$  and PBI vectors induced by  $\Gamma$

$l_1 = 0$

$u_1 = 2$

$\alpha^* = 2$

$\varepsilon = (\frac{1}{n2^n})^2$

$t = 1$

while  $u_t - l_t > \varepsilon$

$\alpha = \frac{u_t + l_t}{2}$

$t = t + 1$

solve feasibility problem  $\langle \beta, \Gamma, d(\cdot), \alpha \rangle$

if  $v \in \Gamma$  such that  $d(B(v) - \beta) \leq \alpha$  exists

then

$u_t = d(B(v) - \beta), \alpha^* = u_t$

else

$l_t = \alpha$

end if

end while

return  $\alpha^*$

The *feasibility problem*  $\langle \beta, \Gamma, d(\cdot), \alpha \rangle$  consists of verifying whether there exists a voting system  $v \in \Gamma$  such that  $d(B(v), \beta) \leq \alpha$ . The following ILP formulation describes it for  $\Gamma = \mathcal{S}$  and the  $d_1$ -metric. Adaptations to  $\mathcal{C}$  or  $\mathcal{W}$  and  $d'_1(\cdot)$  or  $d_\infty(\cdot)$  involve further variables and (modified) constraints, but are otherwise very similar:

---

<sup>23</sup>The description focuses on finding the minimal distance  $\alpha^*$ . A game  $v^* \in \Gamma$  with  $B(v^*) = \alpha^*$  can straightforwardly be obtained from the solution to the feasibility problem  $\langle \beta, \Gamma, d(\cdot), \alpha^* \rangle$ .

$$x_S \in \{0, 1\} \quad \forall S \subseteq N, \quad (7)$$

$$x_S \leq x_T \quad \forall S \subseteq T \subseteq N, \quad (8)$$

$$x_\emptyset = 0 \quad (9)$$

$$x_N = 1 \quad (10)$$

$$y_{i,S} \in \{0, 1\} \quad \forall 1 \leq i \leq n, S \subseteq N \setminus \{i\}, \quad (11)$$

$$y_{i,S} = x_{S \cup \{i\}} - x_S \quad \forall 1 \leq i \leq n, S \subseteq N \setminus \{i\}, \quad (12)$$

$$s_i \geq 0 \quad \forall 1 \leq i \leq n, \quad (13)$$

$$s_i = \sum_{S \subseteq N \setminus \{i\}} y_{i,S} \quad \forall 1 \leq i \leq n, \quad (14)$$

$$s = \sum_{i=1}^n s_i, \quad (15)$$

$$\delta_i \geq 0 \quad \forall 1 \leq i \leq n, \quad (16)$$

$$\delta_i \geq s_i - \beta_i \cdot s \quad \forall 1 \leq i \leq n, \quad (17)$$

$$\delta_i \geq -s_i + \beta_i \cdot s \quad \forall 1 \leq i \leq n, \quad (18)$$

$$\sum_{i=1}^n \delta_i \leq \alpha \cdot s. \quad (19)$$

The binary variables  $x_S$  define a Boolean function  $v$  via  $v(S) = x_S$ ; inequalities (7)–(10) ensure that they represent a simple game. The binary auxiliary variables  $y_{i,S} = x_{S \cup \{i\}} - x_S$  which are introduced in (11)–(12) for all  $i \in N$  and  $\emptyset \subseteq S \subseteq N \setminus \{i\}$  satisfy  $y_{i,S} = 1$  if and only if coalition  $S$  is a swing for voter  $i$ , i.e., contributes  $1/2^{n-1}$  to  $B'_i(v)$ . They are used in order to determine the number of swings  $s_i = 2^{n-1} \cdot B'_i(v)$  for each player  $i$  in equation (14). The total number of swings  $s = \sum_{i=1}^n s_i$  is defined in equation (15). Based on this total number, the individual deviation  $\delta_i = |s_i - \beta_i \cdot s|$  from the target number of swings is captured by inequalities (17) and (18). The feasibility of a  $d_1$ -distance  $\alpha$  is then finally checked by introducing constraint (19). Namely, a simple game  $v \in \mathcal{S}$  whose PBI has  $d_1$ -distance of  $\alpha$  or less exists if and only if the feasible set defined by (7)–(19) is non-empty.

The answer to whether this is the case – and, as a by-product, some  $v \in \mathcal{S}$  with distance at most  $\alpha$  – can be obtained by feeding (7)–(19) into a standard ILP software package in the required format. We have used IBM ILOG CPLEX 12.4 and the hardware described in Section 5.

## Appendix B: Analytical PBI calculations

This appendix presents some technical details on the PBI computations for the sequence of weighted voting games  $\{v^n\}_{n \in \mathbb{N}}$  with

$$v^n = (q^n; w^n) = (2n - a - 4; \underbrace{3, \dots, 3}_a \text{ threes}, \underbrace{2, \dots, 2}_{n-a-1} \text{ twos}, 1),$$

which is considered in the final paragraphs of Section 5.3. Our first lemma determines the number of swings in  $v^n$  for each voter  $i = 1, \dots, n$ , that is, the cardinality of set  $\{S \subseteq N \setminus \{i\} : v^n(S \cup \{i\}) - v^n(S) = 1\}$ , as a function of  $a$  and  $n$ .

**Lemma 1** *The numbers of swings in  $v^n$  are  $2n - a - 2$ ,  $2n - a - 4$ , and  $a$  for all voters with weight 3, 2, and 1, respectively.*

**Proof** It is convenient to exploit the fact that for any  $v \in \mathcal{S}$  the number of voter  $i$ 's swings in  $v$  and in the *dual game*  $v' \in \mathcal{S}$  which is obtained by setting  $v'(S) = 1 - v(S)$  for all  $S \subseteq N$  must coincide. So instead of  $v^n$  consider the game  $v^{n'}$  which involves identical weights but quota  $q' = 4$  instead of  $2n - a - 4$ . Referring to winning and losing coalitions in  $v^{n'}$  we have:

- (i) A voter  $i$  with  $w_i = 3$  renders a losing coalition  $S \subseteq N \setminus \{i\}$  winning by joining if either  $|S| = 1$  or  $S = \{j, k\}$  with  $w_j = 2$  and  $w_k = 1$ . There are  $n - 1$  coalitions of the former and  $n - a - 1$  coalitions of the latter type, amounting to  $2n - a - 2$  swings altogether.
- (ii) A voter  $i$  with  $w_i = 2 \subseteq N \setminus \{i\}$  renders a losing coalition  $S$  winning by joining if either  $S = \{w_j\}$  with  $w_j = 3$  or  $2$ , or  $S = \{j, k\}$  with  $w_j = 2$  and  $w_k = 1$ . There are  $a + (n - a - 2)$  coalitions of the former and  $n - a - 2$  coalitions of the latter type, amounting to  $2n - a - 4$  swings altogether.
- (iii) Voter  $n$  with  $w_n = 1$  renders a losing coalition  $S \subseteq N \setminus \{i\}$  winning by joining if  $S = \{j\}$  with  $w_j = 3$ . There are  $a$  such coalitions.

□

Writing  $\lfloor x \rfloor$  to denote the largest integer not greater than  $x$ , and  $x \bmod y$  to denote the integer remainder when  $x$  is divided by  $y$ , we have the following finding for distances in the  $d_1$ -metric:

**Lemma 2** Choose

$$a(n) = \begin{cases} \lfloor \frac{6n}{7} \rfloor & \text{if } n \bmod 7 \in \{1, 2, 3\}, \\ \lfloor \frac{6n}{7} \rfloor - 1 & \text{if } n \bmod 7 \in \{0, 4, 5, 6\}, \end{cases}$$

and consider

$$v^n = (2n - a - 4; \underbrace{3, \dots, 3}_{a(n) \text{ threes}}, \underbrace{2, \dots, 2}_{n-a(n)-1 \text{ twos}}, 1).$$

Then

$$\lim_{n \rightarrow \infty} n \cdot (d_1(B(v^n), \beta^n)) = \frac{1}{2}.$$

**Proof** Suppose that  $n \bmod 7 = 0$ , i.e.,  $n = 7k + 0$  for some  $k \in \mathbb{N}$ . Then  $a = 6k - 1$  and Lemma 1 yields swing numbers of  $2 \cdot 7k - (6k - 1) - 2$ ,  $2 \cdot 7k - (6k - 1) - 4$ , and  $6k - 1$  for the three voter types, respectively. This implies a total number of

$$(6k - 1)(8k - 1) + (n - 6k)(8k - 3) + (6k - 1) = 56k^2 - 11k$$

swings, and hence a PBI vector of

$$B(v^{7k}) = \frac{1}{56k^2 - 11k} \left( \underbrace{8k - 1, \dots, 8k - 1}_{6k-1 \text{ times}}, \underbrace{8k - 3, \dots, 8k - 3}_k, 6k - 1 \right).$$

This yields

$$B_i(v^{7k}) - \beta_i^{7k} = \begin{cases} -\frac{1}{(14k-1)k(56k-11)} & \text{if } w_i = 3, \\ \frac{28k-3}{(14k-1)k(56k-11)} & \text{if } w_i = 2, \\ -\frac{28k^2-9k+1}{(14k-1)k(56k-11)} & \text{if } w_i = 1, \end{cases}$$

and summing the absolute values of these figures up for the  $6k - 1$  voters with weight 3, the  $k$  voters with weight 2, and the final voter  $n$  one obtains

$$\|B_i(v^n) - \beta^n\|_1 = \frac{2(28k - 3)}{(14k - 1)k(56k - 11)}$$

in case of  $n = 7k$ . This number and results of the similarly tedious computations when  $n \bmod 7 = 1, \dots, 6$  are summarized in Table 11. For each of the seven cases one easily sees that the deviations tend to  $\frac{1}{14k}$ , which is equivalent to  $\frac{1}{2n}$ . □

$n = 7k +$	$B_i(v^n) - \beta_i$ for			$\ B(v^n) - \beta\ _1$
	$w_i = 3$	$w_i = 2$	$w_i = 1$	
0	$-\frac{1}{(14k-1)k(56k-11)}$	$\frac{28k-3}{(14k-1)k(56k-11)}$	$-\frac{28k^2-9k+1}{(14k-1)k(56k-11)}$	$\frac{2(28k-3)}{(14k-1)k(56k-11)}$
1	0	$\frac{2k(14k+1)}{7(4k+1)}$	$-\frac{2(14k+1)}{7(4k+1)}$	$\frac{14k+1}{7(4k+1)}$
2	$\frac{1}{(14k+3)(56k^2+19k+2)}$	$\frac{7(4k+1)}{(14k+3)(56k^2+19k+2)}$	$-\frac{28k^2+13k+1}{(14k+3)(56k^2+19k+2)}$	$\frac{2(28k^2+13k+1)}{(14k+3)(56k^2+19k+2)}$
3	$\frac{1}{(14k+5)(28k^2+17k+3)}$	$\frac{2(7k+3)}{(14k+5)(28k^2+17k+3)}$	$-\frac{2(7k^2+6k+1)}{(14k+5)(28k^2+17k+3)}$	$\frac{4(7k^2+6k+1)}{(14k+5)(28k^2+17k+3)}$
4	$-\frac{1}{7(2k+1)(14k^2+14k+3)}$	$\frac{14k+5}{14(2k+1)(14k^2+14k+3)}$	$-\frac{14k^2+7k+1}{14(2k+1)(14k^2+14k+3)}$	$\frac{14k^2+19k+5}{7(2k+1)(14k^2+14k+3)}$
5	$-\frac{3}{(14k+9)(56k^2+71k+21)}$	$\frac{28k+15}{(14k+9)(56k^2+71k+21)}$	$-\frac{28k^2+25k+6}{(14k+9)(56k^2+71k+21)}$	$\frac{2(28k^2+43k+15)}{(14k+9)(56k^2+71k+21)}$
6	$-\frac{1}{(14k+11)(28k^2+43k+16)}$	$\frac{2(7k+5)}{(14k+11)(28k^2+43k+16)}$	$-\frac{2(7k^2+9k+3)}{(14k+11)(28k^2+43k+16)}$	$\frac{4(7k^2+12k+5)}{(14k+11)(28k^2+43k+16)}$

Table 11:  $d_1$ -distances between  $\beta^n$  and PBI of game  $v^n \in \mathcal{W}$  in Lemma 2 (with  $k \in \mathbb{N}$ )

In case of the  $d_\infty$ -metric, choose

$$a(n) = \lfloor (n+1)/3 \rfloor + \lfloor n/3 \rfloor - 1.$$

In each corresponding game  $v^n$  (see Lemma 2) roughly two thirds of the players have weight 3 each, roughly one third have weight 2, and a single player has weight 1. The games  $v^n$  result in very good solutions of the inverse problem for  $n < 8$  and the best ones we could find for  $n \geq 8$ . Using this in order to obtain an upper bound one can verify the following result in perfect analogy to Lemma 2:

**Lemma 3** *The weighted voting game  $v^{***} \in \mathcal{W}$  whose PBI minimizes  $d_\infty$ -distance to  $\beta^n$  satisfies*

$$d_\infty(v^{***}, \beta^n) \leq b(n) = \begin{cases} \frac{8n-9}{n(4n-7)(2n-1)} & \text{if } n \bmod 3 = 0, \\ \frac{8n-23}{(4n^2-5n-8)(2n-1)} & \text{if } n \bmod 3 = 1, \\ \frac{4}{4n^2-1} & \text{if } n \bmod 3 = 2 \end{cases}$$

for  $n \geq 8$ .

Note that the indicated bound tends to  $\frac{1}{n^2}$ , i.e.,  $\lim_{n \rightarrow \infty} b(n) / \frac{1}{n^2} = 1$ .

## References

- Alon, N. and P. H. Edelman (2010). The inverse Banzhaf problem. *Social Choice and Welfare* 34(3), 371–377.
- Aziz, H., M. Paterson, and D. Leech (2007). Efficient algorithm for designing weighted voting games. In *Multitopic Conference, 2007. INMIC 2007. IEEE International*, pp. 1–6. Available at <http://eprints.dcs.warwick.ac.uk/1547/>.
- Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review* 19(2), 317–343.
- Carreras, F. and J. Freixas (1996). Complete simple games. *Mathematical Social Sciences* 32(2), 139–155.
- Chang, P.-L., V. C. Chua, and M. Machover (2006). L S Penrose's limit theorem: Tests by simulation. *Mathematical Social Sciences* 51(1), 90–106.
- De Nijs, F. and D. Wilmer (2012). Evaluation and improvement of Laruelle-Widgrén inverse Banzhaf approximation. mimeo. Available at <http://arxiv.org/abs/1206.1145>.
- Deĭneko, V. G. and G. J. Woeginger (2006). On the dimension of simple monotonic games. *European Journal of Operational Research* 170(1), 315–318.
- Dubey, P. and L. Shapley (1979). Mathematical properties of the Banzhaf power index. *Mathematics of Operations Research* 4(2), 99–131.

- Fatima, S., M. Wooldridge, and N. Jennings (2008). An anytime approximation method for the inverse Shapley value problem. In *Proceedings of the 7th International Conference on Autonomous Agents and Multi-Agent Systems, Estoril, Portugal*, pp. 935–942. Available at <http://eprints.ecs.soton.ac.uk/15131>.
- Felsenthal, D. and M. Machover (1998). *The Measurement of Voting Power – Theory and Practice, Problems and Paradoxes*. Cheltenham: Edward Elgar.
- Freixas, J. and S. Kurz (2011). On minimal integer representations of weighted games. mimeo. Available at <http://arxiv.org/abs/1103.0868>.
- Freixas, J. and X. Molinero (2009). On the existence of a minimum integer representation for weighted voting systems. *Annals of Operations Research* 166, 243–260.
- Freixas, J. and X. Molinero (2010). Weighted games without a unique minimal representation in integers. *Optimization Methods and Software* 25(2), 203–215.
- Grilli di Cortona, P., C. Manzi, A. Pennisi, F. Ricca, and B. Simeone (1999). *Evaluation and optimization of electoral systems*. SIAM Monographs on Discrete Mathematics and Applications. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).
- Isbell, J. R. (1956). A class of majority games. *Quarterly Journal of Mathematics* 7(1), 183–187.
- Kaniovski, S. (2008). The exact bias of the Banzhaf measure of power when votes are neither equiprobable nor independent. *Social Choice and Welfare* 31(2), 281–300.
- Keijzer, B. d. (2009). On the design and synthesis of voting games. Master’s thesis, Delft University of Technology.
- Keijzer, B. d., T. Klos, and Y. Zhang (2010). Enumeration and exact design of weighted voting games. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, Volume 1, pp. 391–398.
- Kirsch, W., W. Słomczyński, and K. Życzkowski (2007). Getting the votes right. *European Voice* (3–9 May), 12.
- Kurth, M. (2008). Square root voting in the Council of the European Union: Rounding effects and the Jagiellonian Compromise. Available at <http://arxiv.org/abs/0712.2699>.
- Kurz, S. (2012a). On minimum sum representations for weighted voting games. *Annals of Operations Research* (forthcoming).
- Kurz, S. (2012b). On the inverse power index problem. *Optimization* (forthcoming). DOI:10.1080/02331934.2011.587008.
- Kurz, S., N. Maaser, and S. Napel (2012). On the egalitarian weights of nations. mimeo, University of Bayreuth.
- Laruelle, A. and F. Valenciano (2008). *Voting and Collective Decision-Making*. Cambridge: Cambridge University Press.
- Laruelle, A. and M. Widgrén (1998). Is the allocation of power among EU states fair? *Public Choice* 94(3-4), 317–340.
- Leech, D. (2002a). Designing the voting system for the EU Council of Ministers. *Public Choice* 113(3-4), 437–464.
- Leech, D. (2002b). Voting power in the governance of the International Monetary Fund. *Annals of Operations Research* 109(1), 375–397.
- Leech, D. (2003). Power indices as an aid to institutional design: The generalised apportionment problem. In M. J. Holler, H. Kliemt, D. Schmidtchen, and M. E. Streit (Eds.), *Jahrbuch für Neue Politische Ökonomie*, Volume 22. Tübingen: Mohr Siebeck.
- Lindner, I. and M. Machover (2004). L. S. Penrose’s limit theorem: Proof of some special cases. *Mathematical Social Sciences* 47(1), 37–49.



- Lindner, I. and G. Owen (2007). Cases where the Penrose limit theorem does not hold. *Mathematical Social Sciences* 53(3), 232–238.
- Lucas, W. F. (1992). *Fair Voting: Weighted Votes for Unequal Constituencies*. HISTOMAP Module 19. Lexington, MA: COMAP.
- Neyman, A. (1982). Renewal theory without replacements. *Annals of Probability* 10(2), 464–481.
- Nurmi, H. (1982). The problem of the right distribution of voting power. In M. J. Holler (Ed.), *Power, Voting, and Voting Power*. Würzburg: Physica-Verlag.
- Pennisi, A., F. Ricca, P. Serafini, and B. Simeone (2007). Amending and enhancing electoral laws through mixed integer programming in the case of Italy. In E. Yashin (Ed.), *Proceedings of the 8th International Conference on Economic Modernization and Social Development*. HSE, Moscow.
- Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society* 109(1), 53–57.
- Penrose, L. S. (1952). *On the Objective Study of Crowd Behaviour*. London: H. K. Lewis & Co.
- Ricca, F., A. Scozzari, P. Serafini, and B. Simeone (2012). Error minimization methods in biproportional apportionment. *TOP (forthcoming)*. Available at <http://sole.dimi.uniud.it/~paolo.serafini/TOPMINERR.pdf>.
- Shapley, L. S. and M. Shubik (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review* 48(3), 787–792.
- Słomczyński, W. and K. Życzkowski (2006). Penrose voting system and optimal quota. *Acta Physica Polonica B* 37, 3133–3143.
- Słomczyński, W. and K. Życzkowski (2007). From a toy model to the double square root system. *Homo Oeconomicus* 24(3/4), 381–399.
- Słomczyński, W. and K. Życzkowski (2011). Square root voting system, optimal threshold and  $\pi$ . Available at <http://arxiv.org/abs/1104.5213>.
- Taylor, A. D. and W. S. Zwicker (1999). *Simple Games*. Princeton, NJ: Princeton University Press.