The Art and Beauty of Voting Power

Sascha Kurz
Dept. of Mathematics, University of Bayreuth, Germany
sascha.kurz@uni-bayreuth.de

Alexander Mayer
Dept. of Economics, University of Bayreuth, Germany
alexander.mayer@uni-bayreuth.de

Stefan Napel
Dept. of Economics, University of Bayreuth, Germany
stefan.napel@uni-bayreuth.de

July 29, 2022

Abstract

We exhibit the hidden beauty of weighted voting and voting power by applying a generalization of the Penrose-Banzhaf index to social choice rules. Three players who have multiple votes in a committee decide between three options by plurality rule, Borda's rule, antiplurality rule, or one of the many scoring rules in between. A priori influence on outcomes is quantified in terms of how players' probabilities to be pivotal for the committee decision compare to a dictator. The resulting numbers are represented in triangles that map out structurally equivalent voting weights. Their geometry and color variation reflect fundamental differences between voting rules, such as their inclusiveness and transparency.

Keywords: weighted voting · weighted committee games · scoring rules · simple voting games · collective choice

JEL codes: D71 · C71 · C63

We thank an anonymous referee for his or her thoughtful reading and several helpful suggestions.

1 Introduction

Individual voting rights entail a potential to affect collective decisions. Greater numbers of votes controlled by large shareholders, party leaders, delegates to a council or committee, etc. typically increase the respective influence. It is not trivial, though, to tell how much greater the influence of, e.g., a voter wielding 25% of all votes is compared to one wielding only 5%. This holds even when an a priori perspective is adopted, meaning that one purposely leaves aside personal affiliations between the voters and empirical preference information. Various indices try to rigorously quantify voting power in order to address this problem.

For binary collective decision making of the *yes*-or-*no* kind – formalized by simple voting games (cf. von Neumann and Morgenstern 1953, Ch. 10 or Taylor and Zwicker 1999) – prominent examples of voting power indices are the Penrose-Banzhaf, Shapley-Shubik, and Holler-Packel indices (cf. Penrose 1946, Banzhaf 1965, Shapley and Shubik 1954, Holler and Packel 1983). Some of them have been extended to non-binary settings such as the determination of a winner from a set of more than two options by alternative methods of social choice (cf. Kurz, Mayer, and Napel 2021, for instance). The respective winner could be a particular law selected from multiple legal drafts, the managing director of the IMF chosen from a shortlist of three nominees, a presidential candidate who is picked from various primary contenders, and so on.

Applied to a particular voting body such as a parliament or party convention, the IMF Executive Board, the EU Council or the US Electoral College,[1] power indices illuminate some of the discrete mathematical structure that underlies collective choices. They identify possible swings between losing and winning coalitions, which make outcomes depend on a voter's behavior, or more generally they measure potential variation of the winning candidate that derives from a single voter's input to the decision process. Light is cast on either a specific institutional arrangement, when the focus is on influence of distinct members of a voting body relative to another, or on multiple competing arrangements. One may evaluate, for instance, the implications of a change of the majority threshold or a switch from plurality voting to a

[1] We recommend the contributions in Holler and Nurmi (2013) for a good overview of typical applications of power indices. Somewhat atypical applications are discussed by Kovacic and Zoli (2021) and Napel and Welter (2021). Napel (2019) provides a short introduction to power measurement with many further references.

runoff system (e.g., see Maskin and Sen 2016 on plurality voting in US presidential primaries).

So power indices yield insights with political or economic meaning. Common questions are: To what extent can a given shareholder control a corporation and may, perhaps, be held responsibility for its actions? Is the voting power of two parties at least approximately proportional to their seat shares in parliament? Is a given allocation of voting rights to delegates from different constituencies (e.g., US states in the Electoral College, member countries in the EU Council, departments in a university senate, etc.) 'fair' under a particular set of normative premises? Etcetera.

This article, however, is *not* pursuing serious questions of any such kind. We here employ a power index for non-binary decisions with a seemingly superficial and primarily visual purpose: we try to convey the hidden beauty of weighted voting and want to exhibit artistic aspects of the power that voters can derive from their voting weights.

The article's main part therefore consists of several pages of color images. Depending on personal taste, they may be of interest and produce enjoyment without any further explanation. At the same time, they represent the result of hours of computer calculations (several weeks, in fact). They give a graphical picture of the formal structure behind collective decision making by three players – individuals or homogeneous groups of voters – on three candidates.

It will be assumed that winners are determined by a plurality vote, an antiplurality vote, or one of the many scoring rules that lie 'in between', such as Borda's voting rule. Other voting methods like the various rules that focus on pairwise majority comparisons (Copeland's rule, Kemeny-Young rule, etc.) are amenable to the same kind of representation. They generate less scintillating results however (cf. Kurz, Mayer, and Napel 2020).

Appreciators of art and beauty without interest in the formal framework are welcome to jump to Section 4. For all others, we will first provide a short introduction to weighted committee games (Section 2).[2] These games generalize traditional weighted voting from binary majority decisions to social choice from any finite num-

[2]See Kurz et al. (2020) for details and related literature: the article defines weighted committee games, characterizes and counts equivalence classes for selected voting rules, and provides lists of structurally distinct committees. Mayer and Napel (2021) does similarly for the special case of scoring rules. Kurz et al. (2021) generalizes the Penrose-Banzhaf and Shapely-Shubik indices to committee games. For a practical application of the framework see Mayer and Napel (2020).

ber of options. We explain the pertinent generalization of the Penrose-Banzhaf power index and how this can literally provide colorful insights into how voting weights determine voters' influence on collective decisions (Section 3). Possible economic and political implications are briefly pointed out in Section 5.

2 Weighted committees and scoring rules

Binary *weighted voting games* involve a set of n players who respectively wield voting weights denoted by $w_1, w_2, \ldots, w_n \geq 0$. The majority threshold or decision quota is set to $q > 0$: the players can jointly pass any proposal that is made to them if the subset of players who support the motion wield a combined voting weight of at least q. For instance, if players 1, 2 and 3 have weights of $w_1 = 40\%$, $w_2 = 35\%$ and $w_3 = 25\%$ and face a quota $q = 51\%$ then at least two players must support a proposal for it to pass. Subsets $S \subseteq N$ of the set N of players with $\sum_{i \in S} w_i \geq q$ are also referred to as winning coalitions, while subsets $T \subseteq N$ with $\sum_{i \in T} w_i < q$ are known as losing coalitions.

Weighted voting games constitute a special kind of (binary) *simple voting games*. The latter do not necessarily require a link between winning or losing to weights and a quota. They merely assume that the full player set N is winning, the empty set $\emptyset$ is losing, and winning is monotonic with respect to set inclusion, i.e., any superset of a winning coalition is also winning.

Simple voting games are commonly specified in set-theoretic terms. This is done either by directly listing a subclass of winning coalitions (typically those that are minimal with respect to set inclusion) or using an indicator function v that takes a set $S \subseteq N$ as its argument and outputs a 1 if and only if S is winning. A simple voting game can, however, also be described as a mapping from the set of all possible profiles of players' preferences over a status quo option a_1 and an alternative motion a_2 that is voted on to the set of possible outcomes, specifying for each preference profile the collective decision a_1 or a_2 that is adopted. *Weighted committee games* follow this route and allow to handle also non-binary decisions.

In particular, the latter consider a finite set N of $n \geq 2$ players and assume that each player $i \in N$ has strict preferences P_i over a set $A = \{a_1, \ldots, a_m\}$ of $m \geq 2$ options that the committee needs to choose from. The set of all $m \cdot (m-1) \cdot \ldots \cdot 1 = m!$ conceivable strict preference orderings on A is denoted by $\mathcal{P}(A)$. Any collective decision rule can then be conceived of as a mathematical mapping $\rho \colon \mathcal{P}(A)^n \to A$. This translates any

preference profile $\mathbf{P} = (P_1, \ldots, P_n)$ into a single winning alternative $a^* = \rho(\mathbf{P})$. The respective combination (N, A, ρ) of a set of voters, a set of alternatives, and a decision rule is referred to as a committee game or as a *committee* for short.

A committee (N, A, ρ) is called a *weighted plurality committee* if the decision rule ρ amounts to each voter i casting $w_i \geq 0$ votes for its favorite option and then selecting the alternative a^* that received the most votes as winner. Similarly, for a *weighted antiplurality committee* the decision rule ρ amounts to each voter i casting $w_i \geq 0$ negative or dissenting votes for its least preferred option and then the alternative a^* that received the fewest dissenting votes becomes the winner. In case of ties we suppose that they are resolved lexicographically: if, for instance, $A = \{a_1, a_2, a_3, a_4\}$ and these alternatives respectively receive 3, 4, 0, and 4 plurality votes from $n = 11$ voters with a weight of $w_i = 1$ each, then a_2 rather than a_4 is chosen. Declaring both a_2 and a_4 to be winners and tossing a coin to reach a resolute decision would be a possibility too. But randomness would complicate the mathematical exposition without changing the illustrations below.

(Weighted) plurality and antiplurality committees are special cases of (weighted) *scoring committees*. These entail the application of a *scoring rule*: the winning candidate or option a^* always is the one that received the highest total score from the voters. Candidates' scores are determined by their positions in each voter's preference ranking and a given vector $\mathbf{s} = (s_1, s_2, \ldots, s_m)$ with $s_1 \geq s_2 \geq \ldots \geq s_m$ and $s_1 \neq s_m$: when voters' weights are $w_1 = \ldots = w_n = 1$, any alternative $a \in A$ receives s_1 points for every voter that ranks a first, s_2 points for every voter that ranks a second, and so on. When the voters have non-uniform weights $w_1, \ldots, w_n \geq 0$, the respective points derived from how voter $i \in N$ ranks the alternatives are multiplied by w_i.

For illustration, suppose that a committee – perhaps the board of a sports club – involves four voter groups, i.e., players $N = \{1, 2, 3, 4\}$, with group 1 wielding 5 votes, group 2 having 4 votes, group 3 wielding 3 votes, and group 4 having only one vote. The weights are summarized by $\mathbf{w} = (5, 4, 3, 1)$. The voters must select one of three candidates, say, Ann, Bob, or Clara, to lead their club.

Let the players' preferences $\mathbf{P} = (P_1, P_2, P_3, P_4)$ rank the candidates as in the following table:

	P_1	P_2	P_3	P_4
1st best	Bob	Ann	Ann	Bob
2nd best	Clara	Clara	Bob	Clara
3rd best	Ann	Bob	Clara	Ann

Using the scoring vector $\mathbf{s} = (1, 0, 0)$ amounts to a weighted plurality vote: Ann receives a total score of $5 \cdot 0 + 4 \cdot 1 + 3 \cdot 1 + 1 \cdot 0 = 7$; Bob's score is $5 \cdot 1 + 4 \cdot 0 + 3 \cdot 0 + 1 \cdot 1 = 6$; and Clara, being ranked first by nobody, gets a score of 0. Ann wins.

Had above committee used the scoring vector $\mathbf{s} = (1, 1, 0)$ instead, Clara would have won with a score of 10 vs. 7 for Ann and 9 for Bob. The latter vector $\mathbf{s}$ is equivalent to conducting an antiplurality vote because minimizing the number of dissenting votes is the same as maximizing the number of non-dissenting votes captured by $\mathbf{s} = (1, 1, 0)$.

An example of a voting rule in between plurality and antiplurality is *Borda's rule*: voters state their full preferences and each candidate a receives as many points from a given voter i as there are candidates that i ranks below a. For instance, Bob would receive 2 points for each vote wielded by group 1, 0 points from group 2, 1 point for each of the votes held by group 3, and again 2 points from group 4. This gives Bob a total Borda score of $5 \cdot 2 + 4 \cdot 0 + 3 \cdot 1 + 1 \cdot 2 = 15$. That number is greater than the analogous figures of 14 for Ann and 10 for Clara. So Bob would win if scoring vector $\mathbf{s} = (2, 1, 0)$ or Borda's rule were used.

Maximizing the total score given the scoring vector $\mathbf{s} = (2, 1, 0)$ is equivalent to maximizing the total score for vectors $\mathbf{s}' = (1, 1/2, 0)$ or $\mathbf{s}'' = (4, 3, 2)$. Vector $\mathbf{s}'$ merely halves above numbers, while preserving the order of Ann's, Bob's and Clara's totals. Similarly, using $\mathbf{s}''$ raises all candidates' scores by $(5 + 4 + 3 + 1) \cdot 2 = 26$ without changing their order. In particular, scoring winners are invariant to positive affine transformations of the adopted scoring vector $\mathbf{s}$. Hence, whenever a committee picks a winner from three candidates by a scoring rule – plurality, antiplurality, Borda, or any other rule that determines the winner by evaluating the candidates' positions in the applicable preference profile $\mathbf{P}$ with decreasing scores – it is without loss of generality to suppose a vector $\mathbf{s} = (1, s, 0)$ such that $0 \leq s \leq 1$.

When a committee with player set N and voting weights $\mathbf{w} = (w_1, \ldots, w_n)$ decides on a set A of $m = 3$ alternatives and uses a decision rule ρ that amounts to applying the scoring vector $\mathbf{s} = (1, s, 0)$, we write $(N, A, r^s|\mathbf{w})$ instead of (N, A, ρ). We refer to

such committee as a (weighted) *s-scoring committee* (see Mayer and Napel 2021).

We have seen that the special s-scoring committees with $s = 1$, $s = 1/2$, and $s = 0$ amount to weighted plurality, Borda, and antiplurality committees. As above example illustrates, the respective committees differ for the considered voting weights $\mathbf{w} = (5, 4, 3, 1)$. Namely, they select a different winner from three candidates for at least some configuration of preferences. Similarly, two plurality committees ($s = 0$) are different depending on whether weights $\mathbf{w} = (5, 4, 3, 1)$ or weights $\mathbf{w}' = (5, 1, 1, 1)$ apply to the players (club members, shareholders, parties, etc.): for the profile $\mathbf{P}$ at hand, Bob rather than Ann would be selected if $\mathbf{w}$ were replaced by $\mathbf{w}'$.

We call two committees (N, A, ρ) and (N, A, ρ') that never select different winners from set A – no matter which preference profile $\mathbf{P} = (P_1, \ldots, P_n)$ is considered – *equivalent*. This means that the respective mappings $\rho : \mathcal{P}(A)^n \to A$ and $\rho' : \mathcal{P}(A)^n \to A$ are identical, denoted by $\rho \equiv \rho'$. We can have $\rho \equiv \rho'$ even though the verbal descriptions of ρ and ρ' may differ. For instance, ρ may be described as plurality voting with weights $\mathbf{w}' = (5, 1, 1, 1)$ and ρ' as dictatorship of voter 1: the committee in either case always chooses the alternative that is ranked first according to P_1.

When two s-scoring committees $(N, A, r^s|\mathbf{w})$ and $(N, A, r^s|\mathbf{w}')$ with $\mathbf{w} \neq \mathbf{w}'$ are equivalent, i.e., $r^s|\mathbf{w} \equiv r^s|\mathbf{w}'$, we learn that it does not matter which of the two voting weight distributions prevails: decisions will coincide. From the perspective of an outsider who does not care about the labeling of the players, this is also true if weights $\mathbf{w}''$ are used that only label the players differently than $\mathbf{w}$.

Consider, for instance, $\mathbf{w}'' = (1, 3, 4, 5)$ instead of $\mathbf{w} = (5, 4, 3, 1)$ in our example. This represents the same abstract decision structure except that player numbers have changed. In particular, the situation for the preferences $\mathbf{P} = (P_1, P_2, P_3, P_4)$ depicted in the table above for weights $\mathbf{w}$ (with Ann winning under plurality rule, Clara under antiplurality rule, etc.) is the same as that with weights $\mathbf{w}''$ and preferences $\mathbf{P}'' = (P_4, P_3, P_2, P_1)$. We then say that $\mathbf{w}$ and $\mathbf{w}''$ are *structurally equivalent* under the considered s-scoring rule: the implied mappings $r^s|\mathbf{w}$ and $r^s|\mathbf{w}''$ become equivalent after suitably relabeling the players.

Having fixed a scoring rule, such as r^s for $s = 1$, the set of all weights $\mathbf{w} = (w_1, \ldots, w_n)$ that are structurally equivalent to a given reference distribution of weights $\tilde{\mathbf{w}} = (\tilde{w}_1, \ldots, \tilde{w}_n)$ can be grouped together and form an *equivalence class* of weights: if two weight distributions $\mathbf{w} \neq \mathbf{w}'$ belong to the same class, the corresponding s-scoring committees always produce identical decisions (once labels of the players are

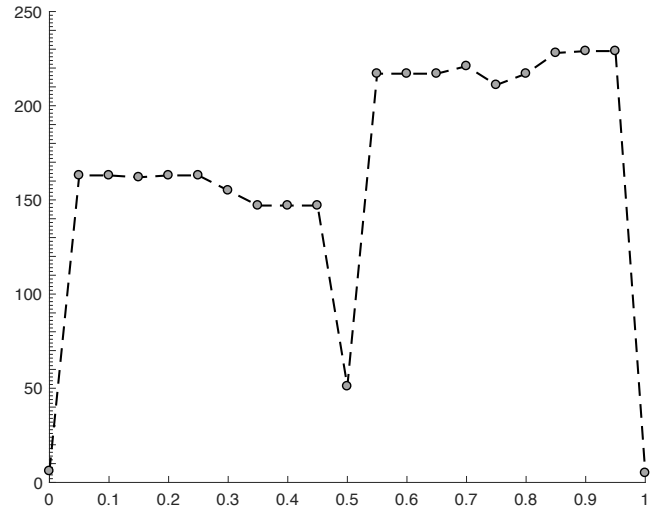

Figure 1: Number of s-scoring committees for $m = 3$ and $s \in \{0, 0.05, 0.1, \ldots, 1\}$

harmonized). If the weight distributions belong to different classes, there exists at least some preference configuration $\mathbf{P}$ that results in different committee decisions.

For antiplurality rule ($s = 1$) and three players ($n = 3$), it turns out that there are only five different equivalence classes – namely those that correspond to reference weights of $\tilde{\mathbf{w}} = (1, 0, 0)$, $(1, 1, 0)$, $(1, 1, 1)$, $(2, 1, 1)$ and $(2, 2, 1)$. Any other distribution of weights among three players is structurally equivalent to one of these, i.e., leads to the same decisions after suitable relabeling (cf. Kurz et al. 2020). Similarly, there are only six structurally different plurality committees for three players. The respective reference weights equal the five just listed for antiplurality rule in addition to $\tilde{\mathbf{w}} = (3, 2, 2)$.

The numbers of structurally distinct s-scoring committees for $s = 1/2$ (Borda) and, more pronouncedly, for $0 < s < 1/2$ or $1/2 < s < 1$ are much higher than those for $s = 0$ and $s = 1$. Exact values have not been published for all s yet, but Mayer and Napel (2021) provide the numbers of equivalence classes for all s that are integer multiples of $1/20$. These numbers range up to 229 and exhibit an M-shaped pattern reproduced in Figure 1.

Knowing that a given weight distribution among three players structurally amounts to, say, $(2, 1, 1)$ can simplify the analysis of the respective committee: the distribution of voting power is as if weights were $(2, 1, 1)$. So are players' manipulation incentives, strategic voting equilibria, the scope for voting paradoxes, etc.

Alas, it is generally an arduous task to determine for a given weight distribution

$\mathbf{w}$ to which scoring equivalence class it belongs (for fixed vector $\mathbf{s}$). The respective equivalence classes form convex polyhedra that are defined by linear inequalities. When we consider three players and restrict attention to their relative voting weights $\bar{\mathbf{w}} = \mathbf{w}/(w_1 + w_2 + w_3)$ (so that $\bar{w}_1 + \bar{w}_2 + \bar{w}_3 = 100\%$), the polyhedra are either points, lines, or area pieces bounded by lines. They jointly cover the triangle highlighted in Figure 2 – the so-called 2-dimensional unit simplex.

Suppose that we have a 'map' of all equivalence classes in this simplex. Then one may start out with an arbitrary weight distribution $\mathbf{w} = (w_1, w_2, w_3)$, compute the corresponding relative weight distribution $\bar{\mathbf{w}}$, locate it in the simplex map, and now identify the applicable class.

Such simplex maps can indeed be constructed. Namely, the figures depicted in Section 4 show the links from all possible weights to equivalence classes, except that we leave out a legend that would identify the respective equivalence classes via reference distributions of weights.[3]

3 Voting power and color

Before we present our figures, let us explain how the selected coloring relates to the a priori voting power of the three players involved. An index of voting power generally takes a description of a voting body – a simple voting game or, in our case, an s-scoring committee of three players deciding on three options – as its input and produces a real number for each player as its output. The respective numbers reflect the players' influence on collective decisions according to a specific conception of a player being influential. They are based on specific probabilistic assumptions about the voting situations faced by the players.

The popular Penrose-Banzhaf power index (Penrose 1946; Banzhaf 1965) equates 'being influential' with the possibility of the considered player changing or swinging a decision at hand if the preferences and behavior of all other players are held fixed. This possibility arises, for instance, in unweighted binary majority decisions when

[3]See Figure 5 in Mayer and Napel (2021). It provides a map of the 51 Borda equivalence classes for $n = m = 3$ and $w_1 \geq w_2 \geq w_3$ with a reference distribution of weights for each class. Maps could be constructed for more than three alternatives, too, but the higher number of preference profiles and perturbations has considerable computational costs. Equivalence classes for scores $0 < s < 1$ change fast: the number of Borda classes rises from 51 to 505 and $\geq$2251 for $m = 3$, 4 and 5 (Kurz et al. 2020). Corresponding analogues of Figure 3 exhibit smoother transitions with even more shades of color.

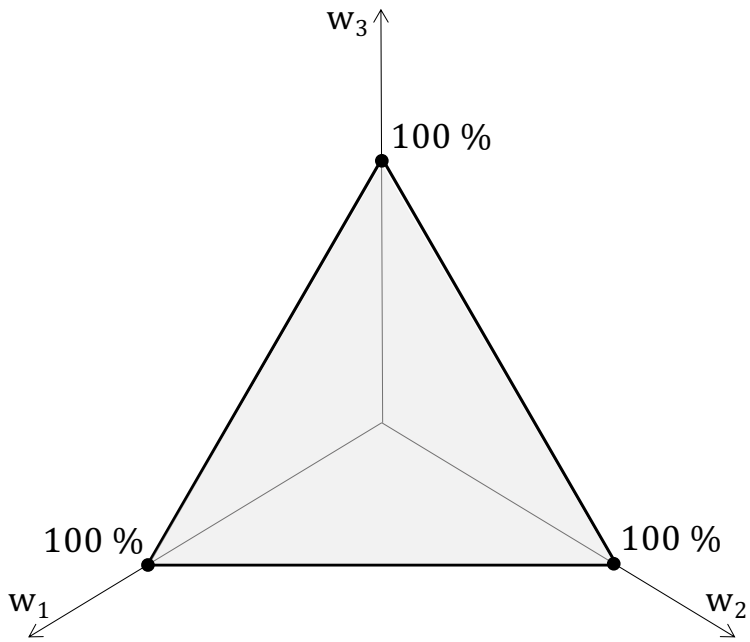

Figure 2: Relative weight distributions among three voters in the unit simplex

the other players are split equally into a *yes*-camp and a *no*-camp, so that the player in question can determine which option receives the majority. In other words, the considered player's vote is pivotal for the outcome. The Penrose-Banzhaf index assesses the probability of pivotality events for a given player under the assumption that all other players vote *yes* or *no* with equal probabilities and independently of another. This is equivalent to assuming that all *yes*-or-*no* configurations or all coalitions $S \subseteq N$ of players who support a change of the status quo are equally likely.[4]

When the collective decision requires a choice from three or more options, such as candidates Ann, Bob and Clara above, one can similarly identify 'being influential' with the committee's decision depending on or varying in the considered player's preferences. For instance, if in our sports club example, player 3 did not rank Ann before Bob and Clara but had Bob as its first preference before Ann and Clara, then

[4]The Shapley-Shubik index (Shapley and Shubik 1954) belongs to the same family of indices but supposes positive correlation of *yes*-or-*no* preferences across voters. In technical terms, it assumes an *impartial anonymous culture (IAC)*, while the Penrose-Banzhaf index reflects an *impartial culture (IC)*. The Holler-Packel index (Holler and Packel 1983) does not consider all coalitions of *yes*-supporters but only *minimal winning coalitions* $S \subseteq N$ in which every *yes*-vote is pivotal for the outcome.

the plurality winner would be Bob rather than Ann. Hence, player 3 is pivotal in the considered voting situation. So is player 2, whereas players 1 and 4 have no scope to individually change the winner for the given preferences of the respective others. Players 1 and 4 are, however, pivotal for many other preference configurations $\mathbf{P} = (P_1, P_2, P_3, P_4) \in \mathcal{P}(A)^4$ that may arise. So also they are influential from an a priori perspective that considers all preference combinations to be possible.

Just as the Penrose-Banzhaf index is based on independent and equiprobable *yes*-or-*no* preferences in the binary case, players' preferences P_i will be assumed to be distributed independently from the others also for more than two options, assigning equal probability to each of the conceivable strict orderings of the options. When assessing the a priori influence implications of voting weights $\mathbf{w} = (5, 4, 3, 1)$, for instance, we will therefore assume player 1 to be as likely to rank (i) Ann before Bob before Clara, as to rank (ii) Ann before Clara before Bob, (iii) Bob before Ann before Clara, (iv) Bob before Clara before Ann, (v) Clara before Ann before Bob, or (vi) Clara before Bob before Ann. We allow the same six possibilities to arise independently also for players 2, 3 and 4. So there are a total of $6 \cdot 6 \cdot 6 \cdot 6 = 6^4 = 1296$ different voting situations that are equally probable when four players decide on three options.

We will focus here on only $n = 3$ players who decide on $m = 3$ options, so that $(m!)^n = 6^3 = 216$ different preference profiles $\mathbf{P} = (P_1, P_2, P_3)$ are possible. Holding a particular player of interest, say player $i \in \{1, 2, 3\}$, fixed, we check for each profile whether a change of i's ranking P_i to one of the alternative five rankings P_i' would make a difference to the collective decision. Whenever this is the case, i.e., the profile $\mathbf{P}'$ that is created by replacing P_i in $\mathbf{P}$ by P_i' yields a collective decision $r^s|\mathbf{w}(\mathbf{P}') \neq r^s|\mathbf{w}(\mathbf{P})$, we count this as a *swing position* for player i. Player i's power index value is then taken to be the ratio of actual swing positions to the maximum conceivable number of such positions.

The latter corresponds to the number of swing positions that a dictator player would hold. For each of the 216 possible preference profiles of three voters on three options, the collective choice under a dictatorship equals the dictator's most preferred alternative. So starting from given preferences of the dictator over three candidate, say ranking (i) above, a switch to four of the five alternative rankings produces a different winner – namely preference changes from (i) to (iii), (iv), (v) or (vi). These perturbations involve a different top preference than (i) and let Bob or Claire win instead of Ann. It follows that a dictator player has $216 \cdot 4 = 864$ swing

positions: they derive from considering 216 distinct voting situations and, for each situation, checking all five ways to spontaneously change the dictator's ranking of the options. Such a change might reflect an idiosyncratic change of mind, perhaps due to new private information on the candidates; it might arise because the player is corrupt and sells its vote to an outside agent; it could simply be a demonstration of the player's power; etc. If a player i in the actual scoring committee should have 432 swing positions, then the corresponding ratio $432/864 = 1/2$ reveals i to be half as powerful as a dictator would be.

Expressing this reasoning in general mathematical terms leads to the (generalized) Penrose-Banzhaf index

$$\mathcal{PBI}_i(N, A, \rho) = \frac{\sum_{\mathbf{P}\in\mathcal{P}(A)^n} \sum_{P_i'\neq P_i\in\mathcal{P}(A)} \Delta\rho(\mathbf{P}; P_i')}{m!^n \cdot (m! - (m-1)!)} \tag{1}$$

of player i's a priori influence or *voting power* in committee (N, A, ρ), as introduced and axiomatically characterized by Kurz et al. (2021).[5] Here $\Delta\rho(\mathbf{P}; P_i')$ denotes an indicator function that is 1 if $\rho(\mathbf{P}') \neq \rho(\mathbf{P})$, and 0 otherwise. Equation (1) is summing over all voting situations (i.e., all conceivable preference configurations $\mathbf{P}$), counts the number of changes of mind by player i (i.e., perturbations of i's preferences P_i to some $P_i' \neq P_i$) that change the collective decision, and then divides this by the total number of swing positions for a hypothetical dictator player ($6^3 \cdot 4 = 864$ for $m = 3$ options and $n = 3$ players). So for an s-scoring committee $(N, A, r^s|\mathbf{w})$ of three players, the triplet

$$\Big(\mathcal{PBI}_1(N, A, r^s|\mathbf{w}), \mathcal{PBI}_2(N, A, r^s|\mathbf{w}), \mathcal{PBI}_3(N, A, r^s|\mathbf{w})\Big),$$

or $\mathcal{PBI}$ for short, quantifies the distribution of voting power in the committee in terms of how close the individual players are to having dictatorial influence. In our sports club example, the power distribution amounts to $\mathcal{PBI} \approx (0.6296, 0.4815, 0.4444, 0.0741)$. That is, player 1 has about 63% of the influence of a dictator while player 4 only has about 7% of the influence of a dictator. The influence of players two and three is just under 50% of that of a dictator.

[5]Replacing the IC assumption that underlies eq. (1) by the IAC assumption (cf. fn. 4) naturally generalizes the Shapley-Shubik index (see Kurz et al. 2021). By contrast, generalization of the Holler-Packel index would first require the definition of a suitable analogue of minimal winning coalitions in weighted committee games. One possibility would be to study each winning alternative $a \in A$ separately and to consider *a-minimal preference profiles* $\mathbf{P}$ where $\rho(\mathbf{P}) = a$ such that $\rho(\mathbf{P}') \neq a$ for any profile $\mathbf{P}'$ in which a is ranked lower by some voter with constant preferences on subset $A \setminus a$.

For graphical purposes one might now associate player 1's power value $\mathcal{PBI}_1$ with the color red, player 2's power $\mathcal{PBI}_2$ with green and player 3's power $\mathcal{PBI}_3$ with blue. Thus we would have linked the scoring rule r^s for a given value of s and a particular distribution $\mathbf{w}$ of voting weights to a particular color using the common RGB color code. For instance, $\mathcal{PBI}(N, A, r^s|(1,0,0)) = (1,0,0)$ for any $0 \leq s \leq 1$ and this would correspond to bright red color. Or the power distribution $\mathcal{PBI} = (^{588}/_{864}, ^{516}/_{864}, ^{312}/_{864}) \approx (0.6806, 0.5972, 0.3611)$ that is derived by Kurz et al. (2021) for $s = ^1/_2$ and weights $\mathbf{w} = (6,5,3)$ would correspond to a dark khaki color.

Although this would be feasible, the figures in Section 4 will not use exactly this coloring option. We will rather make two modifications: first, we will adopt a structural view on committee equivalences, i.e., we do not consider player labels important. Hence we will give the same color to all six points in the unit simplex that represent relative voting weights of, e.g., $\bar{\mathbf{w}} = (^6/_{14}, ^5/_{14}, ^3/_{14})$ after sorting the weights in decreasing order. This implies that the coloring of the weight simplex will be 3-fold radially symmetric around $\bar{\mathbf{w}} = (^1/_3, ^1/_3, ^1/_3)$, as well as mirror symmetric with the three symmetry axes $\bar{w}_1 = \bar{w}_2$, $\bar{w}_2 = \bar{w}_3$ and $\bar{w}_1 = \bar{w}_3$.

Second, we will apply a transformation when turning power triplets $\mathcal{PBI}$ into RGB levels. The motivation is to make better use of the available color palette, to obtain a somewhat lighter image than by, e.g., associating $\bar{\mathbf{w}}$ with dark khaki, and to represent dictatorial power by the dark blue color that has already been used, e.g., by Kurz et al. (2020).

4 Simplex maps of equivalence classes

All images displayed in this section are derived via the following five steps:

1. We fix a scoring vector $\mathbf{s} = (1, s, 0)$ and consider the corresponding scoring rule r^s for collective decisions on $m = 3$ options by $n = 3$ players.

2. We use a finite grid of rational numbers and let the computer loop through all relative voting weight distributions $\bar{\mathbf{w}}$ with $1 \geq \bar{w}_1 \geq \bar{w}_2 \geq \bar{w}_3 \geq 0$ on this grid.

3. For each of the 282 376 weight distributions $\bar{\mathbf{w}}^k$, $k = 1, 2, \ldots, 282\,376$, on the adopted grid, we compute the Penrose-Banzhaf voting power $\mathcal{PBI}^k$ in the respective weighted s-scoring committee $(N, A, r^s|\bar{\mathbf{w}}^k)$.

4. The obtained triplet $(\mathcal{PBI}_1^k, \mathcal{PBI}_2^k, \mathcal{PBI}_3^k)$ is then transformed into red, green and blue intensities $(R, G, B) = \left(\frac{2 \cdot \mathcal{PBI}_3^k}{\max_k \mathcal{PBI}_3^k}, \frac{\mathcal{PBI}_2^k}{\max_k \mathcal{PBI}_2^k}, \frac{\mathcal{PBI}_1^k - \min_k \mathcal{PBI}_1^k}{1 - \min_k \mathcal{PBI}_1^k}\right)$.

5. For each weight distribution $\bar{\mathbf{w}}$, the six points in the simplex (cf. Figure 2) that structurally correspond to $\bar{\mathbf{w}}$ – that is, $(\bar{w}_1, \bar{w}_2, \bar{w}_3)$, $(\bar{w}_2, \bar{w}_1, \bar{w}_3)$, $(\bar{w}_1, \bar{w}_3, \bar{w}_2)$, etc. – are colored with the RGB intensities given by (R, G, B). For instance, $\mathcal{PBI}$ figures of $(1, 0, 0)$ translate into $(R, G, B) = (0, 0, 1)$ and dark blue color.

It is noteworthy that the distribution of voting power in two s-scoring committees $(N, A, r^s|\mathbf{w})$ and $(N, A, r^s|\mathbf{w}')$ can coincide even though the committees are non-equivalent: players are exactly as influential in either but some preference profiles yield different decisions so that $r^s|\mathbf{w} \not\equiv r^s|\mathbf{w}'$. Some of the illustrations in Figure 3 therefore involve fewer different colors than there are distinct equivalence classes for the considered value of s. Moreover, equivalence classes that are represented by a single point in the simplex like the symmetric distribution of relative voting weights $\bar{\mathbf{w}} = (1/3, 1/3, 1/3)$, or a line – e.g., $\bar{\mathbf{w}} = (x, 1 - x, 0)$ for $0 < x < 1/2$ – may not be visible without magnification. We have manually enlarged them only for $s = 0$ and $s = 1$. Bearing these caveats in mind the colored simplices below provide accurate maps of all equivalence classes of scoring committees that exist for a given value of s.

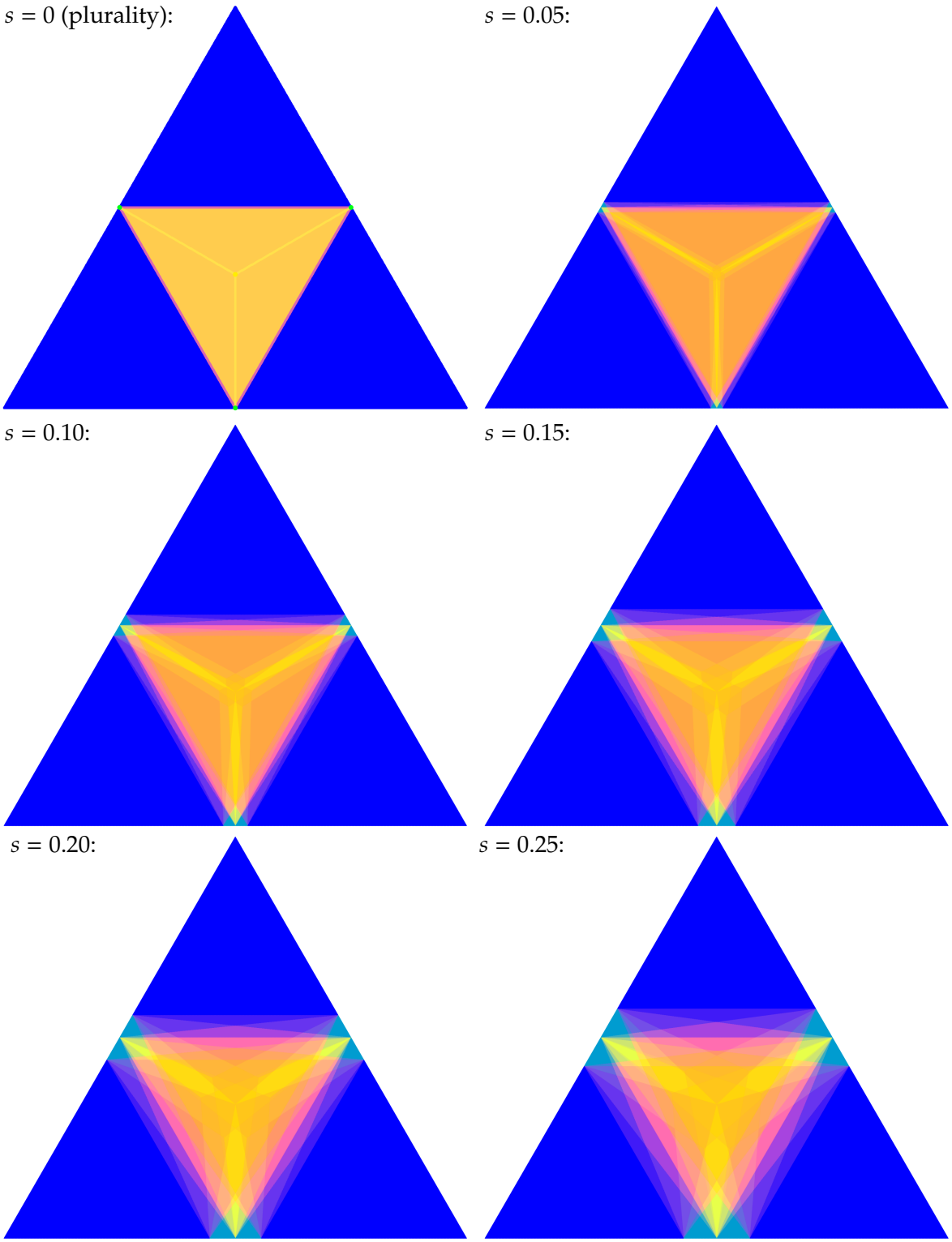

Figure 3: Weighted s-scoring committees in Penrose-Banzhaf coloring

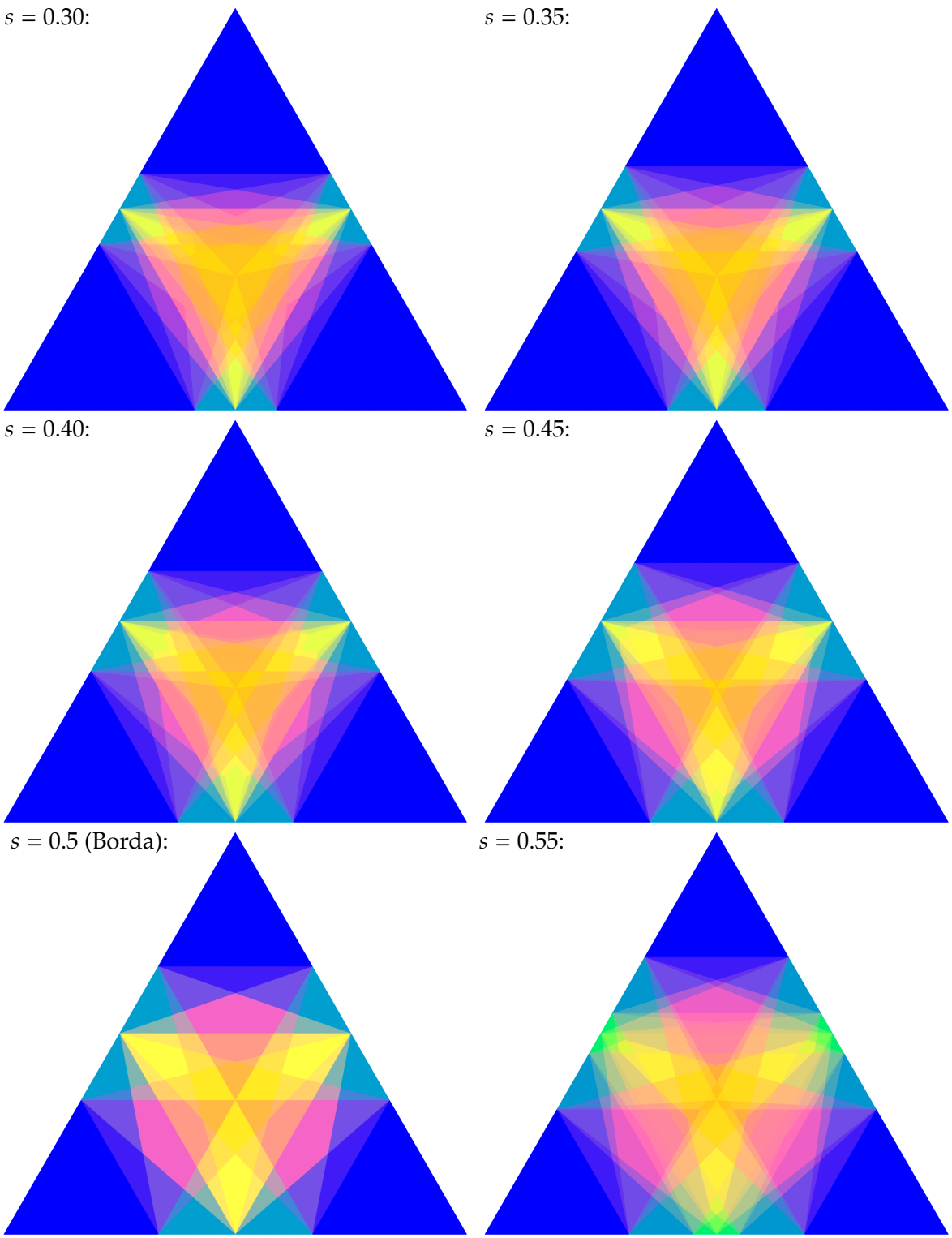

Figure 3 (ctd.): Weighted s-scoring committees in Penrose-Banzhaf coloring

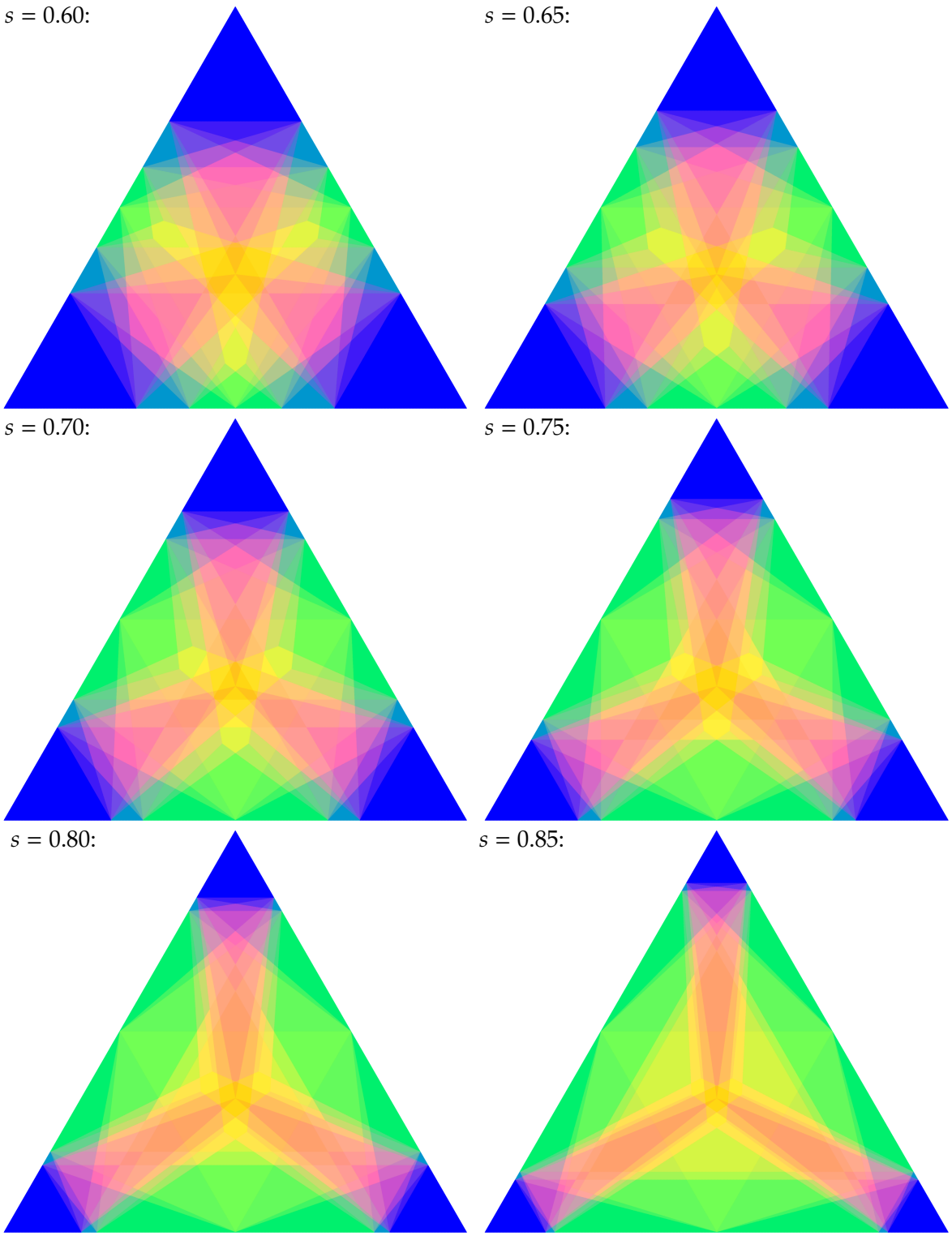

Figure 3 (ctd.): Weighted s-scoring committees in Penrose-Banzhaf coloring

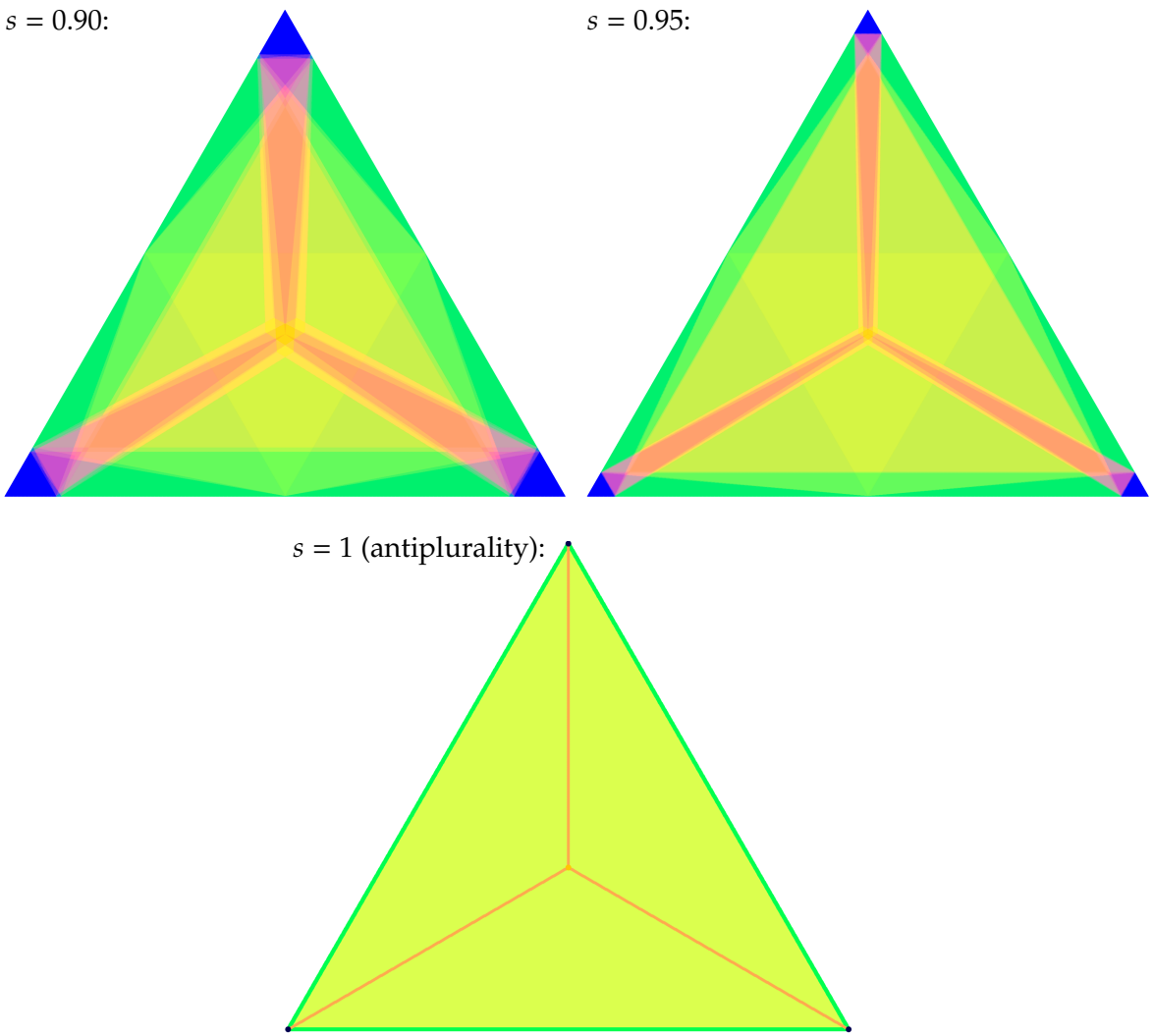

Figure 3 (ctd.): Weighted s-scoring committees in Penrose-Banzhaf coloring

For instance, the large blue triangles inside the panel for $s = 0$, i.e., weighted plurality committees, correspond to $\tilde{\mathbf{w}} = (1, 0, 0)$, i.e., dictatorship of one player. The green midpoints of the simplex's boundary lines represent the equivalence class with $\tilde{\mathbf{w}} = (1, 1, 0)$: two players decide symmetrically, the third never makes a difference. The simplex's light yellow midpoint reflects $\tilde{\mathbf{w}} = (1, 1, 1)$, i.e., three absolutely symmetric players. The remaining three plurality equivalence classes with reference weights of $\tilde{\mathbf{w}} = (2, 1, 1)$, $\tilde{\mathbf{w}} = (2, 2, 1)$ and $\tilde{\mathbf{w}} = (3, 2, 2)$ correspond respectively to the purple lines between the boundary midpoints, the dark yellow lines from the simplex's center to the three boundary midpoints, and the residual orange triangles. Lists of reference weight distributions for other values of s are provided by Mayer

and Napel (2021).

5 Concluding remarks

The illustrations in Section 4 exhibit the hidden beauty of weighted voting in committees. However, the artistically appealing (at least to us) geometry and changing colors have substantive implications. They reveal structural properties of collective decision making in politics and economics.

Take, for instance, the large blue triangles in the panel for $s = 0$. As pointed out already, they correspond to dictatorship of one player. Namely, if some player's voting weight slightly exceeds 50% – because a shareholder has acquired a small majority stake in a corporation, committee seats are awarded in proportion to population shares in an ethnically polarized society with one majority and several large minorities, etc. – then all plurality decisions correspond to the top-preference of that player. This is not the case for many other scoring rules: consider different levels of s and watch how the blue triangle shrinks from panel to panel. The shrinkage documents how basing decisions on more than just the top preference over all candidates makes collective choice more 'inclusive'. For instance, adopting Borda's rule instead of taking plurality decisions turns a previous dictator with $\bar{w}_i$ slightly above 50% into just a very dominant player. Player i can swing the joint decision for many but no longer for all preference configurations under Borda's rule. Antiplurality awards dictatorial influence not even to a player who has a perfect monopoly of votes. The player's relative weight of 100% makes it impossible for the respective worst-ranked candidate to win but with lexicographic tie-breaking relatively few perturbations of the dominant player's preferences alter the winner.

The changing variety of colors in the panels visualize the findings reported in Figure 1: equivalence maps for scoring rules with $0 < s < 1/2$ or $1/2 < s < 1$ involve many more color shades than those for $s = 0$, $s = 1$, and also $s = 1/2$. We can moreover locate the ranges of weights where most of the color changes are concentrated, i.e., where sensitivity to small weight changes is the greatest. In these weight regions the incentives to, say, increase one's corporate shareholdings or to try to attract a party switcher are much greater than in monochrome areas.

Multiplicity of colors in a panel also indicates the scope for achieving a particular distribution of influence as an institutional designer who can determine the distri-

bution of weights (the so-called 'inverse problem' of voting power; cf. Kurz 2012). Think of a federation of three differently sized states: it may be desirable to make states' voting power a specific function of population sizes – e.g., to achieve direct proportionality or proportionality to the square root of population sizes. Though perfect symmetry (light yellow) or dictatorship by one state (dark blue) are always feasible, the chances of finding voting weights that achieve the targeted distribution of influence is arguably smaller for, say, $s = 1$ with only five equivalence classes than for $s = 1/2$ with 51. There is also a tendency for the distribution of relative voting power to match the underlying distribution of relative voting weights better, the more equivalence classes or colors in our illustrations. This relates to the so-called 'transparency' of a voting rule (cf. Kurz et al. 2021, Section 7.4).

We readily admit that illustrations of voting power in three-player committees that decide between three options have neither the complexity nor the aesthetic qualities of Julia sets or Mandelbrot sets, which have crossed the boundaries between art and science much earlier (see, e.g., Peitgen and Richter 1986). But there is definitely more art and beauty in weighted voting and the resulting voting power than typically meets the eye.

References

Banzhaf, J. F. (1965). Weighted voting doesn't work: a mathematical analysis. *Rutgers Law Review 19*(2), 317–343.

Holler, M. J. and H. Nurmi (Eds.) (2013). *Power, Voting, and Voting Power: 30 Years After*. Heidelberg: Springer.

Holler, M. J. and E. W. Packel (1983). Power, luck and the right index. *Zeitschrift für Nationalökonomie (Journal of Economics) 43*(1), 21–29.

Kovacic, M. and C. Zoli (2021). Ethnic distribution, effective power and conflict. *Social Choice and Welfere 57*, 257–299.

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

Kurz, S., A. Mayer, and S. Napel (2020). Weighted committee games. *European Journal of Operational Research 282*(3), 972–979.

Kurz, S., A. Mayer, and S. Napel (2021). Influence in weighted committees. *European Economic Review 132*, 103634.

Maskin, E. and A. Sen (2016). How majority rule might have stopped Donald Trump. *New York Times*. April 28, 2016.

Mayer, A. and S. Napel (2020). Weighted voting on the IMF Managing Director. *Economics of Governance 21*(3), 237–244.

Mayer, A. and S. Napel (2021). Weighted scoring committees. *Games 12*(4), 94.

Napel, S. (2019). Voting power. In R. Congleton, B. Grofman, and S. Voigt (Eds.), *Oxford Handbook of Public Choice*, Volume 1, Chapter 6, pp. 103–126. Oxford: Oxford University Press.

Napel, S. and D. Welter (2021). Simple voting games and cartel damage proportioning. *Games 12*(74), 1–18.

Peitgen, H.-O. and P. H. Richter (1986). *The Beauty of Fractals*. Berlin: Springer.

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

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.

Taylor, A. D. and W. S. Zwicker (1999). *Simple Games*. Princeton, NJ: Princeton University Press.

Von Neumann, J. and O. Morgenstern (1953). *Theory of Games and Economic Behavior* (3rd ed.). Princeton, NJ: Princeton University Press.