Quarrelling in coalitions to increase p-voting power*

László Á. Kóczy†

Abstract

While they use the language of game theory known measures of a priory voting power are hardly more than statistical expectations assuming the random behaviour of the players. Focusing on normalised indices we show that rational players would behave differently from the indices predictions and propose a model that captures such strategic behaviour.

Keywords and phrases: Banzhaf index, Shapley-Shubik index, a priori voting power, rational players.

---

*The author is grateful for discussions with Manfred Holler and Stefan Napel; he also thanks the funding of the OTKA (Hungarian Fund for Scientific Research) for the project "The Strong the Weak and the Cunning: Power and Strategy in Voting Games" (NF-72610) and of the European Commission under a Marie Curie Reintegration Grant (PERG-GA-2008-230879).

†Keleti Faculty of Economics, Budapest Tech, Tavaszmezõ 15-17. H-1084 Budapest. Email: `koczy.laszlo@kgk.bmf.hu`

1 Introduction

Since Shapley and Shubik (1954) adopted the Shapley value to measure a priori voting power game theory has contributed an enormous literature to this topic: established theoretical underpinnings for the existing or rediscovered indices, introduced new ones, but the plethora of power indices hints that there is no single best. Institutional details that cannot be captured by the voting game largely determine, which index is most suited (Laruelle, 1999). Firstly, however, one must decide the purpose of the study. When we compare different voting situations we want to understand the effectiveness of the same player in making or changing decisions. This is captured by the *influence* power, or I-power of players. In this paper we compare the power of different players in the same voting situation, and hence we are interested in the distribution of power. This type of power, P-power presumes an office-seeking behaviour of voters, so that players share a prize (Felsenthal and Machover, 1998, 2004).

Game theory embraced power indices (P-measures) despite the fact that none of the power indices are truly "game theoretical:" Absent of strategic considerations they are, just as I-measures, mere statistical descriptors of random behaviour. They tell how well a trained ape would do in the voting situation shunning the question how much power a rational and intelligent human *could* acquire. In the following we elaborate this observation.

Power indices measure the probability of being the instrumental voter in making a decision. They are defined over voting situations, modelled by simple games, very much like values are applied to general TU games. There is, however one difference. Values stand for fairness – in contrast to both noncooperative games and domination based concepts, where players with much power to change the outcome of the game are rewarded with a high payoff. It is well known that the payoff vector corresponding to the

Shapley-value can be outside the core and hence the coalitional stability of the Shapley-value is not guaranteed. When a player has the power to increase its power index, that is, its decision making power, then it simply has a higher decision making power, which should be reflected in the index. In voting games coalitional deviations do not make sense. It is well known, however, that a commitment of noncooperation between two voters can yield mutual benefits: this is precisely the paradox of quarrelling members (Brams, 1975/2003). Without such a commitment all winning coalitions will form, but such non-cooperation agreements can increase the voting power of a player. Motivated by this paradox we augment voting games with a previous stage, where players can choose the coalitions they want to join in the voting game. This is a simple noncooperative game where "the acquisition of power is the payoff" (Shapley, 1962, p. 59). We show that all known normalised indices are affected by such strategic behaviour.

Our paper is not the first to disallow certain (winning) coalitions in values or power indices. Aumann and Drèze (1975) assume that property rights may make it impossible to form every coalition. Owen (1977, 1982) assume that coalitions are formed exactly in order to *increase* power. Myerson (1977, 1980) presents a model where players communicate via conferences and not all conferences may occur Faigle and Kern (1992). The application of such restrictions to power indices are more recent (Bilbao, Jiménez, and López, 1998).

The structure of the paper is as follows. We start with a brief introduction to voting games and an overview of the known indices. We briefly explain the paradox of quarrelling members, introduce a framework for strategic indices and prove a number of properties.

2 Power indices

We begin by introducing classical models to measure a priory voting power.

A voting situation or voting game is a pair $(N, \mathcal{W})$, where N is the set of voters and $\mathcal{W}$ denotes the set of *winning coalitions*. We consider games where

1. $\varnothing \notin \mathcal{W}$,

2. if $C \subset D \subset E$ and $C, E \in \mathcal{W}$ then $D \in \mathcal{W}$

3. If $S \in \mathcal{W}$ and $T \in \mathcal{W}$ then $S \cap T \neq \emptyset$.

Condition 3 requires the game to be *proper*, that is, a motion and its opposite cannot be approved simultaneously. Condition 2 is a *convexity* condition on the poset formed by the winning coalitions. It is often assumed that $N \in \mathcal{W}$ and then N replaces E in Condition 2; such games are *simple*.

Let Γ denote the collection of proper convex voting games satisfying the above properties.

Let $\mathcal{M}$ denote the set of *minimal winning coalitions*: the set of coalitions without proper winning subsets. This implies that if $S \in \mathcal{M}$ and $i \in S$, then $S \setminus \{i\} \notin \mathcal{W}$. Clearly $\mathcal{M} \subseteq \mathcal{W}$. *Surplus coalitions* are winning, but non-minimal. Let $w : 2^N \to \{0, 1\}$ be a membership function, such that $w(S) = 1$ if and only if $S \in \mathcal{W}$.

We study power in decision making formulated by such voting games. I-power measures the probability that a player contributes to a decision. P-power is the probability that a player contributes to a decision provided that a decision has been reached.

Given a game Γ a *power measure* $\kappa : \Gamma \longrightarrow \mathbb{R}_+^N$ assigns to each player i a non-negative real number κ_i, its *power*; if $\sum_{i \in N} \kappa_i = 1$ then it is also a *power index*.

In the following we explain some of the well-known indices.

The *Shapley-Shubik index* (Shapley and Shubik, 1954) applies the Shapley value (Shapley, 1953) to simple games: Voters arrive in a random order; if and when a coalition turns winning the full credit is given to the last arriving, the *pivotal* player. A player's power is given as the proportion of orderings where it is pivotal, formally

$$\phi_i = \frac{\sum_{S \subseteq N} (|S| - 1)!(w(S) - w(S \setminus \{i\}))}{n!}$$

While in simple games any order will yield a unique pivotal player, when only Condition 2 is satisfied, there may be none. To see this, start the order with a destructive player, whose membership turns any coalition losing. In such cases, to obtain an index a further normalisation is required.

The *Banzhaf measure* (Penrose, 1946; Banzhaf, 1965) is the probability that a party is *critical* for a coalition, that is, the probability that it can turn winning coalitions into losing ones formally

$$\psi_i = \frac{\eta_i(\mathcal{W})}{2^{n-1}},$$

where $\eta_i(\mathcal{W})$ is the number of coalitions in $\mathcal{W}$ in which i is critical. When normalised to 1, we get the *Banzhaf index* β (Coleman, 1971):

$$\beta_i = \frac{\eta_i(\mathcal{W})}{\sum_{j \in N} \eta_j(\mathcal{W})}.$$

Numerous variants of the (normalised) Banzhaf index exist. In the *Johnston index* γ (Johnston, 1978) the credit a critical player gets is inversely proportional to the number of critical players in the coalition. In effect, coalitions of different sizes have the same contribution to the distribution of power. Deegan and Packel (1978) argue that only those coalitions form where the benefits are least divided (Riker, 1962): the *Deegan-Packel index* ρ only considers minimal winning coalitions. Finally the Holler-Packel or *Public Good Index* h (Holler and Packel, 1983) modifies the Deegan-Packel index: here the benefit of forming a winning coalition is given to each and every player in the coalition. With the normalisation in simple games the index is nothing but a normalised Banzhaf index, where only minimal winning coalitions are taken into account.

Although there is some disagreement on what should a power index be like, the power indices in use are very much alike. They give credit precisely to the critical (or swing) players, and give them all the same disregarding their size. The sole difference lies in weighting winning coalitions differently. We consider a general power indices along these lines.

For coalition $C \in 2^N \backslash \{\varnothing\}$ let a^C denote its weight such that $\sum_{C \in 2^N \backslash \{\varnothing\}} a^C = \sum_{C \in \mathcal{W}} a^C = 1$ and let k^C denote the number of critical players in C. The power index $\kappa(N, \mathcal{W})$ can be rewritten as

$$\kappa_i = \sum_{C \in 2^N \backslash \{\varnothing\}} a^C \mu_i^C, \quad \text{where} \tag{2.1}$$

$$\mu_i^C = \begin{cases} \frac{1}{k^C} & \text{if } i \text{ is critical} \\ \frac{1}{|C|} & \text{if no } i \in C \text{ is critical,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

is the credit player i gets for being in the coalition (therefore $\sum_{i \in C} \mu_i^C = 1$). Observe that $a^C \neq 0$ iff C contains critical players.
For instance for the Banzhaf index $a^C = \frac{k^C}{\sum_{C \in \mathcal{W}} k^C}$.

3 Strategic voting

There may be various power indices, but they all work with an exogenously given set $\mathcal{W}$ of winning coalitions and assume that players are always happy to join winning coalitions. This seems indeed natural – why would players give up part of their power? If for instance two players start to "quarrel" and refuse to cooperate making any coalition they both belong to losing, their power should decrease. Not necessarily. The "Paradox of Quarrelling Members" (Kilgour, 1974; Brams, 1975/2003) occurs when two players mutually benefit from refusing to cooperate with each other.

Paradoxical or not is a matter of interpretation, but players can certainly *acquire* a larger share of power by approving/rejecting coalitions. In this paper we extend voting games to allow for such strategic considerations and define strategic power indices.

Let us stress here that we do not expect quarrelling on large scale or that the lack thereof does not disprove our theory. We do not say that these coalitions do not form, but only that they should not be taken into account.

3.1 Examples

As a motivation we present a number of games based on *weighted voting*. Here N is a collection of n interest groups, or *parties* having $w_1, w_2, \ldots, w_n$ individual representatives ($w_i \in \mathbb{N}_+$). Let $w = \sum_{i=1}^{n} w_i$. We assume that a quota of $w \geq q > w/2$ is *required* to pass a bill. A coalition C of parties is winning if and only if $\sum_{i \in C} w_i \geq q$. Since $w > q$ and $w_i \geq 0$ weighted voting games are simple and proper.

Example 1. The game G_1 consists of four players represented by their weights[1]: $3_1, 3_2, 2_1, 2_2$ and voting has a quota of 6. The set winning coalitions is $\mathcal{W} = \left\{\underline{3_1 3_2}, \underline{3_1 3_2} 2_1, \underline{3_1 3_2} 2_2, \underline{3_1 2_1 2_2}, \underline{3_2 2_1 2_2}, 3_1 3_2 2_1 2_2\right\}$ (with critical players underlined). The vector of Banzhaf indices is $\beta = \left\{\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right\}$.

Notice that in coalition $3_1 3_2 2_1$ player 2_1 is not critical, while the two larger players are. If 2_1 can prevent the formation of this coalition, the latter are critical in less coalitions, so in relative terms (thus: in a power *index*) 2_1 gains.

Given $\mathcal{W}' = \left\{\underline{3_1 3_2}, \underline{3_1 3_2} 2_2, \underline{3_1 2_1 2_2}, \underline{3_2 2_1 2_2}, 3_1 3_2 2_1 2_2\right\}$ the recalculated Banzhaf index is $\beta(\mathcal{W}') = \left\{\frac{3}{10}, \frac{3}{10}, \frac{1}{5}, \frac{1}{5}\right\}$. Player 2_1's rejection increased its rel-

[1]Subscripts are used to distinguish players with identical weights from each other.

ative power. It is therefore *not* in player 2_1's interest to join every winning coalition it is invited to. This finding is not really surprising. In coalition $3_1 3_2 2_1$ player 2_1 assisted players 3_1 and 3_2 in forming a winning coalition, but without getting any credit for it.

Minimal winning coalitions may also be subject to blocks:

Example 2. G_2 is a 9-player game with players $5_1, 5_2, 5_3, 1_1, 1_2, 1_3, 1_4, 1_5, 1_6$ and a quota of 11. Here $\mathcal{M} = \{5_1 5_2 5_3, 5_i 5_j 1_k, 5_i 1_1 1_2 1_3 1_4 1_5 1_6\}$, where $k \in \{1, 2, 3, 4, 5, 6\}$ and $i, j \in \{1, 2, 3\}$ with $i \neq j$. Let $\mathcal{W} = \mathcal{M}$. Then the Banzhaf index is given by $\beta = \left\{\frac{7}{39}, \frac{7}{39}, \frac{7}{39}, \frac{1}{13}, \frac{1}{13}, \frac{1}{13}, \frac{1}{13}, \frac{1}{13}, \frac{1}{13}\right\}$.
Now consider $\mathcal{W}' = \{5_1 5_2 5_3, 5_i 5_j 1_k, 5_l 1_1 1_2 1_3 1_4 1_5 1_6\}$, where $k \in \{1, 2, 3, 4, 5, 6\}$, $i, j \in \{1, 2, 3\}$ and $l \in \{2, 3\}$. Then $\beta(\mathcal{W}') = \left\{\frac{13}{71}, \frac{14}{71}, \frac{14}{71}, \frac{5}{71}, \frac{5}{71}, \frac{5}{71}, \frac{5}{71}, \frac{5}{71}, \frac{5}{71}\right\}$. The set $\mathcal{W}'$ does not contain the minimal winning coalition $5_1 1_1 1_2 1_3 1_4 1_5 1_6$, yet the critical player 5_1 is better off as $\frac{13}{71} > \frac{7}{39}$.

While the aforementioned indices claim to measure power, it seems, players have actually little power to influence their power: hence they are no more than probabilistic values. The paradox of quarrelling members and our examples illustrate that some players *can* increase their power by refusing to participate in certain coalitions. If a player credibly refuses to participate in a coalition neither him nor his colleagues should get credit for being critical to a coalition that never forms.

3.2 Model

The idea of quarrelling is generalised to coalitions: a coalition Q quarrels if any of its members wants to quarrel. Player i's strategy s_i therefore corresponds to quarrelling in certain coalitions that i belongs to, thus $s_i \subseteq \{C|\ i \in C\}$ and its strategy space $S_i \subset 2^{\{C|\ i \in C\}}$. Note that due to our convexity condition $C \in s_i$ and $C \subseteq D$ imply $D \in s_i$ therefore not all combinations of quarrelled coalitions are possible.

Only winning coalitions without quarrelling remain winning. Given $s = \{s_i\}_{i\in N}$ the *strategy profile*, they are collected by

$$\mathcal{W}(s) = \{w \in \mathcal{W} |\ w \notin s_i,\ \forall i \in N\} = \left\{ w \in \mathcal{W} \setminus \bigcup_{i\in N} s_i \right\}. \qquad (3.1)$$

Observe that $(N, \mathcal{W}(s))$ is a voting game, thus each strategy profile s determines a voting game. In this game Conditions 1 and 3 clearly hold since no new winning coalitions have been added. On the other hand as the addition of new members to a quarrelling coalition does not make it winning, convexity, that is: Condition 2 holds, too.

Definition 1 (Strategic voting game)**.** The quadruple $(N, S, \mathcal{W}, \kappa)$ consisting of a set of players N, a strategy space $S \subseteq \left\{2^{\{C|i\in C\in\mathcal{W}\}}\right\}_{i\in N}$, a collection of initial winning coalitions $\mathcal{W}$ and a power index κ is called a *strategic voting game.*

As the objective of this game is to maximise power, κ is nothing, but the utility function $\kappa : S \longrightarrow \mathbb{R}^N_+, s \longmapsto \kappa(N, \mathcal{W}(s))$. Strategies are in fact sets of coalitions; in the basic case where all strategies are permitted (no strict inclusion in the definition) the strategy space can be derived from the player set, therefore the triple $(N, \mathcal{W}, \kappa)$ fully defines the game.

The game consists of two stages: a first, noncooperative game of quarrelling and a second, implicit, cooperative game of power allocation. Quarrelling is for good despite incentives to make peace ex-post, which implies that only asymmetric deviations are possible. A strategy profile is a Nash equilibrium if s is a best response to itself, that is, $\kappa(s) \geq \kappa(s'_i, s_{-i})$ for all $s'_i \in S_i$ such that $s'_i \supseteq s_i$. Then:

Definition 2. A *strategic power index* is a vector of equilibrium payoffs, that is $\kappa(s^*) = \kappa(N, \mathcal{W}(s^*))$, where s^* is a Nash equilibrium of the strategic voting game $(N, \mathcal{W}, \kappa)$: for all $i \in N$ and all $s_i \subseteq s^*_i$, $s_i \in S_i$ we have $\kappa_i(s^*) \geq \kappa_i(s_i, s^*_{-i})$.

Such a strategic power index always exists ($\mathcal{W}(s^*) = \varnothing$ is an equilibrium) but is generally not unique. In the sequel we provide a unique refinement for certain indices.

4 Results

4.1 Only minimal winning coalitions

Blocking a coalition B affects a player in two ways. On the one hand for all $C \supseteq B$ the coalition's weight (Recall the definition in Section 2.) becomes $(a^C)' = 0$ and hence the player loses $\sum_{C \supseteq B} a^C \mu_i^C$, on the other hand, due to the normalisation the weight of other coalitions increases, and hence the credit it gets from other coalitions is scaled up by

$$\frac{\sum_{C \in 2^N \setminus \varnothing} a^C}{\sum_{C \in 2^N \setminus \varnothing} a^C - \sum_{C \supseteq B} a^C}. \tag{4.1}$$

Null players, that is, players, who never contribute to a coalition, are unaffected and are therefore ignored in our analysis.

Proposition 3. *Surplus coalitions containing critical players are blocked.*

Proof. Consider a coalition B containing a surplus player i. If i is not critical in B, it is also not critical in $C \supset B$ (as, by monotonicity if $B \setminus \{i\}$ is winning, so is $C \setminus \{i\} \supset B \setminus \{i\}$) and therefore $a^C \mu_i^C = 0$ for all $C \supseteq B$. In sum, neither B nor $C \supset B$ yields any profit for i.

On the other hand $a^B > 0$ (and possibly $a^C > 0$ for some $C \supset B$), so when blocking B the power of player i is scaled up according to Expression 4.1 making the block profitable. □

Corollary 4. *For power* indices *we have $\mathcal{M} \supseteq \mathcal{W}^*$.*

However, not all minimal winning coalitions are quarrel-free (see Example 2).

In the following we allow $a^C > 0$ only if $C \in \mathcal{M}$. Holler and Packel (1983, p. 24.) argue that "since a non-critical member ... has no incentive to vote ... only these coalitions should be considered for measuring a priori voting power." Thus a player cannot count on the formation of coalitions that are not due to his or her power. A similar prediction is made by *aspiration* solution concepts (Bennett, 1983, p. 15.).

4.2 Elementary blocks

Definition 5. Given a strategy profile s we say that the deviation s'_i is *elementary* if $|s'_i| - |s_i| = 1$, that is, if s'_i introduces quarrelling in a single coalition of player $i \in N$.

Proposition 6. *Let s'_i be i's best response to s_{-i}. Then s'_i can be reproduced by a sequence of elementary deviations.*

Proof. Proof by construction. Consider the best response s'_i and let $s_i \setminus s'_i = \{C_1, \ldots, C_k\}$ where, without loss of generality, $\mu_i^{C_1} \geq \cdots \geq \mu_i^{C_k}$.

We show that $\kappa_i(s') \geq \mu_1$. Consider the deviation $s''_i = s'_i \setminus C_1$. By assumption

$$\kappa_i(s'_i, s_{-i}) \geq \kappa_i(s''_i, s_{-i}) \tag{4.2}$$

$$\kappa_i(s'_i, s_{-i}) \geq \frac{\left(\sum_{C \neq s'_i} a^C\right) \kappa_i(s'_i, s_{-i}) + a^{C_1}\mu_i^{C_1}}{\sum_{C \neq s'_i} a^C + a^{C_1}} \tag{4.3}$$

The right hand side is a weighted average of $\kappa_i(s'_i, s_{-i})$ and $\mu_i^{C_1}$, and hence $\kappa_i(s'_i, s_{-i}) \geq \mu_i^{C_1}$. □

In the following we will only consider elementary deviations.

4.3 Friendly equilibrium selection

The strategy profile, where $\mathcal{W}^* = \varnothing$ is clearly a Nash-equilibrium, while this is neither the only one nor the one we want (for one, power indices are undefined here); out of the many Nash equilibria we make a selection.

We now move on to define our refinement.

The literature of power indices has been built on the assumption that all winning coalitions form. We agree that it is a reasonable starting point to assume that unless for good reasons, players will be friendly and not block coalitions. Therefore also in our model we will take this as the status quo; when strategic considerations do not play a major role, the equilibrium remains $s = \varnothing$ and a coalition is only blocked if this increases a player's power. This last observation translates to the fact that for any acceptable equilibrium, there is a sequence of elementary deviations where each of the blocks are introduced leading right back to the status quo. The *friendly set* F, defined below collects the acceptable strategy profiles.

$$s \in F \text{ if } \begin{cases} s_i = \varnothing \ \forall i \in N \\ \exists i \in N, \exists (s_i', s_{-i}) \in F, \text{ such that } \kappa_i(s) > \kappa_i(s_i', s_{-i}). \end{cases}$$

We select *friendly equilibria* $s^* \in F$ that are Nash equilibria and are maximal for inclusion. The *equilibrium set of winning coalitions* is $\mathcal{W}^* = \mathcal{W}(s^*)$ and the *strategic κ power index* is defined as

$$\kappa^* = \kappa(s^*) = \kappa(N, \mathcal{W}^*).$$

Now observe that for minimal winning coalitions $C \neq D$ we have neither $C \subset D$ nor $D \subset C$, therefore by blocking C a player will not block D and vice versa, a player has the possibility to block each minimal winning coalition separately. In sum, our model can be reduced to players picking which coalitions they do not want to form. This result makes it particularly easy to work with coalitions rather than strategies. Then an equilibrium is simply $\mathcal{W}^*$ instead of $\mathcal{W}(s^*)$ and let $\mathcal{F} = \{\mathcal{W}(s) | s \in F\}$.

Player i profitably blocks coalition B iff

$$\kappa_i(N, \mathcal{W} \setminus \{B\}) \quad > \quad \kappa_i(N, \mathcal{W}) \tag{4.4}$$

$$\frac{\sum_{C \in \mathcal{W}} a^C}{\sum_{C \in \mathcal{W}} a^C - a^B} \left(\sum_{C \in \mathcal{W}} a^C \mu_i^C - a^B \mu_i^B \right) \quad > \quad \sum_{C \in \mathcal{W}} a^C \mu_i^C \tag{4.5}$$

After some rearrangements we get

$$\frac{\sum_{C \in \mathcal{W}} a^C \mu_i^C}{\sum_{C \in \mathcal{W}} a^C} = \kappa_i(N, \mathcal{W}) > \mu_i^B, \tag{4.6}$$

which gives the following result.

Lemma 7. *A block by player i is profitable if and only if the blocked coalition gives less credit to player i than the average credit it gets, that is, than its power index.*

Proposition 3 can also be seen as a corollary of this lemma.

Lemma 7 also suggests a relation to the theory of *aspirations* (Bennett, 1983), although this relation turns out to be superficial. In the theory of aspirations it is not some coalition's payoff that is bargained over: it is the players that make their claims and unless their claims are satisfied certain coalitions will or will not form. Here this claim is expressed by their power index, the "credit they receive in general" and players demand the same credit in coalitions. Unfortunately the link between the two concepts does not go much beyond that. While a power index satisfies $\sum_{i \in N} \kappa_i^* = 1$ a vector of aspirations will almost always be larger. Bennett (1983, p. 15.) provides the following example:

Example 3. A game with 5 players with weights 2, 2, 1, 1, and 1, and a quota of 5. Here the unique partnered, balanced, equal gains aspiration is $(0.4, 0.4, 0.2, 0.2, 0.2)$, while the public good index is $h = (\frac{4}{17}, \frac{4}{17}, \frac{3}{17}, \frac{3}{17}, \frac{3}{17})$.

Now we move on to our main result.

Theorem 8. *The friendly equilibrium set of winning coalitions is uniquely defined and is given by*

$$\mathcal{W}^* = \bigcap_{s \in F} \mathcal{W}(s). \tag{4.7}$$

In order to prove this theorem we need some additional results.

Proposition 9. *Let $C_i, C_j \in \mathcal{W}$ be coalitions that both contain both i and j and such that i and j want to block C_i and C_j respectively. Then either i wants to block C_j or j wants to block C_i.*

Proof. Assume that the proposition is false. This means the following. Player j blocks C_j, hence $\mu_j^{C_j} < \kappa_j(\mathcal{W})$. By our assumption i does not block, hence $\mu_i^{C_j} \geq \kappa_i(\mathcal{W})$. Therefore $\mu_j^{C_j} < \mu_i^{C_j}$. Similarly i blocks C_i, hence $\mu_i^{C_i} < \kappa_i(\mathcal{W})$. By our assumption j does not block, hence $\mu_j^{C_i} \geq \kappa_j(\mathcal{W})$. In sum $\mu_i^{C_i} < \mu_i^{C_j}$ and $\mu_j^{C_i} < \mu_j^{C_j}$. Since C_i and C_j are minimal winning coalitions $\mu_i^{C_i} = \mu_j^{C_i} = \frac{1}{|C_i|}$ and $\mu_j^{C_j} = \mu_i^{C_j} = \frac{1}{|C_j|}$. Contradiction □

Proposition 10. *For all $\mathcal{W}_i, \mathcal{W}_j \in \mathcal{F}$ we have $\mathcal{W}_i \cap \mathcal{W}_j \in \mathcal{F}$.*

Proof. The proof is by induction on the differences between $\mathcal{W}_i$ and $\mathcal{W}_j$.

First we deal with the elementary step. Assume $\mathcal{W}_i = \{A, C_1, C_2, \ldots C_m\}$, $\mathcal{W}_j = \{B, C_1, C_2, \ldots C_m\}$, that is, the two sets only differ in 1 element each. This ensures that their intersection is non-trivial. $\mathcal{W}_i$ and $\mathcal{W}_j$ are descendants of a common ancestor $\mathcal{W}_0 = \{A, B, C_1, C_2, \ldots C_m\}$, but after blocking B and A, respectively by some players i and j. The proposition merely states that either blocking A is profitable from $\mathcal{W}_i$ or blocking B is profitable from $\mathcal{W}_j$.

$\mathcal{W}_i$ is the result of blocking B by i. If $j \notin B$ then $\kappa_j(\mathcal{W}_0) \leq \kappa_j(\mathcal{W}_i)$. We know that j blocks A at $\mathcal{W}_0$ and hence $\kappa_j(\mathcal{W}_0) > \mu_j^A$. Hence $\kappa_j(\mathcal{W}_i) > \mu_j^A$, which implies that j also blocks A at $\mathcal{W}_i$. Thus $\mathcal{W}_{ij} = \{C_1, C_2, \ldots C_m\} \in \mathcal{F}$.

The symmetric case gives the corresponding result for i and B at $\mathcal{W}_j$.

Finally we must consider the case where none of the previous two cases applied, that is where $j \in B$ and $i \in A$. As only a member can block a coalition, we also have $j \in A$ and $i \in B$. Therefore we can apply Proposition 9 to show that i blocks at $\mathcal{W}_j$ or j at $\mathcal{W}_i$, which, as before, gives the result.

We have discussed all possible cases which completes the first part of the proof. Now we move on to the general case. Assume that we have shown the result for all pairs of sets with differences up to $k-1$.

Now consider $\mathcal{W}_i = \{A_1, A_2, \ldots, A_k, C_1, C_2, \ldots C_m\}$ as well as $\mathcal{W}_j = \{B_1, B_2, \ldots, B_l, C_1, C_2, \ldots C_m\}$, where $A_1, A_2, \ldots, A_k$ and $B_1, B_2, \ldots, B_l$ represent the blocks that did *not* take place and $l \leq k$. (Possibly $A_p = B_q$ for some p and q.) The question is whether this difference can be eliminated.

By definition if $\mathcal{W}_i \in \mathcal{F}$ there exists a sequence of blocks starting from $\mathcal{W}_0$ that lead to $\mathcal{W}_i$ and a similar sequence exists to $\mathcal{W}_j$. Let $\mathcal{W}_i^0$ and $\mathcal{W}_j^0$ be the first elements that are not common, without loss of generality, as results of blocking B_1 and A_1 respectively. By the elementary step $\mathcal{W}_j^1 = \mathcal{W}_i^0 \cap \mathcal{W}_j^0$ belongs to $\mathcal{F}$.[2] Now take the next set $\mathcal{W}_2$ along the path to $\mathcal{W}_i$, $\mathcal{W}_i^1$. By the same argument $\mathcal{W}_i^1 \cap \mathcal{W}_j^1$ also belongs to $\mathcal{F}$. Repeating this argument we travel parallel to the path and in the penultimate step we get $\mathcal{W}_j^p \in \mathcal{F}$. For the last time by the same argument $\mathcal{W}_i \cap \mathcal{W}_j^p = \{A_2, \ldots, A_k, C_1, C_2, \ldots C_m\}$ also belongs to $\mathcal{F}$. If $l < k$, our inductive assumption can be used to complete the proof.

In case $l = k$ it is necessary to apply the same argument once more, but on the other side: to show that $\{B_2, \ldots, B_l, C_1, C_2, \ldots C_m\} \in \mathcal{F}$. □

Proof of Theorem. By Proposition 10 pairwise intersections of elements of $\mathcal{F}$ also belong to $\mathcal{F}$. As the number of winning coalitions is finite the result on pairwise intersections implies that $\mathcal{W}^*$ as defined in Equation 4.7 belongs to $\mathcal{F}$. Clearly $\mathcal{W}^* \subseteq w$ for all $w \in \mathcal{F}$. Therefore $\mathcal{W}^*$ is the smallest friendly set and it is trivially an equilibrium. □

Corollary 11. *The strategic power index κ^* is well-defined.*

[2]Our notation is slightly misleading as $\mathcal{W}_j^1$ is not necessarily on the path to $\mathcal{W}_j$, but this should not lead to confusion.

5 Conclusion

We have developed a model that measures power taking the rational, utility maximising behaviour of players into account. We have also shown that none of the well-known power indices account for this behaviour. It appears that these supposedly game theoretic concepts are not more than statistical measures of random behaviour.

There are at least two possibilities to resolve this conflict. The one we chose is to modify existing power indices so that no credit is given for coalitions that do not form. The advantage of this solution is that it is directly motivated by the problem and gives a perfect answer to it without affecting the concepts a great deal.

While this is the option we choose here there is an interesting alternative. Observe that blocking a winning coalition may be advantageous to some players, but it will hurt others in the coalition. The only players whose power will surely increase are those *outside* the coalition. This indicates that overall members of the coalition loose by not forming the coalition. Hence forming the coalition increases the power of the members and therefore there exists distributions of this power that benefit all members. Giving room for renegotiation would lead us to cooperative, probably set-like solutions and would make us lose the advantages of a single-point solution concept.

Two other choices we have made are to assume that blocking coalition C also blocks $D \supset C$ and to work with power indices defined over minimal winning coalitions only. Blocking single coalitions would not preserve null players who could gain power for "mediation" (turning a blocked coalition into a winning one by their entry – of course this coalition would be blocked soon, too) and would allow non-minimal winning coalitions that are not surplus coalitions as they would only consist of critical players. While our original model considered a variant of this alternative, in order to avoid such odd phenomena one has to separate the notions of winning a feasible coalition.

Finally, the uniqueness of the friendly equilibrium for power indices also looking at surplus coalitions remains an open problem. With the aforementioned model counterexamples can be presented here a systematic search for them was in vain, now we believe the result to hold, but the present proof does not directly extend to those indices.

References

AUMANN, R. J., AND J. H. DRÈZE (1975): "Cooperative games with Coalition Structures," *International Journal of Game Theory*, 3, 217–237.

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

BENNETT, E. (1983): "The aspiration approach to predicting coalition formation and payoff distribution in sidepayment games," *International Journal of Game Theory*, 12(1), 1–28.

BILBAO, J. M., A. JIMÉNEZ, AND J. J. LÓPEZ (1998): "The Banzhaf power index on convex geometries," *Mathematical Social Sciences*, 36, 157–173.

BRAMS, S. J. (1975/2003): *Game theory and politics*. Dover, Mineola, N.Y., 2nd edn.

COLEMAN, J. S. (1971): "Control of Collectives and the Power of a Collectivity to Act," in *Social Choice*, ed. by B. Lieberman, pp. 192–225. Gordon and Breach, New York.

DEEGAN, J., AND E. W. PACKEL (1978): "A New Index of Power for Simple n-Person Games," *International Journal of Game Theory*, 7(2), 113–123.

Faigle, U., and W. Kern (1992): “The Shapley Value for Cooperative Games Under Precendence Constraints,” *International Journal of Game Theory*, 21(3), 249–266.

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

——— (2004): “A priori voting Power: What is it all about?,” *Political Studies Review*, 2(1), 1–23.

Holler, M. J., and E. W. Packel (1983): “Power, Luck and the Right Index,” *Journal of Economics (Zeitschrift für Nationalökonomie)*, 43(1), 21–29.

Johnston, R. J. (1978): “On the Measurement of Power: Some Reactions to Laver,” *Environment and Planning A*, 10(8), 907–914.

Kilgour, D. M. (1974): “A Shapley Value for Cooperative Games with Quarrelling,” in *Game Theory as a Theory of Conflict Resolution*, ed. by A. Rapoport, Theory & Decision Library, pp. 193–206. D. Reidel Publishing Co., Dordrecht.

Laruelle, A. (1999): “On the Choice of a Power Index,” Discussion Paper WP-AD 99-10, Instituto Valenciano de Investigaciones Económicas, Valencia.

Myerson, R. B. (1977): “Graphs and cooperation in games,” *Mathematics of Operations Research*, 2, 225–229.

——— (1980): “Conference Structures and Fair Allocation Rules,” *International Journal of Game Theory*, 9(3), 169–182.

Owen, G. (1977): "Values of games with a priori unions," in *Mathematical economics and game theory: Essays in honor of Oskar Morgenstern*, ed. by R. Henn, and O. Moeschlin, vol. 141 of *Lecture Notes in Economics and Mathematical Systems*, pp. 76–88. Springer, Berlin.

——— (1982): "Modification of the Banzhaf-Coleman Index for Games with a Priori Unions," in *Power, Voting and Voting Power*, ed. by M. J. Holler, pp. 232–238. Physica-Verlag, Würzburg–Wien.

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

Riker, W. H. (1962): *The Theory of Political Coalitions*. Yale University Press, New Haven.

Shapley, L. S. (1953): "A value for n-person games," in *Contributions to the Theory of Games II*, ed. by H. W. Kuhn, and A. W. Tucker, vol. 28 of *Annals of Mathematics Studies*, pp. 307–317. Princeton University Press, Princeton.

——— (1962): "Simple Games: An outline of the descriptive theory," *Behavioral Science*, 7, 59–66.

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.