Worst-case Bounds on Power vs. Proportion in Weighted Voting Games with an Application to False-name Manipulation

Yotam Gafni[1] , **Ron Lavi**[1,2] , **Moshe Tennenholtz**[1]

[1] Technion - Israel Institute of Technology
[2] University of Bath, UK
{yotam.gafni@campus, ronlavi@ie, moshet@ie}.technion.ac.il

Abstract

Weighted voting games are applicable to a wide variety of multi-agent settings. They enable the formalization of power indices which quantify the coalitional power of players. We take a novel approach to the study of the power of big vs. small players in these games. We model small (big) players as having single (multiple) votes. The aggregate relative power of big players is measured w.r.t. their votes proportion. For this ratio, we show small constant worst-case bounds for the Shapley-Shubik and the Deegan-Packel indices. In sharp contrast, this ratio is unbounded for the Banzhaf index. As an application, we define a false-name strategic normal form game where each big player may split its votes between false identities, and study its various properties. Together our results provide foundations for the implications of players' size, modeled as their ability to split, on their relative power.

1 Introduction

Weighted Voting Games (WVGs) are a class of cooperative games that naturally appear in diverse settings, such as parliaments, councils and firm shareholders. In recent years, they were found to naturally appear in multi-agent systems such as VCG auctions [Bachrach *et al.*, 2011] and other online economic systems. WVGs are defined by a set of players, their weights, and a threshold T. A set of players forming a coalition must have an aggregate weight of at least T. It is natural to ask: What is a player's power to influence decisions, or, alternatively, what is a player's share of the benefit of forming a coalition? This power measure does not necessarily comply with the player's proportional weight. For example, consider a WVG with a large threshold T, a big player with weight $T-1$ and a small player with weight 1. Despite the large discrepancy in their weights, a consensus is required for any motion to pass, suggesting they have equal power. This view of power considers a player's pivotal role as "king-maker"– "To the victors go the spoils".

Due to this reason, cooperative game theory studies power indices to capture the true effective power of players in WVGs. This literature views the power index of a player as a numeric predictor of utility. The most prominent power indices include the Shapley-Shubik index [Shapley and Shubik, 1954] which stems from the more general Shapley value [Shapley, 1952], and the Banzhaf index [Banzhaf III, 1964]. Other power indices emphasize different aspects of the power structure, such as the Deegan-Packel index [Deegan and Packel, 1978], which we also study. Our work lies in the intersection of three strands of WVG literature:

1.1 Big vs. Small Players and Group Power

We lay out a WVG model with big vs. small players and study the power of big players compared to their vote proportion, assuming that all player weights are natural numbers and consider all players with weight larger than 1 as "big" and all players with weight 1 "small". The inequality in voting – "big vs. small" – is a main drive for the study of power indices, going back to the formation of the US electoral college. Riker [1986] points out that Luther Martin of Maryland, a staunch anti-federalist and one of the US founding fathers, analyzed the then forming electoral college in a manner similar to the Banzhaf index. In the compilation by Storing [2008], p. 50, Martin claims:

> The number of delegates ought not to be in exact proportion to the number of inhabitants, because the influence and power of those states whose delegates are numerous, will be greater [even relative to their proportion] when compared to the influence and power of the other states...

The contrast between big vs. small players exists not only in traditional voting settings but also in modern contexts. For example, we see multiple situations in which several small websites (where here "small" is in terms of their number of users) aggregate their market power by forming a unified service platform to compete with a big incumbent website. Similarly we see aggregation of computational power (e.g., mining pools in Bitcoin, consortia of cloud computing services). The other direction of a big player splitting itself to multiple small identities also exists, even when such a split is costly in terms of advertising and maintaining the brand, e.g., flight search engines and web hosting services.

Shapley and Shubik [1954] demonstrate that in the settings of one big player and many small players, the power of the big player can be higher than its proportional weight. They do not answer (nor ask) the question of how large this ratio can be. They also do not analyze the opposite direction, of whether this ratio is bounded below by some constant, possibly smaller than 1. As we explain below, our results generalize and answer these questions. Beyond the case of one big player and many small players, which we completely characterize, our model and results extend in two aspects:

- *Arbitrary number of big (and small) players* - We obtain results regarding the power vs. proportion of any specific one big player and regarding the *aggregate* power of big players. As the distinction to big and small immediately suggests, the relative power of these groups now becomes the focus. Examples for such groups are the top 1% wealthy people, the G7 countries, Bitcoin's large miners. These settings typically involve a few big players and many small players. Milnor and Shapley [1978] and Neyman [1981] take this to an extreme by considering so-called "oceanic games" where there are a few significant large players and a continuum of small players. In contrast, our results are not asymptotic and hold for any arbitrary number of big and small players.
- *The Banzhaf and Deegan-Packel indices* - Since different power indices structuring naturally encapsulates different aspects of strategic power, it is important to compare the results of different indices given our model, and more so as the Banzhaf index gives qualitatively different results than the other indices.

1.2 Power vs. Proportion

Our main theoretical results, given in Sections 3-5, characterize the ratio between the aggregate power of the big players and their aggregate proportional weight for different power indices. Most previous literature analyzes ways to adjust voters' voting weights in order to equate voting power to actual weight. An early suggestion by [Penrose, 1946] is that, in the UN, states should be assigned seats proportional to the square root of their population, in order to achieve proportional representation for each citizen, worldwide, regardless of her state. [Słomczyński and Zyczkowski, 2006; Słomczyński and Życzkowski, 2007] further suggest an improvement in the form of the *double square root* voting system, where on top of assigning seats proportional to the square root, the voting threshold (quota) of the representative body itself is determined so to optimize proportionality.

More recently, attention was drawn to whether a good choice of voting threshold (quota) can attain proportionality [Zick *et al.*, 2011; Zick, 2013; Oren *et al.*, 2014; Bachrach *et al.*, 2016b; Bachrach *et al.*, 2016a]. Theoretical guarantees, experimental results, and probabilistic models were suggested, for which this occurs. For example, these works collectively establish that under some probabilistic assumptions, setting the threshold T to be about 50% of the total sum of weights results in power being equal to proportion with high probability. In contrast, our worst-case analysis of this problem does not depend on probabilistic assumptions that might not hold in reality, due to the independence assumptions or specific properties of the distributions. We also show examples where the threshold is very close to 50% and the power is far from proportional. In addition, it may not be possible to tune the threshold T because of exogenous dictates (e.g., important parliament votes, where a two-thirds majority is required) or because the model aims to capture some underlying reality (e.g., over the Internet) that constrains T.

A third approach focuses on probabilistic modelling of the WVG weights. For example, Jelnov and Tauman [2014] show that if player weights are sampled uniformly from the unit simplex, the expected Shapley-Shubik power of a player relative to its proportion goes to 1 with rapid convergence in the number of players. Lindner and Machover [2004] study a different model where the ratio of the Shapley-Shubik index to proportional weight in infinite chains of game instances asymptotically approaches 1. Chang *et al.* [2006] follow up with an experimental analysis of a similar model thus further verifying the previous conceptual conclusions.

1.3 False-name Manipulation in WVG

As an application of our main results for power vs. proportion, we define in section 6 a false-name strategic normal-form game where each big player may split its votes between false identities. Aziz *et al.* [2011] are the first to study power indices in the context of false-name manipulation, showing upper and lower bounds on a player's gain (or loss) from splitting its votes into two parts, for the Shapley-Shubik and Banzhaf indices. They also address a range of computational issues, among them the decision problem of splitting into two equal parts which is NP-hard for both indices. Faliszewski and Hemaspaandra [2009] show that the decision problem of benefiting from splitting into two equal parts to be in PP, and Rey and Rothe [2014] show it is PP-complete for the Shapley-Shubik index, PP-complete for the Banzhaf index with three equal splits, and PP-hard for both indices with general splits.

Our results contribute to the above literature on false-name splits by unifying it with the two previously mentioned aspects ('big vs. small' and 'power vs. proportion') and by generalizing on two additional fronts:

- *General splits* - We consider splits into multiple identities, rather than splits into two or three identities. Lasisi and Allan [2017] initiated work on this more general problem, where they show some upper and lower bounds for the individual power gain from general splits compared to the original power. These bounds assume that only a single agent splits, where as the bounds we provide hold under any combination of strategic manipulations by the agents.
- *Global bounds on manipulation* - By our results for the power vs. proportion we extract useful global bounds on manipulation. Regularly power is compared before and after splits. Since previous work shows the most basic questions in regard to successful power manipulation to be computationally hard, developing global performance bounds is important.

2 Preliminaries

Weighted Voting Games (WVGs), starting with weighted majority games [Morgenstern and Von Neumann, 1953; Shapley, 1962], aim to capture a situation where several players need to form a coalition. Each player has a weight, and a subset of players can form a coalition if their sum of weights passes a certain threshold. In this paper, we make a distinction between "big" versus "small" players, where small players have a weight of one. Formally,

Definition 1. *(Adapted from Shapley [1962]) A weighted voting game is a tuple $\{A, m, T\}$ with*

$$A = \{a_1, ..., a_r\}, M = \overbrace{\{1, ..., 1\}}^{m}, \quad 1 \leq T \leq m + \sum_{j=1}^{r} a_j,$$

where $a_1, ..., a_r, m, T \in \mathbb{N}$, there are r "big players", m "small players" of weight 1, and a coalition threshold T.[1] *We at times denote the small players as $1_1, ...1_m$. Note that A is a multiset. When we write $A \setminus \{i\}$, for some weight i, at most one occurrence of i is removed from A.*

A basic question in WVGs is how to split the gains from forming a coalition among its members. One possible notion of fairness is to split gains in a way that is approximately proportional to the weights of the coalition members.

Definition 2. *The proportional value of a weighted voting game is* $P(A, m, T) = \frac{\sum_{j=1}^{r} a_j}{m + \sum_{j=1}^{r} a_j}$.

However reality tells us that many times the "power" of players is different than their proportional weight. A well-established literature on power indices formally studies this by looking at our setting as a cooperative game. For the analysis we have the following value function $v(S)$ which describes whether a subset of players $S \subseteq A \cup M$ is able to form a coalition:

$$v(S) = \begin{cases} 1 & \sum_{s \in S} s \geq T \\ 0 & otherwise. \end{cases}$$

We next compare the aggregate power of the big players, using several three well-known power indices, to their aggregate proportional weight.

3 The Shapley-Shubik Power Index

Define the ordered tuple $A + M = (a_1, ..., a_r, 1_1, ..., 1_m)$. Let S_{m+r} be the group of all permutations operating on $m + r$ objects. For some $\sigma \in S_{m+r}$, $\sigma(A + M)$ is the ordered tuple which results by applying the permutation σ to $A + M$. We usually omit the term $A + M$ when it is clear from the context. Define $\sigma|_p, \bar{\sigma}|_p$ as the set of all players (strictly, non-strictly) preceding player p in permutation $\sigma(A + M)$. The permutation pivotal player indicator function for a player p is

$$\mathbb{1}_{p,\sigma} = v(\bar{\sigma}|_p) - v(\sigma|_p).$$

In words, the indicator $\mathbb{1}_{p,\sigma}$ is equal to one if the players preceding p in the permutation $\sigma(A+M)$ do not form a coalition and adding p enables the coalition formation. In such a case, we say that p is pivotal for σ. Note that $\mathbb{1}_{p,\sigma} \in \{0, 1\}$ and that each permutation has exactly one pivotal player.

Definition 3. *(Adapted from [Winter, 2002]) The Shapley-Shubik power index of a weighted voting game (A,m,T) is*

$$\phi_p(A, m, T) = \mathbb{E}_{\sigma \sim UNI(S_{m+r})}[\mathbb{1}_{p,\sigma}],$$

for a player p (whether a big player a_i or a small player 1_i), where UNI is the uniform distribution over a discrete set.

This definition is a special case of the Shapley value applied to WVGs. Three well-known properties of this power index are symmetry, efficiency, and non-negativity, which we utilize in our proofs.

Definition 4. *The Shapley-proportional ratios are the global supremum (infiumum) over all weighted voting games*

$$\bar{R}_\phi = \sup_{A,m,T} \frac{\sum_{j=1}^{r} \phi_{a_j}(A, m, T)}{P(A, m, T)}, \underline{R}_\phi = \inf_{A,m,T} \frac{\sum_{j=1}^{r} \phi_{a_j}(A, m, T)}{P(A, m, T)}.$$

Note that $\underline{R}_\phi \geq 0$ because of non-negativity.

Example 1 ($\bar{R}_\phi$ is at least 2)**.** *For some $k \geq 2$, consider $A = \{k\}, m = k - 1, T = k$. Then $\phi_{a_1}(A, m, T) = 1$, while $P(A, m, T) = \frac{1}{2} + \frac{1}{4k-2}$.*

In fact, this asymptotic lower bound is tight:

Theorem 1. $\bar{R}_\phi = 2$.

To prove Theorem 1, we first show in lemma 1 a recursive relation for the Shapley-Shubik index. We are then able to give an inductive proof of Theorem 1, to appear in the full version of this paper.

Lemma 1. *The following recursion holds for ϕ_1*

$$\phi_1(A, m, T) = \begin{cases} \frac{1}{m+r} & T = 1 \\ \frac{1}{m+r}\Bigg(\sum_{\substack{1 \leq i \leq r \\ a_i < T}} \phi_1(A \setminus \{a_i\}, m, T - a_i) + \\ \qquad (m-1)\phi_1(A, m-1, T-1)\Bigg) & T > 1 \end{cases}$$

The following example shows that $\underline{R}_\phi = 0$.

Example 2. *For any $k \geq 2$, choose $A = \{k\}, m = k, T = 2k$. Then $\phi_{a_1}(A, m, T) = \frac{1}{k+1}$ while $P(A, m, T) = \frac{1}{2}$.*

In example 2, a big player has less power in terms of the Shapley-Shubik index than its proportional weight. In the next example the opposite holds:

Example 3. *For a single player, it may hold that its individual power to proportional weight ratio is unbounded: Consider $A = \{2, k\}, m = 1, T = k + 3$. Then $\phi_{a_1}(A, m, T) = \frac{1}{3}$ while $\frac{a_1}{m + \sum_{j=1}^{r} a_j} = \frac{2}{k+3}$.*

Nevertheless, there does exist an upper bound on the Shapley-Shubik index of any individual big player:

Theorem 2. $\phi_{a_i}(A, m, T) \leq \frac{a_i}{m + r}$.

[1]There is some loss of generality by fixing the parameters a_i, T to be exact multiples of the weight of the small player. Our model can be slightly generalized as follows: Let s be some minimal weight corresponding to some operational or regulatory minimal size of a venture, or to an electoral threshold for parliaments. Any player with an integer weight $s \leq w < 2s$ is termed "small" as small players are the ones that cannot split. The model as presented corresponds to the case $s = 1$ for tractability and readability but we believe that our results hold for the more general model as well.

4 The Banzhaf Index

We show a contrary result for the Banzhaf index [Banzhaf III, 1964], where an asymptotic example has an unbounded ratio of aggregate big players' power over their proportion.

Definition 5. *(See [Dubey and Shapley, 1979]) Let $P(S)$ be the power set of S, and UNI be the uniform distribution over a discrete set. The Absolute Banzhaf index of a WVG is:*

$$\beta'_{a_i}(A,m,T) = \mathbb{E}_{S\sim UNI(P(A\setminus\{a_i\}\cup M)}\left[v(S\cup\{a_i\}) - v(S)\right]$$
$$\beta'_1(A,m,T) = \mathbb{E}_{S\sim UNI(P(A\cup M\setminus\{1\})}\left[v(S\cup\{1\}) - v(S)\right].$$

The Normalized Banzhaf index is:

$$\beta_{a_i} = \frac{\beta'_{a_i}(A,m,T)}{\sum_{j=1}^{r}\beta'_{a_j} + m\beta'_1(A,m,T)}.$$

While the Shapley-Shubik index gives equal probabilities to all permutations over players, the Banzhaf index gives equal probabilities to all subsets of players. The normalization is needed to achieve the efficiency property, where summation over the indices of all players sums up to exactly one. The absolute Banzhaf indices may sum to less or more than 1. For an individual absolute Banzhaf index, by definition

$$\beta'_{a_i}(A,m,T) \le 1 \tag{1}$$

Definition 6. *The Banzhaf-proportional ratios are the global supremum over all WVGs*

$$\bar{R}_{\beta'} = \sup_{A,m,T}\frac{\sum_{i=1}^{r}\beta'_{a_i}(A,m,T)}{P(A,m,T)},$$

A similar definition holds for $\bar{R}_\beta$.

While the power of the big players cannot be much larger than their proportional weight according to the Shapley-Shubik index, the Bhanzaf index gives a different result:

Theorem 3. *$\bar{R}_{\beta'}, \bar{R}_\beta$ are unbounded.*

The proof of Theorem 3, given in the full version of the paper, shows that the ratio in the following example goes to infinity with k.

Example 4. *Consider $A = \{2k\}, m = k^{1.5}, T = \frac{2k+k^{1.5}}{2}$. Then for $k = 1600$, we have*

$$P(A,m,T) = \frac{3200}{3200+1600^{1.5}} = \frac{1}{21}, \beta'_{a_1}(A,m,T) \approx 0.999999.$$

The intuition for the calculation is as follows. Since there is only one big player, and all other players have a weight of one, only the size of the subset of other players matters for the index. Choosing a subset S of small identical players with uniform probability is like letting each small player participate in the chosen subset with probability $\frac{1}{2}$ (a Bernoulli trial). So, the size of the subset is sampled from the Binomial distribution with parameters $B(m, \frac{1}{2})$. Measure concentration properties of the symmetric binomial distribution around its mean imply that with high probability the size of the sampled set is close enough to $\frac{m}{2}$ so that the big player in the example is pivotal. Details of this calculation can be directly extracted from the argument in the proof of Theorem 3. The normalized Banzhaf index is for these parameters is:

$$\beta_{a_1}(A,m,T) \approx \frac{0.999999}{0.999999+1600^{1.5}\cdot 3.52795*10^{-33}} \approx 1.$$

This yields 21 for both the ratio of absolute Banzhaf to proportion and normalized Banzhaf to proportion.

Example 4 is in the spirit of Section b of [Penrose, 1946] and the asymptotic results analysed in Section 7 of [Dubey and Shapley, 1979]. As far as we know, the exact bound that we derive along with its the formal analysis are new.

Theorem 7 of [Aziz *et al.*, 2011] states that if a single player splits her votes between exactly two identities, its power as measured by the Banzhaf index cannot increase by a factor larger than 2. In contrast, Theorem 3 above shows that, with general splits, a player might end up *decreasing* its power by an unbounded factor. This paints an overall non-favorable picture for false-name manipulations as measured by the Banzhaf index.

5 The Deegan-Packel Index

Definition 7. *An all-pivotal set in a WVG is a set S such that for any player $s \in S$, $v(S) - v(S\setminus s) = 1$. Let the set of all-pivotal subsets be AP. The Deegan-Packel index is:*

$$\rho_{a_i}(A,m,T) = \mathbb{E}_{S\sim UNI(AP)}\left[\frac{\mathbb{1}_{a_i\in S}}{|S|}\right],$$

where UNI is the uniform disribution over a discrete set.

Thus, the Deegan-Packel index is similar to the Banzhaf index but takes into account only the all-pivotal subsets that are, in a sense, the minimal coalitions. It also considers the size of the coalition, so a participation in a large coalition results in less "power" than being a part of a small coalition. One can verify that this index is efficient, i.e., the sum of indices of all players is always exactly one.

Definition 8. *The Deegan-Packel-proportional ratio is the global supremum over all WVGs*

$$\bar{R}_\rho = \sup_{A,m,T}\frac{\sum_{i=1}^{r}\rho_{a_i}(A,m,T)}{P(A,m,T)}.$$

Example 5. *$\bar{R}_\rho \ge 2$: Let $A = \{k\}, m = k-1, T = k$. Then $\rho_{a_1}(A,m,T) = 1$, while $P(A,m,T) = \frac{k}{2k-1}$.*

Theorem 4. *$\bar{R}_\rho \le 3$.*

The theorem follows from lemmas given in appendices E,F. It exploits properties of the all-pivotal coalition in two threshold regimes: If the threshold is large, we show that enough small players must participate in an all-pivotal coalition, making the relative power of the big players in such a coalition small. If the threshold is small, we are able to use algebraic manipulations over binomials to derive the bound.

6 A Power Index False-name Game

In this section, we consider a framework general to all power indices under the possibility of votes split by big players in the voting game. Our discussion focuses on the Shapley-Shubik power index, where we give a conjecture with empirical results. We begin with a notation. Given a natural number a, define the integer partitions of a as

$$Partitions(a) = \bigcup_{i=1}^{a} \left\{ \{b_1, ..., b_i\} \middle| \sum_{j=1}^{i} b_j = a, \forall_{j=1}^{i} b_j \in \mathbb{N} \right\}$$

In words, the partitions of a are all the different sets of natural numbers such that their sum is a. Note that we allow several big players to split into multiple identities each, which is stronger than many other incentive analyses of false-name attack where only one strategic player is considered.

Definition 9 (The false-name weighted voting game for a power index $\alpha \in \{\phi, \beta, \rho\}$)**.** *Let $\{A, m, T\}$ be a WVG. We define a non-cooperative game with $|A|$ strategic players (which are the big players in the WVG). The strategy space of each strategic player is $Partitions(a_i)$. Given strategies $s_i = \{b_i^1, ..., b_i^{c_i}\}$ for $1 \le i \le r$, let $B = \{b_1^1, ..., b_1^{c_1}, ..., b_r^1, ..., b_r^{c_r}\}$. The payoff for player i is*

$$u_i^{\alpha}(s_1, ..., s_r) = \sum_{j=1}^{c_i} \alpha_{b_i^j}(B, m, T).$$

Let $c = \sum_{i=1}^{r} c_i$ stand for the total number of elements in B.

We wish to understand how the option to split (submit false-name bids) changes the power of the strategic players. The previous section sheds light on this question:

Theorem 5. *When $\alpha = \phi$ (the Shapley-Shubik index), then for any tuple of strategies $s_1, ..., s_r$,*

$$\sum_{i=1}^{r} u_i^{\phi}(s_1, ..., s_r) \le 2P(A, m, T).$$

In particular, this happens in any mixed or pure Nash or correlated equilibrium of the game.

A similar result holds for the Deegan-Packel power index.

Proof. By Theorem 1, considering B as determined by the set of strategies $s_1, ..., s_r$ as the set of big players in the theorem (and where the number of big players is c), we have

$$\sum_{i=1}^{r} \sum_{j=1}^{c_i} \phi_{b_i^j}(B, m, T) \le 2 \frac{\sum_{i=1}^{r} \sum_{j=1}^{c_i} b_i^j}{m + \sum_{i=1}^{r} \sum_{j=1}^{c_i} b_i^j} = 2 \frac{\sum_{j=1}^{r} a_j}{m + \sum_{j=1}^{r} a_j}.$$

□

Thus, strategically splitting vote weights using false-name manipulations cannot increase the overall power of the big players to be more than double their proportional weight. On the other hand, we conjecture that such strategic manipulations cannot harm their overall Shapley-Shubik power:

Conjecture 1. *For the Shapley-Shubik index ϕ with any weighted voting game $\{A, m, T\}$ and a choice of strategies resulting in a corresponding weighted voting game $\{B, m, T\}$, it holds that*

$$\sum_{i=1}^{r} \phi_{a_i}(A, m, T) \le 2 \sum_{i=1}^{r} \sum_{j=1}^{c_i} \phi_{b_i^j}(B, m, T).$$

Remark 1.

- *Section 6.2 supports the conjecture with empirical results obtained by an exhaustive search over small WVGs.*
- *Theorem 1 is a special case of the conjecture, where each player i's strategy choice is $\overbrace{\{1, ..., 1\}}^{a_i}$.*
- *Theorem 6 of [Aziz* et al.*, 2011] states that a single player that splits her votes to exactly two identities cannot decrease its Shapley-Shubik index by more than a factor of $\frac{n+1}{2}$. Our conjecture gives a much stronger bound for the aggregate power of big players: the worst decrease of aggregate power, caused by any combination of splits, is by a constant factor of 2.*

Combining Theorem 5 and Conjecture 1 yields:

Corollary 1. *For all strategies $s_1, ..., s_r$ in the Shapley false-name WVG $\{A, m, T\}$, if Conjecture 1 holds,*

$$\frac{1}{2} \sum_{i=1}^{r} \phi_{a_i}(A, m, T) \le \sum_{i=1}^{r} u_i^{\phi}(s_1, ..., s_r) \le 2P(A, m, T).$$

To conclude, false-name attacks can unboundedly increase the aggregate Shapley-Shubik power index of the big players, e.g., by splitting to singletons (Example 2). However, no attack can increase the power to more than twice the power resulting from the simple attack of splitting to singletons (Theorem 5). False-name attacks can also decrease the Shapley-Shubik index (Example 1). However, we believe, as expressed in Conjecture 1, that no false-name attack can decrease the aggregate Shapley-Shubik power to be less than one-half of the original power.

6.1 The Worst-case Effects of False-name Manipulation for a Single Player

While the total utility of the big players is conjectured to not lose much by splits, a single player may multiplicatively lose arbitrarily much in B compared to A. This is evident by Example 3, but we give two additional examples that do not require all players to fully split. In the first example, the player that chooses not to split loses by this choice. In the second example, the player that chooses to split loses by this choice.

Example 6. *Consider $A = \{k, k, k\}, B = \{k, \overbrace{1, ..., 1}^{k}, \overbrace{1, ..., 1}^{k}\}, m = k, T = 4k$. Then*

$$\phi_{a_1}(A, m, T) = \frac{1}{k+3} \qquad \phi_{b_1^1}(B, m, T) = \frac{1}{3k+1}.$$

Thus, with $k \ge 6$, the ratio is higher than 2. The example can be generalized to exceed any bound r, with $A = \overbrace{\{k, ..., k\}}^{r+1}, B = \{k, \overbrace{\overbrace{1, ..., 1}^{k}, ..., \overbrace{1, ..., 1}^{k}}^{r}\}, m = k, T = (r+2)k$.

Example 7. *Consider* $A = \{k, k\}, B = \{\overbrace{1, ..., 1}^{k}, k\}, m = 0, T = k+1$. *Then*

$$\phi_{a_1}(A, m, T) = \frac{1}{2}, \qquad \sum_{j=1}^{a_1} \phi_{b_1^j}(B, m, T) = \frac{1}{k+1}.$$

The basic upper bound on the power of a single big player that Theorem 2 yields continues to hold under the possibility of splits, and it decreases as the number of splits increase:

Corollary 2. *For a player* i*, and any strategy choice* $s_1, ..., s_r$ *of the players, it holds that:*

$$\boldsymbol{u_i^\phi(s_1, ..., s_r)} = \sum_{j=1}^{c_i} \phi_{b_i^j}(B, m, T) \overset{THM\ 2}{\leq} \sum_{j=1}^{c_i} \frac{b_i^j}{m + \sum_{k=1}^{r} c_k} \leq \sum_{j=1}^{c_i} \frac{b_i^j}{m + r + (c_i - 1)} = \frac{\boldsymbol{a_i}}{\boldsymbol{m + r + (c_i - 1)}}.$$

This further yields another result.

Corollary 3. *For the settings where there is a single big player* a_1 *and* m *small players, and any threshold* T*, the big player has a strategy that guarantees at least* $\frac{1}{2}$ *the power of its best possible strategy.*

Proof. We start by showing that $P(A, m, T) \geq \frac{1}{2} \sup_{s_1} u_1^\phi(s_1)$, and then give a strategy s_1 that attains $u_1^\phi(s_1) = P(A, m, T)$.

If $a_1 > m$, $P(A, m, T) > \frac{1}{2} \geq \frac{1}{2} \sup_{s_1} u_1^\phi(s_1)$, by the efficiency property of the Shapley-Shubik index.

Assume $a_1 \leq m$. The big player is the only strategic player in the game. By Corollary 2, for any strategy s_1,

$$u_1^\phi(s_1) \leq \frac{a_1}{m+1} \leq \frac{2a_1}{m + a_1} = 2P(A, m, T).$$

Thus, the proportional value for the big player is at least half as good as the best possible strategy. By the symmetry property of the Shapley-Shubik index, $s_1 = \{\overbrace{1, ..., 1}^{a_1}\}$ ("full split") guarantees the proportional value. □

6.2 Experimental Results

To support Conjecture 1, we ran an exhaustive validation over all WVGs with $m < 25$, $\sum_{j=1}^{r} a_j < 25$. A total of $5,833,920$ WVGs were checked against all valid sub-partitions of them, resulting in a total of $1,246,727,916$ valid pairs being compared. The maximal ratio attained was $1.958333'$. The minimal ratio attained was 0.08. The exhaustive search consists of two parts. First, using dynamic programming over a dual recursion to that of Lemma 1 (see Appendix I in the full version of the paper), we built a full recursion table of all Shapley-Shubik indices for the WVGs in the range. Then, for each WVG we considered all valid partition strategy sets B.

While the maximal ratio over all instances of the experimental analysis was close to 2, in most instances the ratio was much closer to 1. Figure 1 shows a histogram of the number of cases (on the y-axis) for different possible ratios between 0 and 2 (on the x-axis). As can be seen from the figure, the ratio is concentrated around 1.

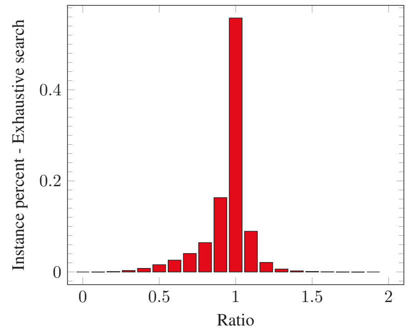

Figure 1: Ratio of big players' power before and after splits

7 Discussion and Future Directions

Many questions remain open. We find Conjecture 1 hard to prove even in limited settings. For example, consider the WVG $\{B, m, T\}$ where $m < T < \sum_{i=1}^{|B|} b_i$, i.e., the overall weights of the small players are less than the threshold, which itself is less than the overall weights of the big players. In this case it is possible to show that if we take $A = \{\sum_{i=1}^{|B|} b_i\}$, i.e., a single big player, then that player has a Shapley-Shubik index of 1 in the WVG $\{A, m, T\}$. The conjecture's inequality in that case then states that the sum of Shapley-Shubik indices of the big players in B is larger or equal to $\frac{1}{2}$. This reads as a very clean combinatorial problem: If we draw a permutation at random over a multiset of integers $m \times \{1\} \cup B$, with $m < T < \sum_{i=1}^{|B|} b_i$, then the probability that the pivotal player (crossing the threshold T) is "big" is higher than the probability that it is "small". This can be even simplified further if we assume all big players are of identical size k.

Section 6 is developed in regards to the Shapley-Shubik index. The negative nature of the results in section 4 make a similar treatment of the Banzhaf index superfluous, but the Deegan-Packel index might induce a similar conjecture and experimental results. The model itself could be generalized so that the threshold value T and big players' values A are not a multiple of the small players value, or into some other idea of looser distinctions between big and small players. Tight bounds can be derived for the Deegan-Packel index, and similar results explored for other power indices in common use, such as these by Johnston [1978], Holler and Packel [1983] and Coleman [1971]. Generalizing our results to a larger class of cooperative games is also interesting.

Acknowledgments

Yotam Gafni and Moshe Tennenholtz were supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 740435).

Ron Lavi was partially supported by the ISF-NSFC joint research program (grant No. 2560/17).

References

[Aziz *et al.*, 2011] Haris Aziz, Yoram Bachrach, Edith Elkind, and Mike Paterson. False-name manipulations in weighted voting games. *Journal of Artificial Intelligence Research*, 40:57–93, 2011.

[Bachrach *et al.*, 2011] Yoram Bachrach, Morteza Zadimoghaddam, and Peter Key. A cooperative approach to collusion in auctions. *SIGecom Exch.*, 10(1):17–22, March 2011.

[Bachrach *et al.*, 2016a] Yoram Bachrach, Yuval Filmus, Joel Oren, and Yair Zick. Analyzing power in weighted voting games with super-increasing weights. In Martin Gairing and Rahul Savani, editors, *Algorithmic Game Theory*, pages 169–181, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.

[Bachrach *et al.*, 2016b] Yoram Bachrach, Yuval Filmus, Joel Oren, and Yair Zick. A characterization of voting power for discrete weight distributions. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, page 74–80. AAAI Press, 2016.

[Banzhaf III, 1964] John F Banzhaf III. Weighted voting doesn't work: A mathematical analysis. *Rutgers L. Rev.*, 19:317, 1964.

[Chang *et al.*, 2006] Pao-Li Chang, Vincent Chua, and Moshé Machover. L.s. penrose's limit theorem: Tests by simulation. *Mathematical Social Sciences*, 51:90–106, 01 2006.

[Coleman, 1971] James S Coleman. Control of collectivities and the power of a collectivity to act. *Social choice*, pages 269–300, 1971.

[Deegan and Packel, 1978] J. Deegan and Edward Packel. A new index of power for simple n-person games. *International Journal of Game Theory*, 7:113–123, 01 1978.

[Dubey and Shapley, 1979] Pradeep Dubey and Lloyd S Shapley. Mathematical properties of the banzhaf power index. *Mathematics of Operations Research*, 4(2):99–131, 1979.

[Faliszewski and Hemaspaandra, 2009] Piotr Faliszewski and Lane Hemaspaandra. The complexity of power-index comparison. *Theoretical Computer Science*, 410(1):101 – 107, 2009.

[Holler and Packel, 1983] Manfred J. Holler and Edward W. Packel. Power, luck and the right index. *Zeitschrift für Nationalökonomie / Journal of Economics*, 43(1):21–29, 1983.

[Jelnov and Tauman, 2014] Artyom Jelnov and Yair Tauman. Voting power and proportional representation of voters. *International Journal of Game Theory*, 43(4):747–766, 2014.

[Johnston, 1978] Ronald John Johnston. On the measurement of power: Some reactions to laver. *Environment and Planning A*, 10(8):907–914, 1978.

[Lasisi and Allan, 2017] Ramoni O Lasisi and Vicki H Allan. False-name manipulation in weighted voting games: Empirical and theoretical analysis. *Computational Intelligence*, 33(3):478–506, 2017.

[Lindner and Machover, 2004] Ines Lindner and Moshé Machover. L.s. penrose's limit theorem: Proof of some special cases. *Mathematical Social Sciences*, 47:37–49, 01 2004.

[Milnor and Shapley, 1978] J. W. Milnor and L. S. Shapley. Values of large games ii: Oceanic games. *Mathematics of Operations Research*, 3(4):290–307, 1978.

[Morgenstern and Von Neumann, 1953] Oskar Morgenstern and John Von Neumann. *Theory of games and economic behavior*. Princeton university press, 1953.

[Neyman, 1981] Abraham Neyman. Singular games have asymptotic values. *Mathematics of Operations Research*, 6(2):205–212, 1981.

[Oren *et al.*, 2014] Joel Oren, Yuval Filmus, Yair Zick, and Yoram Bachrach. Power distribution in randomized weighted voting: the effects of the quota. https://arxiv.org/abs/1408.0442, 2014.

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

[Rey and Rothe, 2014] Anja Rey and Jörg Rothe. False-name manipulation in weighted voting games is hard for probabilistic polynomial time. *The Journal of Artificial Intelligence Research*, 50(1):573–601, May 2014.

[Riker, 1986] William H Riker. The first power index. *Social Choice and Welfare*, pages 293–295, 1986.

[Shapley and Shubik, 1954] Lloyd S Shapley and Martin Shubik. A method for evaluating the distribution of power in a committee system. *American political science review*, 48(3):787–792, 1954.

[Shapley, 1952] Lloyd S Shapley. A value for n-person games. Technical report, Rand Corp Santa Monica CA, 1952.

[Shapley, 1962] Lloyd S Shapley. Simple games: an outline of the descriptive theory. *Behavioral Science*, 7(1):59–66, 1962.

[Storing, 2008] Herbert J Storing. *The complete anti-federalist*, volume 2. University of Chicago Press, 2008.

[Słomczyński and Zyczkowski, 2006] Wojciech Słomczyński and K. Zyczkowski. Penrose voting system and optimal quota. *Acta Physica Polonica B*, 37:3133–3143, 2006.

[Słomczyński and Życzkowski, 2007] Wojciech Słomczyński and Karol Życzkowski. From a toy model to the double square root voting system. *Homo Oeconomicus*, 24(3-4):381–399, 2007.

[Winter, 2002] Eyal Winter. Chapter 53 the shapley value. volume 3 of *Handbook of Game Theory with Economic Applications*, pages 2025 – 2054. Elsevier, 2002.

[Zick *et al.*, 2011] Yair Zick, Alexander Skopalik, and Edith Elkind. The shapley value as a function of the quota in weighted voting games. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 490–496, 2011.

[Zick, 2013] Yair Zick. On random quotas and proportional representation in weighted voting games. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, pages 432–439, 2013.