The Winner-Take-All Dilemma*

Kazuya Kikuchi† Yukio Koriyama‡

October 29, 2021

Abstract

We consider collective decision making when the society consists of groups endowed with voting weights. Each group chooses an internal rule that specifies the allocation of its weight to the alternatives as a function of its members' preferences. Under fairly general conditions, we show that the winner-take-all rule is a dominant strategy, while the equilibrium is Pareto dominated, highlighting the dilemma structure between optimality for each group and for the whole society. We also develop a technique for asymptotic analysis and show Pareto dominance of the proportional rule. Our numerical computation for the US Electoral College verifies its sensibility.

*For their fruitful discussions and useful comments, we are thankful to Pierre Boyer, Micael Castanheira, Quoc-Anh Do, Shuhei Kitamura, Michel Le Breton, Takehito Masuda, Shintaro Miura, Takeshi Murooka, Matías Núñez, Alessandro Riboni, and the seminar participants at the 11th Pan Pacific Game Theory Conference, CREST, Hitotsubashi Summer Institute, Institut Henri Poincaré, Institute of Social and Economic Research at Osaka University, Osaka School of International Public Policy Lunch Seminar, Parisian Political Economy Workshop, Philipps-University of Marburg, Transatlantic Theory Workshop, Univerisity of Montpelier, University of Tokyo, Waseda University and the 2019 Autumn Meeting of the Japanese Economic Association. Financial support by Investissements d'Avenir, ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047 and DynaMITE: Dynamic Matching and Interactions: Theory and Experiments, ANR-13-BSHS1-0010, Waseda University Grant for Special Research Projects (2018K-014), JSPS KAKENHI Grant Numbers JP17K13706 and JP15H05728 are gratefully acknowledged.

†Hosei University. E-mail: kazuya.kikuchi68@gmail.com.

‡CREST, Ecole Polytechnique, Institut polytechnique de Paris. E-mail: yukio.koriyama@polytechnique.edu.

JEL classification: C72, D70, D72

Keywords: Representative democracy, winner-take-all rule, proportional rule, Prisoner's Dilemma.

1 Introduction

In many situations of collective decision making, including representative democracy, the society consists of distinct groups and decisions are made based on opinions aggregated within the groups. For example, in the United States presidential election, each state allocates its electoral votes based on the statewide popular vote. Another example is legislative voting, in which each party indicates how the legislators should vote based on the opinions of the party members.

In such situations, the social decision depends on the rules that the groups use to aggregate their members' opinions. However, if the groups choose their rules based on private motives, the resulting social decisions may not be desirable. The resulting decisions may even make all groups worse off than they could have been. This paper studies the relationship between groups' incentives and their welfare consequences.

Existing institutions use different rules, many of which pertain to how to allocate the weight assigned to each group. On the one hand, the *winner-take-all rule* devotes all the weight to the alternative preferred by the majority of its members. This rule has been used to allocate electoral votes in all but two states in the most recent US presidential elections. A council of national ministers, each with a weighted vote (e.g., the Council of the European Union), is another example, provided that the ministers represent their countries' majority interests. Party discipline which is frequently observed in legislative voting is also an example of the winner-take-all rule used by parties.

On the other hand, the *proportional rule* allocates a group's weight in proportion to the number of members who prefer the respective alternatives. In a wide range of parliamentary institutions at the regional, national and international levels, each group (e.g., constituency, prefecture or state) elects a set of representatives whose composition proportionally reflects its citizens' preferences. Alternatively, when the representatives are viewed as standing for parties rather than for states or prefectures, the proportional rule corresponds to a party's rule that allows its representatives to vote for or against proposals based on their own preferences, provided the composition of the party's representatives proportionally reflects the opinions of all party members.

The weight allocation rules are often exogenously given to all groups, but there are also cases where each group chooses its own rule. For instance, in national parliaments, how the representatives are elected from the respective constituencies is stipulated by national law. By contrast, parties often have control over how their representatives vote, by punishing those who violate the party lines. As another example, the US Constitution stipulates that it is up to each state to decide the way in which the presidential electoral votes are allocated (Article II, Section 1, Clause 2).

If groups are allowed to choose their own rules, each group may have an incentive to allocate the weight so as to increase the influence of its members' opinions on social decisions, at the cost of the other groups' influence. It is not clear whether such an incentive at the group level is compatible with desirable properties of the overall preference aggregation, such as Pareto efficiency. A society consisting of distinct groups thus faces a dilemma between each group's private incentive and the overall social objectives. To address this issue, we model the choice of rules as a noncooperative game.

In this paper, we consider a model of collective decision making where a society consists of groups endowed with voting weights. Each group chooses the rule to allocate its weight to the binary alternatives, and the winner is the one with the most weight. A *rule* for a group is a function that maps members' preferences (e.g., group-wide popular vote) to an allocation of the weight to the alternatives. Any Borel-measurable function is allowed, including the examples of the winner-take-all and proportional rules stated above. A *profile* is a specification of rules for all groups. We study the game in which the groups independently choose their rules, so as to maximize their members' expected welfare.

The main result of this paper is that the game is an n-player *Prisoner's Dilemma* (Theorem 1). The winner-take-all rule is a *dominant strategy*, i.e., it is an optimal strategy for each group, regardless of the rules chosen by the other groups. However, if each group has less than half of the total weight, then the winner-take-all profile is *Pareto dominated*, i.e., some other profile makes *every* group better off. In brief, no group has an incentive to deviate from the winner-take-all rule, but every group would be better off if all groups jointly move to another profile. The dilemma structure exists for any number of groups (>2) and with fairly little restriction on the joint distribution of preferences (Assumption 1). Members' preferences are allowed to be biased and correlated within and across groups. For example, the model can be applied to parties with distinct but overlapping political goals, or to states with different levels of support for specific alternatives, such as blue, red or swing states in the US elections.

The observation that the winner-take-all rule is an optimal strategy for groups is not new. As we will discuss in detail in Sections 1.1 and 3.1, earlier studies have already pointed out such incentives for groups in various voting situations. The theoretical prediction about group behavior is also consistent with the fact that it has been dominantly employed by the states in the US Electoral College since the 1830s in order to allocate presidential electoral votes, and also with the widely observed party discipline behaviors in assemblies. Despite the various problems or limitations that have been pointed out concerning the winner-take-all rule,[1] it is still used prevalently.[2]

The main contribution of this paper is to establish that under quite general circumstances, the winner-take-all profile is Pareto dominated, i.e., *every* group would be better off if all groups simultaneously changed their rules. This point should be distinguished from the conventional knowledge that the direct popular vote (i.e., majority voting by all individuals) max-

[1]There are multiple arguments against the winner-take-all rule. First, the winner of the election may be inconsistent with that of the popular vote (May (1948), Feix et al. (2004)). Such a discrepancy has happened five times in the history of the US presidential elections, including recently in 2000 and 2016. Second, it may cause reduced dimensionality: (i) the parties have an incentive to concentrate campaign resources only in the battleground states, and (ii) the voters' incentive to turn out or to invest in information may be small and/or uneven across states, since the probability of each voter being pivotal is so small under the winner-take-all rule, and even smaller in non-swing states. Although campaign resource allocation and voter turnout are important issues, they are beyond the scope of this paper.

[2]One recent attempt of reform took place in 2004 in Colorado, when a ballot initiative for an amendment to the state constitution was raised, proposing the proportional rule. The amendment did not pass, garnering only 34.1% approval.

imizes the *utilitarian* welfare of the society, as it maintains the possibility that *some* groups (e.g., small states) may benefit from the winner-take-all profile. We provide a counterexample in Example 1: a small group is strictly better off under the winner-take-all profile compared with both the direct popular vote and the proportional profile. Indeed, protecting minority states' interests is an oft-used argument by advocates of the Electoral College and its adherence to the winner-take-all rule. The welfare criterion used in Theorem 1 is Pareto dominance, which is obviously stronger than the utilitarian welfare evaluation: there exists a profile under which *every* group is better off than the winner-take-all profile. Example 1 shows that the dominant profile is not necessarily the proportional one nor the direct popular vote. If such is the case, what profile Pareto dominates the winner-take-all profile? A full characterization of the Pareto set is provided in Lemma 1.

To further study welfare properties, we turn to an asymptotic and normative analysis of the model. We consider situations where the number of groups is sufficiently large, and the preferences are independent across groups and distributed symmetrically with respect to the alternatives. Under these conditions, we show that the proportional profile *Pareto dominates* every other symmetric profile (i.e., one in which all groups use the same rule), including the winner-take-all one. The assumptions on the preference distribution abstract from the fact that some groups may prefer specific alternatives. Such an abstraction would be reasonable on the grounds that normative judgment about rules should not favor particular groups because of their characteristic preference biases. To see how many groups are typically sufficient for the asymptotic result, we provide numerical computations in a model based on the US Electoral College, using the current apportionment of electoral votes. The numerical comparisons indicate that the proportional profile does Pareto dominate the winner-take-all profile in the model with fifty states and a federal district.

While the above result suggests that the proportional profile asymptotically performs well in terms of efficiency, it is silent about the equality of individuals' welfare. We apply our model to study how rules affect the distribution of welfare, by examining an asymmetric profile called the *congressional district profile*. This profile is inspired by the Congressional District Method currently used by Maine and Nebraska, in which two electoral votes are allocated by the winner-take-all rule, and the remaining ones are awarded to the winner of each district-wide popular vote.[3] We show that the congressional district profile achieves a more equal distribution of welfare than any symmetric profile by making individuals in smaller groups better off.

A technical contribution of this paper is to develop an asymptotic method for analyzing players' expected welfare in weighted voting games. One of the major challenges in analyzing such games is their discreteness. By the nature of combinatorial problems, obtaining an analytical result often requires a large number of classifications by cases, which may include prohibitively tedious and complex tasks in order to obtain general insights. We overcome this difficulty by considering asymptotic properties of games in which there are a sufficiently large number of groups. This technique allows us to obtain an explicit formula that captures the asymptotic behavior of the probability of success for each individual, which holds for a wide class of distributions of weights among groups (the correlation lemma: Lemma 2).

1.1 Literature Review

The incentives for groups to use the winner-take-all rule have been studied by several papers. Hummel (2011) and Beisbart and Bovens (2008) analyze models of the US presidential elections. Gelman (2003) and Eguia (2011a,b) give theoretical explanations as to why voters in an assembly form parties or voting blocs to coordinate their votes. Their findings are coherent with our observation that the winner-take-all rule is a dominant strategy. In particular, Beisbart and Bovens (2008) and Gelman (2003) compare the winner-take-all and proportional profiles. Under the current apportionment of electoral votes in the US, Beisbart and Bovens (2008) nu-

[3]The idea of allocating a part of the votes by the winner-take-all rule and allowing the rest to be awarded to distinct candidates can be seen as a compromise between the winner-take-all and the proportional rules. Symbolically, the two votes allocated by the winner-take-all rule is the same number as the Senators in each state, while the rest is equal to the number of the House representatives. The idea behind such a mixture is in line with the logic supporting bicameralism, which is supposed to provide checks and balances between the states' autonomy and federal governance.

merically compares these profiles, in terms of inequality indices on citizens' voting power and the mean majority deficit, on the basis of *a priori* and *a posteriori* voting power measures. Gelman (2003) compares the case with coalitions of equal sizes in which voters coordinate their votes to the case without such coordination. Our analysis is based on Pareto dominance between profiles, and provides results which hold under general distribution of groups' weights or sizes. In that sense, Beisbart and Bovens's positive analysis is complementary to our normative analysis of properties of the proportional profile.

De Mouzon et al. (2019) provides a welfare analysis of popular vote interstate compacts, and shows that, for a regional compact, the welfare of member states is single-peaked as a function of the number of the participating states, while it is monotonically decreasing for the non-member states. The second effect dominates in terms of social welfare, unless a large majority (approximately more than $2/\pi \simeq 64\%$) of the states join the compact, implying that a small- or middle-sized regional compact is welfare detrimental. For a national compact, the total welfare is increasing, as it turns out that even non-members would mostly benefit from the compact, implying that the social optimum is attained when a majority joins the compact, i.e., the winner is determined by the national popular vote. Their findings are coherent with ours: if the winner-take-all rule is applied only to a subset of the groups, then the member states enjoy the benefit at the expense of the welfare loss of the non-member states, and the total welfare decreases. The social optimum is attained when the entire nation uses the popular vote. The possibility of the national popular vote as a coordination device is discussed also in Cloléry and Koriyama (2020).

The history, objectives, problems, and reforms of the US Electoral College are summarized, for example, in Edwards (2004), Bugh (2010) and Wegman (2020). One of the problems of the Electoral College most often scrutinized is its reduced dimensionality. The incentive of the candidates to concentrate their campaign resources in swing and decisive states is modeled in Strömberg (2008), which is coherent with the findings of the seminal paper in probabilistic voting by Lindbeck and Weibull (1987). Strömberg (2008) also finds that uneven resource allocation and unfavorable treatment of minority states would be mitigated by implementing a national popular vote, which is coherent with the classical findings by Brams and Davis (1974). Voters' incentive to turn out is investigated by Kartal (2015), which finds that the winner-take-all rule discourages turnout when the voting cost is heterogeneous.

Constitutional design of weighted voting is studied extensively in the literature. Seminal contributions are found in the context of power measurement: Penrose (1946), Shapley and Shubik (1954), Banzhaf (1968) and Rae (1946). Excellent summaries of theory and applications of power measurement are given by, above all, Felsenthal and Machover (1998) and Laruelle and Valenciano (2008). The tools and insights obtained in the power measurement literature are often used in the apportionment problem: e.g., Barberà and Jackson (2006), Koriyama et al. (2013), and Kurz et al. (2017).

Our analysis can be interpreted in the context of Bayesian mechanism design, by considering groups in our model as agents whose preference intensities are private information. Under this interpretation, Theorem 1 translates into an impossibility theorem which states that there is no social choice function that is Bayesian incentive compatible, Pareto efficient and non-dictatorial. This is consistent with the results obtained in earlier papers in Bayesian mechanism design, such as Börgers and Postl (2009), Azrieli and Kim (2014) and Ehlers et al. (2020). The precise statement of the impossibility theorem (Proposition 3) and a discussion of the mechanism design literature will appear in Subsection 3.2.

2 The Model

We consider collective decision making when a society consists of groups endowed with voting weights. We first describe the weighted voting mechanism (Section 2.1). We then construct a non-cooperative game in which each group chooses an internal rule that specifies the allocation of its weight to the alternatives as a function of its members' preferences (Section 2.2). Finally, we introduce social choice functions which include the weighted voting mechanism as a special case (Section 2.3).

2.1 Weighted Voting

Let us begin with the description of the social decision process. We consider a society partitioned into n disjoint groups: $i \in \{1, 2, \cdots, n\}$. Each group i is endowed with a voting weight $w_i > 0$.

The society makes a decision between two alternatives, denoted -1 and $+1$, through the following two voting stages: (i) each individual votes for his preferred alternative; (ii) each group allocates its weight between the alternatives, based on the group-wide voting result. The winner is the alternative that receives the majority of overall weight.

Let $\theta_i \in [-1, 1]$ denote the vote margin in group i at the first voting stage. That is, θ_i is the fraction of members of i preferring alternative $+1$ minus the fraction preferring -1.[4]

At the second stage, each group's allocation of weight is determined as a function of the group-wide margin.

Definition 1. A *rule* for group i is defined as a Borel-measurable[5] function:

$$\phi_i : [-1, 1] \to [-1, 1].$$

The value $\phi_i(\theta_i)$ is the group-wide weight margin, i.e., the fraction of the weight w_i allocated to alternative $+1$ minus that allocated to -1, given that the vote margin is θ_i. That is, the rule allocates $w_i\phi_i(\theta_i)$ more weight to alternative $+1$ than alternative -1.[6]

Let

$$\Phi = \{\phi_i | \text{Borel-measurable}\}$$

be the set of all admissible rules.

Examples of rules. Among all admissible rules, the following examples deserve particular attention.

[4]For example, $\theta_i = 0.2$ means that 60% of members of i prefer $+1$ and 40% prefer -1.

[5]Borel-measurability is needed to ensure that $\phi_i(\theta_i)$ is a well-defined random variable, when θ_i is a random variable.

[6]For example, if $w_i = 50$ and $\phi_i(\theta_i) = 0.2$, it means that the rule allocates 30 (resp. 20) units of weight to the alternative $+1$ (resp. -1) so that the weight margin in favor of the alternative $+1$ is $50 \times 0.2 = 30 - 20$.

(i) *Winner-take-all rule*: $\phi_i^{\mathrm{WTA}}(\theta_i) = \operatorname{sgn} \theta_i$.

(ii) *Proportional rule*: $\phi_i^{\mathrm{PR}}(\theta_i) = \theta_i$.

(iii) *Mixed rules*: $\phi_i^a(\theta_i) = a\phi_i^{\mathrm{WTA}}(\theta_i) + (1-a)\phi_i^{\mathrm{PR}}(\theta_i)$, $0 \le a \le 1$.

The winner-take-all rule devotes all the weight of a group to the winning alternative in the group. The proportional rule allocates the weight in proportion to the vote shares of the respective alternatives in the group. The mixed rule ϕ^a allocates the fixed ratio a of the weight by the winner-take-all rule and the remaining $1-a$ part by the proportional rule.

The social decision is the alternative that receives the majority of overall weight. In the case of a tie, we assume that each alternative is chosen with probability $\frac{1}{2}$. Thus, given the rules $\phi = (\phi_i)_{i=1}^n$ and the group-wide vote margins $\theta = (\theta_i)_{i=1}^n$, the social decision $d_\phi(\theta)$ is determined as follows:

$$d_\phi(\theta) = \begin{cases} \operatorname{sgn} \ \sum_{i=1}^n w_i\phi_i(\theta_i) & \text{if } \sum_{i=1}^n w_i\phi_i(\theta_i) \neq 0, \\ \pm 1 \text{ equally likely} & \text{if } \sum_{i=1}^n w_i\phi_i(\theta_i) = 0. \end{cases} \tag{1}$$

2.2 The Game

We now define the non-cooperative game Γ in which the n groups choose their own rules simultaneously.

The game is played under incomplete information about individuals' preferences, and hence about the group-wide vote margins. Each group chooses a rule so as to maximize the expected welfare of its members. Since rules are fixed prior to realization of the preferences, a pure strategy of the game is a function from the realization of members' preferences to the allocation of the weight.

Let Θ_i be a random variable that takes values in $[-1, 1]$ and represents the vote margin in group i.[7] We impose little restriction on the joint distribution of the random vector $\Theta = (\Theta_i)_{i=1}^n$. The precise assumption on the distribution will be stated later in this section (Assumption 1).

The ex post payoff for group i is the average payoff for its members from the social decision. For simplicity, we assume that each individual

[7] Throughout the paper, we use capital Θ_i for the representation of a random variable, and small θ_i for the realization.

obtains payoff 1 if he prefers the social decision and payoff -1 otherwise.[8] The average payoff of members of group i equals Θ_i or $-\Theta_i$ depending on whether the social decision is $+1$ or -1; more concisely, it is:

$$\Theta_i d_\phi(\Theta).$$

The ex ante payoff for group i, denoted $\pi_i(\phi)$, is the expected value of the above expression:

$$\pi_i(\phi) = \mathbb{E}\left[\Theta_i d_\phi(\Theta)\right]. \tag{2}$$

Let $\pi_i(x_i, \phi_{-i}|\theta_i)$ denote the interim payoff for group i if it chooses the weight margin $x_i \in [-1, 1]$ given the realization of the vote margin θ_i. It is obtained as the weighted average of the ex post payoffs θ_i and $-\theta_i$ from decisions $+1$ and -1 with the conditional probabilities:[9]

$$\begin{aligned} &\pi_i(x_i, \phi_{-i}|\theta_i) \\ &= \theta_i \mathbb{P}\left\{ w_i x_i + \textstyle\sum_{j\neq i} w_j \phi_j(\Theta_j) > 0 \middle| \Theta_i = \theta_i \right\} \\ &\quad - \theta_i \mathbb{P}\left\{ w_i x_i + \textstyle\sum_{j\neq i} w_j \phi_j(\Theta_j) < 0 \middle| \Theta_i = \theta_i \right\}. \end{aligned} \tag{3}$$

The ex ante and interim payoffs are thus related as follows:

$$\begin{aligned} &\phi_i \text{ maximizes } \pi_i(\phi_i, \phi_{-i}) \\ &\Leftrightarrow x_i = \phi_i(\theta_i) \text{ maximizes } \pi_i(x_i, \phi_{-i}|\theta_i) \text{ for almost every } \theta_i \in [-1, 1]. \end{aligned}$$

To summarize, the *game* Γ is the one in which: the players are the n groups; the strategy set for each group i is the set Φ of all rules; the payoff function for group i is π_i defined in (2).

The following is the assumption on the joint distribution of the group-wide margins.

Assumption 1. The joint distribution of group-wide margins $(\Theta_i)_{i=1}^n$ is

[8] The assumption is without loss of generality, as we show in Remark 2 that the model can be generalized to heterogeneous preference intensities.

[9] The term corresponding to the event of a tie (i.e., $w_i x_i + \sum_{j\neq i} w_j \phi_j(\Theta_j) = 0$) does not appear in the formula below, since we assume that the tie is broken fairly.

absolutely continuous and has full support $[-1,1]^n$.

Assumption 1 permits a wide variety of joint distributions of individuals' preferences, in which intra- and inter-group correlations and biases are possible. First, the assumption imposes no restriction on preference correlations within each group. Second, individuals' preferences may also be correlated across groups, since the group-wide margins $(\Theta_i)_{i=1}^n$ can be correlated. This allows us to capture situations where, for instance, residents of different states or members of different parties have common interest on some issues. Third, preferences may be biased toward a particular alternative, since Θ_i can be asymmetrically distributed. For instance, blue (resp. red) states in the US might be described as groups whose group-wide margins have a distribution biased to the left (resp. right). In contrast, swing states might be described as groups whose distributions are concentrated around zero.

Remark 1. *Success probability and voting power.* Our definition of group payoffs has the following interpretation based on the members' preferences. Let M_i be the set of individuals in group i, and $X_{im} \in \{-1,+1\}$ be the preferred alternative of member $m \in M_i$ in group i. Let us here redefine Θ_i as a *latent variable* that parametrizes the distribution of the random preferences in group i. Specifically, suppose X_{im} are independently and identically distributed conditional on the realization $(\theta_i)_{i=1}^n$ with the following probabilities for all $i = 1, \cdots, n$ and $m \in M_i$:

$$\begin{cases} \mathbb{P}\left\{X_{im} = +1 | \Theta_1 = \theta_1, \cdots, \Theta_n = \theta_n\right\} = (1+\theta_i)/2, \\ \mathbb{P}\left\{X_{im} = -1 | \Theta_1 = \theta_1, \cdots, \Theta_n = \theta_n\right\} = (1-\theta_i)/2. \end{cases} \tag{4}$$

Then, as the group size becomes large ($|M_i| \to \infty$), the Law of Large Numbers implies that the group-wide margin $\frac{1}{M_i}\sum_{m \in M_i} X_{im}$ indeed converges to Θ_i almost surely, which is consistent with our original definition of Θ_i

as the group-wide margin. Moreover,

$$\begin{aligned}
&\mathbb{P}\left\{X_{im} = d_\phi(\Theta)\right\} \\
&= \mathbb{E}\left[\mathbb{P}\left\{X_{im} = d_\phi(\Theta)|\Theta\right\}\right] \\
&= \mathbb{E}\left[\mathbb{P}\left\{X_{im} = 1, d_\phi(\Theta) = 1|\Theta\right\} + \mathbb{P}\left\{X_{im} = -1, d_\phi(\Theta) = -1|\Theta\right\}\right] \\
&= \mathbb{E}\left[\mathbb{P}\left\{d_\phi(\Theta) = 1|\Theta\right\}\frac{1+\Theta_i}{2} + \mathbb{P}\left\{d_\phi(\Theta) = -1|\Theta\right\}\frac{1-\Theta_i}{2}\right] \\
&= \frac{1}{2}\left(1 + \mathbb{E}\left[\mathbb{P}\left\{d_\phi(\Theta) = 1|\Theta\right\}\Theta_i + \mathbb{P}\left\{d_\phi(\Theta) = -1|\Theta\right\}(-\Theta_i)\right]\right) \\
&= \frac{1}{2}\left(1 + \mathbb{E}\left[\Theta_i d_\phi(\Theta)\right]\right).
\end{aligned}$$

Therefore, $\pi_i(\phi) = \mathbb{E}\left[\Theta_i d_\phi(\Theta)\right]$ is an affine transformation of the probability that the preferred alternative of a member m in group i coincides with the social decision ($X_{im} = d_\phi(\Theta)$), which is called *success* in the literature of voting power measurement (Laruelle and Valenciano (2008)). The objective of the group, formulated as the maximization of π_i, is thus equivalent to maximization of the probability of success.

Under the winner-take-all profile ϕ^{WTA}, π_i is closely related to the classical voting power indices studied in the literature. If $(\Theta_i)_{i=1}^n$ are independently, identically and symmetrically distributed (thus each group's preferred alternative is independently and equally distributed over $\{-1, +1\}$, called *Impartial Culture*), then π_i corresponds to the Banzhaf-Penrose index (Banzhaf (1965), Penrose (1946)) and $\mathbb{P}\left\{X_{im} = d_\phi(\Theta)\right\}$ to the Rae index (Rae (1946)), up to a multiplication by the constant $\mathbb{E}\left[|\Theta_i|\right]$. If $(\Theta_i)_{i=1}^n$ are perfectly correlated and symmetrically distributed (called *Impartial Anonymous Culture*; see, for example, Le Breton et al. (2016)), then π_i corresponds to the Shapley-Shubik index (Shapley and Shubik (1954)). □

Remark 2. *Heterogeneous preference intensities.* We have assumed that all individuals have the same preference intensities (i.e., each individual receives a unit payoff whenever she prefers the social decision), and that each group's objective is to maximize the ex ante average payoff of its members. However, our formal definition (2) can be generalized to heterogeneous preference intensities. It only suffices for the group-wide payoff from the social decision to be more generally defined, not necessarily as the average of members' payoffs with identical preference intensities.

To be more precise, suppose each group i receives a random payoff U_i^+ or U_i^- depending on whether the social decision is $+1$ or -1, where U_i^+ and U_i^- are assumed to take *any* values in $[0,1]$. Redefine the variable Θ_i as the payoff difference: $\Theta_i := U_i^+ - U_i^-$. Then the group's ex ante payoff from the social decision under profile ϕ is

$$\begin{aligned} u_i(\phi) &= \mathbb{E}\left[U_i^+ \frac{1+d_\phi(\Theta)}{2} + U_i^- \frac{1-d_\phi(\Theta)}{2}\right] \\ &= \frac{1}{2}\mathbb{E}\left[\Theta_i d_\phi(\Theta)\right] + \frac{1}{2}\mathbb{E}\left[U_i^+ + U_i^-\right] \\ &= \frac{1}{2}\pi_i(\phi) + \text{constant}. \end{aligned}$$

Since this is a positive affine transformation of $\pi_i(\phi)$, our model captures the general case where each group maximizes the expected group-wide payoff u_i. In particular, the group-wide payoffs U_i^+ and U_i^- can be any functions of members' payoffs including heterogeneous preference intensities.

Furthermore, by considering groups in our model as agents whose preference intensities are private information, we can consider our model in the context of more general and abstract Bayesian mechanism design problems. The n-agent setting will be useful in Section 3.2 where we clarify the underlying logic behind the results we obtain in Section 3.1. For that purpose, we introduce a formal definition of the social choice function in the following Subsection 2.3. □

2.3 Social Choice Functions

Let $\triangle(\{-1,+1\})$ be the set of all random variables taking values in $\{-1,+1\}$. A *social choice function (SCF)* is a Borel-measurable function[10]

$$d : [-1,1]^n \to \triangle(\{-1,+1\}).$$

The SCF assigns to each profile of realized vote margins $\theta = (\theta_i)_{i=1}^n \in [-1,1]^n$ a social decision $d(\theta)$ which may randomize between alternatives

[10] More precisely, an SCF is a function $d(\theta,\omega)$ of two variables, $\theta \in [-1,1]^n$ and $\omega \in \Omega$ for a sample space Ω, such that: for each θ, $d(\theta,\cdot) : \Omega \to \{-1,+1\}$ is a random variable; for each ω, $d(\cdot,\omega) : [-1,1]^n \to \{-1,+1\}$ is a Borel-measurable function.

-1 and $+1$. The decision function d_ϕ in game Γ is an example of an SCF.[11]

With a slight abuse of notation, we denote by $\pi_i(d)$ the ex ante payoff for group i under SCF d. By extending formula (2), we have the following expression:

$$\pi_i(d) = \mathbb{E}\left[\Theta_i d(\Theta)\right].$$

Our main analysis in Section 3.1 is based on game Γ, but the results have implications to mechanism design problems with general SCFs, which we summarize in Section 3.2.

3 The Dilemma

3.1 The Main Result

In game Γ, a rule (or strategy) ϕ_i for group i *weakly dominates* another rule ψ_i if $\pi_i(\phi_i, \phi_{-i}) \geq \pi_i(\psi_i, \phi_{-i})$ for any ϕ_{-i}, with strict inequality for at least one ϕ_{-i}. A rule ϕ_i is a *weakly dominant strategy* for group i if it weakly dominates every rule not equivalent to ϕ_i where we call two rules ϕ_i and ψ_i *equivalent* if $\phi_i(\theta_i) = \psi_i(\theta_i)$ for almost every θ_i (with respect to Lebesgue measure on $[-1, 1]$).

A profile ϕ *Pareto dominates* another profile ψ if $\pi_i(\phi) \geq \pi_i(\psi)$ for all i, with strict inequality for at least one i. If ϕ is not Pareto dominated by any profile, it is called *Pareto efficient*. Pareto dominance between SCFs is defined in the same way, based on the payoff functions $\pi_i(d)$ (see Section 2.3).

We first consider the case in which there is no 'dictator' group that can determine the winner by putting all its weight to one alternative (Theorem 1). Later we consider the case with such a group (Proposition 2).

Assumption 2. Each group has less than half the total weight: $w_i < \frac{1}{2}\sum_{j=1}^{n} w_j$ for all $i = 1, \cdots, n$.

Theorem 1. *Under Assumptions 1 and 2, game Γ is a Prisoner's Dilemma:*

[11] Randomness of $d_\phi(\theta)$ occurs when the weighted vote is tied.

(i) the winner-take-all rule ϕ_i^{WTA} is the weakly dominant strategy[12] *for each group i;*

(ii) the winner-take-all profile ϕ^{WTA} is Pareto dominated.

We use the following lemma to prove the theorem. An SCF d is called a *weighted majority rule* if there exists a vector $(\lambda_i)_{i=1}^n \in \mathbb{R}_+^n \setminus \{\mathbf{0}\}$ such that:

$$d(\theta) = \operatorname{sgn} \sum_{i=1}^{n} \lambda_i \theta_i \text{ for almost every } \theta \in [-1,1]^n.$$

In game Γ, a profile ϕ is called a *generalized proportional profile* if there exists a vector $(\lambda_i)_{i=1}^n \in [0,1]^n \setminus \{\mathbf{0}\}$ such that for each i,

$$\phi_i(\theta_i) = \lambda_i \theta_i \text{ for almost every } \theta_i \in [-1,1].$$

Two profiles ϕ and ψ are called *equivalent* if $d_\phi(\theta) = d_\psi(\theta)$ for almost every $\theta \in [-1,1]^n$.

Lemma 1. *(Characterization of the Pareto set) Under Assumption 1, the following statements hold:*

(i) An SCF d is Pareto efficient in the set of all SCFs if and only if it is a weighted majority rule.

(ii) In game Γ, a profile $\phi = (\phi_i)_{i=1}^n$ is Pareto efficient in the set of all profiles if and only if it is equivalent to a generalized proportional profile.

The proof of Lemma 1 is relegated to the Appendix.

Proof of Theorem 1. Part (i). We first check that

$$\pi_i(\phi_i^{\mathrm{WTA}}, \phi_{-i}) \geq \pi_i(\phi_i, \phi_{-i}) \tag{5}$$

for any (ϕ_i, ϕ_{-i}). By (3), if $\theta_i > 0$ (resp. $\theta_i < 0$), then the interim payoff $\pi_i(x_i, \phi_{-i}|\theta_i)$ is non-decreasing (resp. non-increasing) in $x_i \in [-1,1]$. We

[12] By the definition of weak dominance, ϕ_i^{WTA} is the unique weakly dominant strategy up to equivalence of rules.

thus have $\pi_i(\phi_i^{\mathrm{WTA}}(\theta_i), \phi_{-i}|\theta_i) \geq \pi_i(\phi_i(\theta_i), \phi_{-i}|\theta_i)$ for any (ϕ_i, ϕ_{-i}) and $\theta_i \neq 0$. Since $\Theta_i = 0$ occurs with probability 0, This implies (5).

Now we show that for any profile ϕ_{-i} in which each $\phi_j(\Theta_i)$ $(j \neq i)$ has full support $[-1, 1]$ (e.g., ϕ_j^{PR}), the strict inequality

$$\pi_i(\phi_i^{\mathrm{WTA}}, \phi_{-i}) > \pi_i(\phi_i, \phi_{-i}) \tag{6}$$

holds for any rule ϕ_i that differs from ϕ_i^{WTA} on a set $A \subset [-1, 1]$ of positive measure. To see this, note that for such ϕ_{-i} and any θ_i, the conditional distribution of $\sum_{j\neq i} w_j \phi_j(\Theta_j)$ given $\Theta_i = \theta_i$ has support

$$I = \left[-\sum_{j\neq i} w_j, \sum_{j\neq i} w_j\right].$$

Since $w_i < \sum_{j\neq i} w_j$ by Assumption 2, as x_i moves in $[-1, 1]$, $w_i x_i$ moves in interval I. Formula (3) thus implies that if $\theta_i > 0$ (resp. $\theta_i < 0$), then $\pi_i(x_i, \phi_{-i}|\theta_i)$ is strictly increasing (resp. decreasing) in $x_i \in [-1, 1]$. Hence $\pi_i(\phi_i^{\mathrm{WTA}}(\theta_i), \phi_{-i}|\theta_i) > \pi_i(\phi_i(\theta_i), \phi_{-i}|\theta_i)$ at any $\theta_i \in A$. Since Θ_i has full support, result (6) follows.

Part (ii). By the characterization of the Pareto set (Lemma 1(ii)), it suffices to check that ϕ^{WTA} is not equivalent to any generalized proportional profile. Suppose, on the contrary, that ϕ^{WTA} is equivalent to a generalized proportional profile with coefficients $\lambda \in [0, 1]^n \setminus \{\mathbf{0}\}$. Then, since $(\Theta_i)_{i=1}^n$ has full support,

$$d_{\phi^{\mathrm{WTA}}}(\theta) = \operatorname{sgn} \sum_{i=1}^{n} w_i \lambda_i \theta_i \text{ at almost every } \theta \in [-1, 1]^n. \tag{7}$$

Since no group dictates the social decision, the coefficients λ_i are positive for at least two groups. Without loss of generality, assume $\lambda_1 > 0$ and $\lambda_2 > 0$. Now, fix θ_i for $i \neq 1, 2$ so that they are sufficiently small in absolute value. Then, according to (7), for (almost any) sufficiently small $\varepsilon > 0$, $d_{\phi^{\mathrm{WTA}}}(\theta) = +1$ if $\theta_1 = 1 - \varepsilon$ and $\theta_2 = -\varepsilon$, while $d_{\phi^{\mathrm{WTA}}}(\theta) = -1$ if $\theta_1 = \varepsilon$ and $\theta_2 = -1 + \varepsilon$. This contradicts the fact that $d_{\phi^{\mathrm{WTA}}}(\theta)$ depends only on the signs of $(\theta_i)_{i=1}^n$. □

Theorem 1 shows that, while the dominant strategy for each group is the winner-take-all rule, the dominant-strategy equilibrium is Pareto dominated by a generalized proportional profile. This typical Social Dilemma (or, n-player Prisoner's Dilemma) situation suggests that a Pareto efficient outcome is not expected to be achieved under decentralized decision making, and a coordination device is necessary in order to attain a Pareto improvement.

The observation that groups have an incentive to use the winner-take-all rule is not new. Beisbart and Bovens (2008) consider Colorado's deviation from the winner-take-all rule to the proportional rule, following the state's attempt in 2004 to amend the state constitution, and show that the citizens in Colorado are worse off under both *a priori* and *a posteriori* measures. Hummel (2011) shows that a majority of the voters in a state is worse off by unilaterally switching to the proportional rule from the winner-take-all profile.

Our results are also consistent with the findings in the literature of the coalition formation games in which individuals may have incentive to raise their voices by forming a coalition and aligning their votes. Gelman (2003) illustrates that individuals are better off by forming a coalition and assign all their weights to one alternative. Eguia (2011a) considers a game in which the members in an assembly decide whether to accept the party discipline to align their votes, and shows that the voting blocs form in equilibrium if preferences are sufficiently polarized. Eguia (2011b) considers a dynamic model and shows the conditions under which voters form two polarized voting blocs in a stationary equilibrium.

A novelty of Theorem 1 lies in its generality. Earlier studies have introduced a specific structure either on the distribution of the preferences and/or of the weights, or on the set of the rules that groups can use.[13] In contrast, we only impose fairly mild conditions on the preference distribution (in particular, Assumption 1 imposes no restriction on across-group correlation), on the weight distribution (Assumption 2 imposes no specific

[13]Beisbart and Bovens (2008) consider Colorado's strategic choice between the winner-take-all and the proportional rules in the US Electoral College. Hummel (2011) either introduces a correlation structure in the preference distribution or assumes weights to be constant in other states. Gelman (2003) shows interesting computations, but all claims are based on observations from examples. Eguia (2011a) introduces a three-group preference structure, left, right and independent, and Eguia (2011b)'s main results focus on a nine-voter example, and the internal rules are assumed to be (super) majority rules.

weight structure such as one big group and several smaller ones, or equally sized groups), and on the set of the available rules (Definition 1 admits all Borel-measurable rules, not just the winner-take-all and the proportional rules).

Most importantly, the generality of our model allows for a welfare analysis which does not require introduction of a specific structure on the weight and/or the preference distribution, or the set of available rules. Since our model incorporates *all* Borel-measurable profiles, the Pareto set obtained in Lemma 1 leads us to an explicit characterization of the set of *first-best* outcomes which can be attained.

The key welfare implication of our result is that the dominant-strategy equilibrium is Pareto dominated by generalized proportional profiles. This provides us with two important insights in welfare analysis of groupwise preference aggregation problems. First, the game is a Prisoner's dilemma so that a coordination device is necessary for a Pareto improvement. Second, once such a device is available, our characterization lemma tells us that, at the first-best, the society should use rules that are proportional in nature, so that the cardinal information of the group-wide preferences is transmitted without distortion.

It is worth emphasizing that the result does not imply merely utilitarian (i.e., benthamite) inefficiency of the equilibrium profile. The profile is Pareto dominated, implying that it is in *every* group's interest to move from the winner-take-all equilibrium to another profile. From the utilitarian perspective, it is straightforward to see that the social optimum is obtained by the *popular vote*, i.e., direct majority voting by all individuals. However, this observation is not sufficient to establish that the winner-take-all profile is Pareto dominated.[14] After all, the utilitarian optimum is merely one point in the Pareto set.

The following example illustrates that the winner-take-all profile is not always Pareto dominated by either the popular vote or the proportional profile.

Example 1. Consider a society which consists of two large groups with an equal weight and one small group. For an illustrative purpose, let us

[14]Obviously, utilitarian optimality does not imply Pareto dominance.

consider three American states: Florida, New York and Wyoming. Their populations and weights are summarized in Table 1.

Table 1: Comparison of the expected payoffs in an example of the society which consists of three states: Florida, New York and Wyoming. Weights are the electoral votes assigned in the Electoral College in 2020. Population is an estimation of the voting-age population in 2018 (in thousands). Source: US Census Bureau.

State	Weight	Population	$\pi_i\left(\phi^{\text{WTA}}\right)$	$\pi_i\left(\phi^{\text{PR}}\right)$	$\pi_i\left(\phi^{\text{POP}}\right)$	$\pi_i\left(\hat{\phi}\right)$
Florida	29	15,047	0.250	0.332	0.343	0.271
New York	29	13,684	0.250	0.332	0.323	0.271
Wyoming	3	422	0.250	0.034	0.008	0.271
Per capita average			0.250	0.328	0.329	0.271

As defined above, ϕ^{WTA} and ϕ^{PR} are the winner-take-all and proportional profiles. The vote margins $(\Theta_i)_{i=1,2,3}$ are drawn from the uniform distribution on $[-1,1]$ independently across the states. The payoff of the popular vote $\pi_i\left(\phi^{\text{POP}}\right)$ is defined as the ex ante expected payoff of a representative voter in each state, which is obtained by letting the social decision d be the popular vote winner in (2).

Since there is no dictator state (i.e., Assumpion 2 is satisfied) in this example, any pair of two states is a minimal winning coalition under the winner-take-all profile, implying that the expected payoffs are exactly the same across states under ϕ^{WTA}.

The two larger states are better off under the proportional profile ϕ^{PR}, while the smaller state is worse off. This is because the social decision is more likely to coincide with the alternative preferred by the majority of the large states under ϕ^{PR}. As a consequence, the differences in the weights are reflected more directly on the differences in the expected payoffs.

Even though the small state is better off under ϕ^{WTA} in this particular example, it is worth underlining that whether the winner-take-all profile favors small states as compared to the proportional profile depends on the weight distribution. For example, if one state is a dictator (i.e. violating Assumption 2), the payoffs of the two other states are zero under ϕ^{WTA}, while they are (probably small but) strictly positive under ϕ^{PR}.

Under the popular vote ϕ^{POP}, the expected payoff of the small state is even smaller than under ϕ^{PR}. This comes from the fact that the weight assigned to the small state is larger than the large states in the *per capita* measure. In this example, Wyoming has more weight than it would if assigned proportionally to the population.[15] Under the popular vote, the citizens in the small state lose such an advantage assigned through the weights. We can also observe that the utilitarian (benthamite) welfare is maximized under the popular vote ϕ^{POP} by comparing the per capita average of the expected payoffs.

Finally, let $\hat{\phi}$ be the generalized proportional profile with coefficients $\lambda_i = 1/w_i$. We observe that it Pareto dominates ϕ^{WTA}. Remember that our characterization lemma tells us that a profile is Pareto efficient if and only if it is equivalent to a generalized proportional profile. We can show that among the profiles which Pareto dominate the equilibrium profile ϕ^{WTA}, one is obtained by letting $\lambda_i = 1/w_i$, because the expected payoffs are equal across the states in this example, and we can obtain the particular point in the Pareto set with the equal Pareto coefficients by setting $\lambda_i = 1/w_i$.

This example illustrates that the winner-take-all, proportional profiles, and the popular vote may be all Pareto imcomparable. Even though Theorem 1 shows that the winner-take-all profile is Pareto dominated, it may not be dominated by either the proportional profile or the popular vote. This may happen when the number of groups is small. For the cases in which there are sufficiently many groups, we provide clear-cut insights in Section 4 by using an asymptotic model and numerical simulations. □

To summarize, we have the following propositions.

Proposition 1. *Under Assumption 1, the proportional profile ϕ^{PR} and the popular vote ϕ^{POP} are both Pareto efficient.*

Proof. Trivially, the proportional profile is a generalized proportional profile by letting $\lambda_i = 1$ for all i. The outcome of the popular vote coincides with that of the generalized proportional profile with $\lambda_i = n_i/w_i$ for all i. By Lemma 1, we obtain the result. □

[15]The digressive proportionality is a consequence of the rule specified in the US Constitution. The number of electoral votes of each state is the sum of the numbers of Senate members (constant) and of the House (proportional to population in principle). Under such a rule, per capita weight is decreasing in population.

Proposition 2. *Under Assumption 1, the winner-take-all profile ϕ^{WTA} is Pareto dominated if and only if Assumption 2 is satisfied.*

Proof. The "if" part is already proven in Theorem 1 (ii). To show the "only if" part, suppose that Assumption 2 is violated. Then, there exists a dictator state i^* that can determine the winner by putting all its weight to the alternative preferred by the majority of the state. Hence, ϕ^{WTA} is equivalent to the generalized proportional profile with coefficients $\lambda_{i^*} > 0$ and $\lambda_i = 0$ for all $i \neq i^*$. By Lemma 1, we obtain the result. □

3.2 An impossibility theorem underlying the WTA Dilemma

In order to provide an interpretation of the result obtained in Theorem 1 in the context of mechanism design, we state an impossibility theorem that underlies the winner-take-all dilemma.

We show in Remark 2 above that there is a direct analogy between the non-cooperative voting game Γ considered in Theorem 1 and the Bayesian collective decision problem in which each agent's preferences including the intensity level are private information. In order to elucidate the logic behind our theorem, it is thus useful to consider a model of social choice function of which the cardinal preferences are the input.

Consider a society which consists of n agents ($i = 1, \cdots, n$) and which makes a collective decision between two alternatives $+1$ and -1. As we described in Remark 2, each group is a player in the voting game Γ, while it can be seen more generally as an agent whose preference intensity is represented by a von Neumann-Morgenstern utility function. Let U_i^+ (resp. U_i^-) be agent i's utility from the alternative $+1$ (resp. -1). We assume that the utilities are random variables whose support is included in a bounded interval, which we suppose as $[0, 1]$ without loss of generality.

Agent i's *type* is represented by the utility difference $\Theta_i := U_i^+ - U_i^-$. Then, Θ_i is a random variable taking a value in $[-1, 1]$. The type is private information: each agent observes only his own type. We only impose absolute continuity and full support of the joint distribution (Assumption 1). This allows for correlations of types, and ex ante asymmetries with respect to the agents and the alternatives.

A social choice function (SCF) d is defined as in Section 2.3. For each profile of realized types $\theta = (\theta_i)_{i=1}^n$, $d(\theta)$ is a random variable which takes a value either $+1$ or -1. An SCF is *dictatorial* if there exists an agent i such that $d(\theta)$ assigns probability one to the alternative sgn θ_i for almost every $\theta \in [-1, 1]^n$. Note that a weighted majority rule is dictatorial if and only if $\lambda_i > 0$ for one i and $\lambda_j = 0$ for all $j \neq i$.

We consider the *direct mechanism* associated with SCF d. Each of n agents simultaneously reports a type, based on which an alternative is chosen according to d. A strategy for agent i is a Borel-measurable function $\sigma_i : [-1, 1] \to [-1, 1]$ that assigns to each realization of type $\theta_i \in [-1, 1]$ a reported type $\sigma_i(\theta_i) \in [-1, 1]$. A strategy σ_i is called *truthful* if $\sigma_i(\theta_i) = \theta_i$ for almost every θ_i. Given a strategy profile $\sigma = (\sigma_i)_{i=1}^n$, the ex ante payoff for agent i induced by d is:

$$\pi_i(\sigma; d) = \mathbb{E}\left[\Theta_i d(\sigma(\Theta))\right]$$

where $\sigma(\Theta) = (\sigma_j(\Theta_j))_{j=1}^n$ is the profile of reported types.

The game Γ defined in Section 2.2 is thus exactly the one induced by the direct mechanism associated with the weighted majority rule d with coefficients $\lambda_i = w_i$ $(i = 1, \cdots, n)$. Call group i in game Γ as agent i, and its group-wide vote margin Θ_i as the agent's type. The strategy set Φ for group i in that game is the same as the strategy set for agent i in the direct mechanism. The definition (1) of the social decision $d_\phi(\theta)$ in game Γ is exactly the same as the decision $d(\phi(\theta))$ in the direct mechanism in which the strategy profile σ coincides with ϕ. Therefore, the ex ante payoff functions in the two models also coincide.

An SCF is *Bayesian incentive compatible* (BIC) if the profile of truthful strategies is a Bayesian Nash equilibrium of the direct mechanism. By the revelation principle, it is without loss of generality to consider only direct mechanisms.

The following is the impossibility result which underlies the WTA dilemma.

Proposition 3. *Under Assumption 1, an SCF is Pareto efficient and Bayesian incentive compatible if and only if it is dictatorial.*

Proof. It is obvious that every dictatorial SCF is Pareto efficient and Bayesian incentive compatible. By Lemma 1(i), it suffices to check that if a weighted majority rule d is not dictatorial, then it is not Bayesian incentive compatible. In the proof of Theorem 1(i), we have shown that in game Γ, if ϕ_{-i} is such that each $\phi_j(\Theta_j)$ ($j \neq i$) has full support $[-1, 1]$, the unique (up to equivalence) best response for group i is the winner-take-all rule. Thus, in the direct mechanism for d, the unique (up to equivalence) best response for each agent i against the profile in which all other agents play a truthful strategy is $\sigma_i(\theta_i) = \text{sgn}\ \theta_i$, which is again not a truthful strategy. Thus the profile of truthful strategies is not a Bayesian Nash equilibrium. □

The essence of the impossibility described in Proposition 3 lies in the fundamental incompatibility between Pareto efficiency and equilibrium behavior in the cardinal preference aggregation problem.

In order to understand where the incompatibility comes from, consider the classical Gibbard-Satterthwaite Theorem, which states impossibility of achieving both strategyproofness and non-dictatorship in the *ordinal* preference aggregation problem. Relaxing the strategyproofness condition to Bayesian incentive compatibility requires the introduction of expected payoff, as BIC is defined on the solution concept of Bayesian Nash equilibrium. This means that we need to consider a *cardinal* preference aggregation problem.

An impossibility result analogous to the Gibbard-Satterthwaite theorem is no longer obtained, when the SCF takes cardinal preferences as its input. This is shown by a counterexample: the winner-take-all profile is a non-dictatorial SCF which satisfies Bayesian incentive compatibility. BIC alone is not sufficient to imply dictatorship.

Essentially, only ordinal information can be aggregated when Bayesian incentive compatibility is required. To see why, suppose that two preferences types u and v are in affine transformation, that is, there exists $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u = \alpha v + \beta \mathbb{1}$ (call such a transformation as *purely cardinal*). If the outcome differs by reporting between u and v, the incentive compatibility of either u or v should be violated. To be more precise, consider a purely cardinal change in preferences. The agent's preferences on the lotteries over the alternatives are unchanged. If the lottery over the alternatives changes by a purely cardinal change of the agent's report, it means that she can manipulate the outcome even though her preferences over the lotteries are unchanged, implying a violation of incentive compatibility. Therefore, by requiring Bayesian incentive compatibility, the outcome should be equivalent up to purely cardinal changes, and thus only ordinal information can be aggregated at most.

Our observation that BIC implies ordinal aggregation is coherent with the results obtained in the literature. Azrieli and Kim (2014) characterize the second-best social choice functions and show that they are ordinal qualified weighted majority rules. Ehlers et al. (2020) provide a thorough analysis of the conditions under which BIC implies ordinality.

On the other hand, Pareto efficiency requires aggregation of cardinal information. To see why, remember that a Pareto efficient allocation should solve the maximization problem of the social welfare weighted by Pareto coefficients. By definition, weighted social welfare depends continuously on the cardinal preferences of each agent.[16] Therefore, cardinal information concerning the agents' preferences (such as intensity) should be reflected continuously to the social outcome when Pareto efficiency is required.

In sum, requiring incentive compatibility implies ordinal aggregation, while requiring Pareto efficiency implies cardinal aggregation. The fundamental property behind Proposition 3 is the incompatibility between the two types of aggregation. The result is coherent with the impossibility theorem obtained in Börgers and Postl (2009) in case of three alternatives and two agents. Ehlers et al. (2020) provides a general result for any number of alternatives by showing that a weighted utilitarian SCF is dictatorial if and only if it satisfies BIC under an independence condition (Theorem 8).

When applied to the group-wide voting problem, Proposition 3 provides an interpretation of our main result stated in Theorem 1. Behind the dilemma structure stated in Theorem 1 lies the impossibility of reconciling both Bayesian incentive compatibility and Pareto efficiency, proven in a more abstract mechanism design context in Proposition 3.

[16]Moreover, it is due to its linearity that we could provide a full characterization of the Pareto frontier by the weighted majority rules in Lemma 1.

4 Asymptotic and Computational Results

4.1 Asymptotic Analysis

We saw above that the game is a Prisoner's Dilemma. In this section, we provide further insights on the welfare properties, by focusing on the following situations in which: (i) the number of groups is sufficiently large, and (ii) the preferences of the members are distributed symmetrically. These properties allow us to provide an asymptotic and normative analysis.

Often the difficulty of analysis arises from the discrete nature of the problem. Since the social decision D_ϕ is determined as a function of the sum of the weights allocated to the alternatives across the groups, computing the expected payoffs may require the classification of a large number of success configurations which increases exponentially as the number of groups increases, rendering the analysis prohibitively costly. We overcome this difficulty by studying asymptotic properties. In order to check the sensibility of our analysis, we provide Monte Carlo simulation results later in the section, using an example of the US Electoral College.

In order to study asymptotic properties, let us consider a sequence of weights $(w_i)_{i=1}^{\infty}$, exogenously given as a fixed parameter.

Assumption 3. The sequence of weights $(w_i)_{i=1}^{\infty}$ satisfies the following properties.

(i) $w_1, w_2, \cdots$ are in a finite interval $[\underline{w}, \bar{w}]$ for some $0 \leq \underline{w} < \bar{w}$.

(ii) As $n \to \infty$, the statistical distribution G_n induced by $(w_i)_{i=1}^{n}$ weakly converges to a distribution G with support $[\underline{w}, \bar{w}]$.[17]

Assumption 3 guarantees that for large n, the statistical distribution of weights G_n is sufficiently close to some well-behaved distribution G, on which our asymptotic analysis is based.

Additionally, we impose an impartiality assumption for our normative analysis:

[17]The statistical distribution function G_n induced by $(w_i)_{i=1}^{n}$ is defined by $G_n(x) = \#\{i \leq n | w_i \leq x\}/n$ for each x. G_n weakly converges to G if $G_n(x) \to G(x)$ at every point x of continuity of G.

Assumption 4. The variables $(\Theta_i)_{i=1}^{\infty}$ are drawn independently from a common symmetric distribution F.

As in Felsenthal and Machover (1998), a normative analysis requires impartiality, and a study of fundamental rules in the society, such as a constitution, should be free from specific dependence on the ex post realization of the group characteristics. Assumption 4 allows our normative analysis to abstract away the distributional details. Of course, a normative analysis is best complemented by a positive analysis which takes into account the actual characteristics of the distributions, as in Beisbart and Bovens (2008).

Following the symmetry of the preferences, our analysis also focuses on *symmetric* profiles, in which all groups use the same rule: $\phi_i = \phi$ for all i. With a slight abuse of notation, we write ϕ both for a *single rule* ϕ and for the *symmetric profile* $(\phi, \phi, \cdots)$, which should not create confusion as long as we refer to symmetric profiles. As for the alternatives, it is natural to consider that the label should not matter when the group-wide vote margin is translated into the weight allocation, given the symmetry of the preferences.

Assumption 5. We assume that the rule is monotone and neutral, that is, ϕ is a non-decreasing, odd function: $\phi(\theta_i) = -\phi(-\theta_i)$.

Let $\pi_i(\phi; n)$ denote the expected payoff for group $i (\leq n)$ under profile ϕ when the set of groups is $\{1, \cdots, n\}$ and each group j's weight is w_j, the jth component of the sequence of weights. The definition of $\pi_i(\phi; n)$ is the same as $\pi_i(\phi)$ in the preceding sections; the new notation just clarifies its dependence on the number of groups n.

The main welfare criterion employed in this section is the asymptotic Pareto dominance.

Definition 2. For two symmetric profiles ϕ and ψ, we say that ϕ *asymptotically Pareto dominates* ψ if there exists N such that for all $n > N$ and all $i = 1, \cdots, n$,

$$\pi_i(\phi; n) > \pi_i(\psi; n).$$

4.2 Pareto Dominance

The following is the main result in our asymptotic analysis.

Theorem 2. *Under Assumptions 1-5, the proportional profile asymptotically Pareto dominates all other symmetric profiles. In particular, it asymptotically Pareto dominates the dominant-strategy equilibrium of the game, i.e., the symmetric winner-take-all profile.*

We use the following lemma to prove Theorem 2. The proof of Lemma 2 is relegated to the Appendix. The proof of part (ii) uses a more general result, Lemma 3, stated in the next subsection, whose proof also appears in the Appendix.

Lemma 2. *Under Assumptions 1-5, the following statements hold.*

(i) For any symmetric profile ϕ,

$$\pi_i(\phi; n) = 2\int_0^1 \theta_i \mathbb{P}\left\{ -w_i\phi(\theta_i) < \sum_{j\le n,\, j\ne i} w_j\phi(\Theta_j) \le w_i\phi(\theta_i)\right\} dF(\theta_i).$$

(ii) For any symmetric profile ϕ, as $n \to \infty$,

$$\sqrt{2\pi n}\,\pi_i(\phi; n) \to 2w_i\sqrt{\frac{\mathbb{E}[\Theta^2]}{\int_{\underline{w}}^{\bar{w}} w^2 dG(w)}}\,\mathrm{Corr}\,[\Theta, \phi(\Theta)],$$ [18]

uniformly in $w_i \in [\underline{w}, \bar{w}]$, where Θ is a random variable having the same distribution F as Θ_i. The limit depends on the profile ϕ only through the factor $\mathrm{Corr}[\Theta, \phi(\Theta)]$.

Proof of Theorem 2.
The heart of the proof is in the correlation result shown in part (ii) of Lemma 2. It follows that if correlation of $\phi(\Theta)$ with Θ is higher than that of $\psi(\Theta)$, then for each group i, there exists N_i such that if the number of

[18] Since Θ and $\phi(\Theta)$ are symmetrically distributed, the correlation is given by $\mathrm{Corr}\,[\Theta, \phi(\Theta)] = \mathbb{E}[\Theta\phi(\Theta)]/\sqrt{\mathbb{E}[\Theta^2]\mathbb{E}[\phi(\Theta)^2]}$ unless $\phi(\Theta)$ is almost surely zero. If $\phi(\Theta)$ is almost surely zero, then the correlation is zero.

groups (n) is greater than N_i, group i ($\leq n$) will be better off under ϕ than ψ.

Note that the convergence in part (ii) of Lemma 2 is uniform in $w_i \in [\underline{w}, \bar{w}]$. This implies that the convergence is uniform in $i = 1, 2, \cdots$.[19] Thus there is N with the above property, without subscript i, which applies to all groups $i = 1, 2, \cdots$. Therefore, if correlation of $\phi(\Theta)$ with Θ is higher than that of $\psi(\Theta)$, then ϕ asymptotically Pareto dominates ψ.

Since the perfect correlation $\mathrm{Corr}[\Theta, \phi^{\mathrm{PR}}(\Theta)] = 1$ is attained by the proportional rule, Theorem 2 follows. □

The above results show that the winner-take-all rule is characterized by its strategic dominance, while the proportional rule is characterized by its asymptotic Pareto dominance. The following proposition provides a complete Pareto order among all the linear combinations of the two rules.

Remember that we defined the mixed rules in Section 2 above. For $0 \leq a \leq 1$, a fraction a of the weight is assigned to the winner of the group-wide vote, while the rest, $1 - a$, is distributed proportionally to each alternative:

$$\phi^a(\theta_i) = a\phi^{\mathrm{WTA}}(\theta_i) + (1-a)\phi^{\mathrm{PR}}(\theta_i).$$

Proposition 4. *Under Assumptions 1-4, mixed profile ϕ^a asymptotically Pareto dominates mixed profile $\phi^{a'}$ for any $0 \leq a < a' \leq 1$. In particular, the proportional profile asymptotically Pareto dominates any mixed profile ϕ^a for $0 < a < 1$, which in turn asymptotically Pareto dominates the winner-take-all profile. In other words, all mixed profiles can be ordered by asymptotic Pareto dominance, from the proportional profile as the best, to the winner-take-all profile as the worst.*

Proof. In Appendix. □

The winner-take-all rule is not only asymptotically Pareto inefficient,

[19] A more detailed explanation of this step is the following. By Lemma 2 (i), $\sqrt{2\pi n}\pi_i(\phi; n))$ asymptotically behaves as $2\sqrt{2\pi n}\int_0^1 \theta\mathbb{P}\{-w_i\phi(\theta) < \sum_{j\leq n} w_j\phi(\Theta_j) \leq w_i\phi(\theta)\}dF(\theta)$, where whether the sum $\sum_{j\leq n} w_j\phi(\Theta_j)$ includes the ith term or not is immaterial in the limit. The estimate of $\sqrt{2\pi n}\pi_i(\phi; n)$ therefore has the form $f_n(w_i)$, where $f_n(x) := 2\sqrt{2\pi n}\int_0^1 \theta\mathbb{P}\{-x\phi(\theta) < \sum_{j\leq n} w_j\phi(\Theta_j) \leq x\phi(\theta)\}dF(\theta)$. Lemma 2 (ii) implies that $f_n(x)$ converges uniformly in $x \in [\underline{w}, \bar{w}]$, which in turn implies that the convergence of $\sqrt{2\pi n}\pi_i(\phi; n) \approx f_n(w_i)$ is uniform in $i = 1, 2, \cdots$.

but the worst among the symmetric mixed profiles. Is it worse than *any* other symmetric profile? We provide an answer in Remark 3 below.

Remark 3. *What is the worst profile?* Theorem 2 leaves the natural question of whether the winner-take-all profile is the worst among all symmetric profiles, in terms of asymptotic Pareto dominance. The answer is negative. To see this, note first that, for the winner-take-all profile, the correlation in Lemma 2 is strictly positive: $\mathrm{Corr}[\Theta, \phi^{\mathrm{WTA}}(\Theta)] = \mathbb{E}(|\Theta|)/\sqrt{\mathbb{E}(\Theta^2)} > 0$. On the other hand, for the symmetric profile ϕ^0 in which the rule is defined by $\phi^0(\theta) = 0$ for almost all θ, the correlation is obviously zero. This rule assigns exactly half of the weight to each alternative, regardless of the group-wide vote. Thus the profile ϕ^0 is the worst among all symmetric profiles, as the social decision is made by a coin toss almost surely, yielding expected payoff 0 to all groups. In the rest of this section, we exclude such a trivial profile from our consideration.

4.3 Congressional District Method

The analysis in the preceding subsection suggests that the proportional profile is optimal in terms of Pareto efficiency. However, our model also implies that this profile produces an unequal distribution of welfare; in fact, this unequal nature pertains to all symmetric profiles. The Correlation Lemma 2 (ii) shows that for these profiles, the expected payoff for a group is asymptotically proportional to its weight, providing high expected payoffs to groups with a large weight.

In this subsection, we examine whether such inequality can be alleviated without impairing efficiency by using an asymmetric profile, based on the Congressional District Method, currently used in Maine and Nebraska. This profile allocates a fixed amount c of each group's weight by the winner-take-all rule and the rest by the proportional rule:

$$w_i \phi^{\mathrm{CD}}(\theta_i, w_i) = c\phi^{\mathrm{WTA}}(\theta_i) + (w_i - c)\phi^{\mathrm{PR}}(\theta_i).$$

We consider the *congressional district profile* ϕ^{CD} in which the rule is used by all groups. Note that the profile is not symmetric in the sense that we defined at the beginning of this section. Since the weight allocation rule depends on w_i, which is heterogeneous across groups, ϕ_i is not the same function of θ_i for all i. Therefore, we cannot apply Theorem 2 in order to obtain a Pareto dominance relationship. However, we can obtain a small-group advantage result (Theorem 3) and a Lorenz dominance result (Theorem 4). To ensure that the profile is well-defined, we impose that the lower bound of weights $\underline{w}$ is strictly positive and $c \in (0, \underline{w}]$.

Theorem 3. *Under Assumptions 1-5, let us consider the congressional district profile with parameter $c \leq \underline{w}$. For any symmetric profile ϕ, there exists $w^* \in [\underline{w}, \bar{w}]$ with the following property: for any $\varepsilon > 0$, there is N such that for all $n > N$ and $i = 1, \cdots, n$,*

$$w_i < w^* - \varepsilon \Rightarrow \pi_i(\phi^{\mathrm{CD}}; n) > \pi_i(\phi; n),$$
$$w_i > w^* + \varepsilon \Rightarrow \pi_i(\phi^{\mathrm{CD}}; n) < \pi_i(\phi; n).$$

The proof of Theorem 3 uses the following lemma, which shows that the correlation lemma holds for a class of profiles such that the weight allocation rules have the following specific form of separability. Its proof and the Local Limit Theorem used in the proof are relegated to the Appendix.

Assumption 6. Let $\phi = (\phi_i)_{i=1}^{\infty}$ be a profile. There exist functions h_1, h_2, h_3 such that

$$w_i \phi_i(\theta_i, w_i) = h_1(w_i) h_2(\theta_i) + h_3(w_i) \operatorname{sgn} \theta_i, \text{ for all } i$$

where (i) h_1 is bounded, (ii) h_2 is an odd function such that the support of the distribution of $h_2(\Theta_i)$ contains 0, and (iii) h_3 is continuous but not constant.[20]

It is straightforward to show that Assumption 6 is satisfied for any symmetric profile as well as the congressional district profile. For a symmetric profile ϕ, let $h_1(w_i) = w_i$, $h_2(\theta_i) = \phi(\theta_i) - r \operatorname{sgn} \theta_i$, and $h_3(w_i) = w_i r$ where $r > 0$ is any positive number in the support of the distribution of

[20]Under this form, $\phi_i(\cdot, \cdot)$ is the same for all i so that we can omit subscript i whenever there is no confusion.

$\phi(\Theta)$.[21] For the congressional district profile ϕ^{CD}, let $h_1(w_i) = w_i - c$, $h_2(\theta_i) = \theta_i - \operatorname{sgn}\theta_i$, and $h_3(w_i) = w_i$.

Lemma 3. *Under Assumptions 1-5, let ϕ be a profile which satisfies Assumption 6. Then, as $n \to \infty$,*

$$\sqrt{2\pi n}\pi_i(\phi; n) \to \frac{2w_i\mathbb{E}[\Theta\phi(\Theta, w_i)]}{\sqrt{\int_{\underline{w}}^{\bar{w}} w^2\mathbb{E}[\phi(\Theta, w)^2]dG(w)}},$$

uniformly in $w_i \in [\underline{w}, \bar{w}]$, where Θ is a random variable having the same distribution F as Θ_i.

Proof of Theorem 3. By Lemma 3, the expected payoff for group i under a symmetric profile ϕ tends to a linear function of w_i. Let A^ϕ be the coefficient:

$$\begin{aligned} \lim_{n\to\infty} \sqrt{2\pi n}\pi_i(\phi; n) = & \frac{2w_i\mathbb{E}[\Theta\phi(\Theta)]}{\sqrt{\mathbb{E}[\phi(\Theta)^2]\int_{\underline{w}}^{\bar{w}} w^2 dG(w)}} \\ =&: A^\phi w_i. \end{aligned} \tag{8}$$

For the congressional district profile, remember the definition:

$$\begin{aligned} w_j\phi^{\text{CD}}(\theta_j, w_j) &= c\phi^{\text{WTA}}(\theta_j) + (w_j - c)\,\phi^{\text{PR}}(\theta_j) \\ &= c\operatorname{sgn}(\theta_j) + (w_j - c)\,\theta_j. \end{aligned}$$

We claim that the limit function is affine in w_i:

$$\lim_{n\to\infty} \sqrt{2\pi n}\pi_i(\phi^{\text{CD}}; n) = Bw_i + C. \tag{9}$$

To see that, let us apply Lemma 3 again:

$$\begin{aligned} \lim_{n\to\infty} \sqrt{2\pi n}\pi_i(\phi^{\text{CD}}; n) &= 2 \cdot \frac{w_i\mathbb{E}\left[\Theta\phi^{\text{CD}}(\Theta, w_i)\right]}{\sqrt{\int_{\underline{w}}^{\bar{w}} w^2\mathbb{E}\left[\phi^{\text{CD}}(\Theta, w)^2\right]dG(w)}} \\ &= 2 \cdot \frac{c\mathbb{E}\left[|\Theta|\right] + (w_i - c)\,\mathbb{E}\left[\Theta^2\right]}{\sqrt{\int_{\underline{w}}^{\bar{w}} w^2\mathbb{E}\left[\phi^{\text{CD}}(\Theta, w)^2\right]dG(w)}}. \end{aligned}$$

[21] This is possible since $\phi(\Theta)$ is symmetrically distributed, and since we exclude the trivial case in which $\phi(\Theta) = 0$ almost surely.

Since $|\theta| \geq \theta^2$ with a strict inequality for $0 < |\theta| < 1$, the full support condition for Θ implies $\mathbb{E}[|\Theta|] > \mathbb{E}[\Theta^2]$, and thus the intercept C is positive. The coefficient of w_i is:

$$B = \frac{2\mathbb{E}[\Theta^2]}{\sqrt{\int_{\underline{w}}^{\bar{w}} w^2 \mathbb{E}\left[\phi^{\mathrm{CD}}(\Theta, w)^2\right] dG(w)}}.$$

If $A^\phi < B$, combined with $C > 0$, the right-hand side of (9) is above that of (8). Then, set $w^* = \bar{w}$. If $A^\phi > B$, again combined with $C > 0$, the two limit functions (8) and (9) intersect only once at a positive value $\hat{w}$. Let $w^* = \max\{\underline{w}, \min\{\hat{w}, \bar{w}\}\}$.

Since the convergences (8) and (9) are uniform in w_i, for any $\varepsilon > 0$ there is N with the property stated in Theorem 3. □

Theorem 3 implies that the congressional district profile makes the members of groups with small weights better off, compared with *any* symmetric profile. If the weight is an increasing function of the group size, it means that the congressional district profile is favorable for the members of small groups.

The intuitive reason why the congressional district profile is advantageous for small groups is as follows. Under this profile, the ratio of weights cast by the winner-take-all rule (i.e., c/w_i) is higher for small groups than for large groups. Therefore, the rules used by the smaller groups are relatively close to the dominant strategy, inducing a relative advantage for the small groups. We provide a numerical result in the following subsection using an example of the US Electoral College.

In addition to Theorem 3, we can also show that the congressional district profile distributes payoffs more equally than any symmetric profile does, in the sense of Lorenz dominance. A profile of per capita payoffs for the groups, $\pi = (\pi_1, \cdots, \pi_n)$, is said to *Lorenz dominate* another profile $\pi' = (\pi'_1, \cdots, \pi'_n)$ if the share of payoffs acquired by any bottom fraction of groups is larger in the former profile than in the latter.[22] Lorenz dom-

[22]Formally, for each $x \in [-1, 1]$, let $H_\pi(x)$ be the total population share of those groups whose per capita welfare is not greater than x under the payoff profile π. Then H_π is a distribution function. The *Lorenz curve* of H_π is the graph of the function $\int_0^{H_\pi^{-1}(p)} x dH_\pi(x) / \int_0^1 x dH_\pi(x)$, $0 \leq p \leq 1$, where we define $H_\pi^{-1}(p) = \sup\{x | H_\pi(x) \leq p\}$.

inance, whenever it occurs, agrees with equality comparisons by various inequality indices including the coefficient of variation, the Gini coefficient, the Atkinson index, and the Theil index (see Fields and Fei (1978) and Atkinson (1970)). To see why the congressional district profile is more e-qual than any symmetric profile, recall equations (8) and (9) in the proof of Theorem 3, which assert that when the number of groups is large, the per capita payoff for group i is approximately $A^{\phi} w_i$ for the symmetric profile, and it is approximately $Bw_i + C$ for the congressional district profile. The constant term $C > 0$ for the congressional district profile assures equal additions to all groups' payoffs, which results in a more equal distribution than when there is no such term. More precisely, we can prove the following statement. The proof is relegated to the Appendix.

Theorem 4. *Under Assumptions 1-5, let us consider the payoff profile under the congressional district profile:* $\pi\left(\phi^{\mathrm{CD}}; n\right) = \left(\pi_i\left(\phi^{\mathrm{CD}}; n\right)\right)_{i=1}^{n}$. *Let* ϕ *be any symmetric profile and* $\pi(\phi; n) = (\pi_i(\phi; n))_{i=1}^{n}$ *the payoff profile under* ϕ. *For sufficiently large* n, $\pi\left(\phi^{\mathrm{CD}}; n\right)$ *Lorenz dominates* $\pi(\phi; n)$.

4.4 Computational Results

The results in the previous subsection concern cases with a large number of groups. The question remains as to whether the conclusions obtained there are also valid for a finite number of groups. In this section, we provide a numerical computation result using an example of the US presidential election.

There are 50 states and one federal district. The weights $(w_i)_{i=1}^{51}$ are the numbers of electoral votes assigned in the 2020 election. The first and second columns of Table 2 show the distribution of weights among the states.

We assume IAC* (Impartial Anonymous Culture*): the statewide popular vote margins Θ_i are independent and uniformly distributed on $[-1, 1]$, first introduced by May (1948) and studied thoroughly by, for example, De Mouzon et al. (2019). For any profile ϕ, we can compute the per capita

A payoff profile π *Lorenz dominates* another profile π' if the Lorenz curve of H_π lies above that of $H_{\pi'}$.

payoff for state i via the formula:

$$\pi_i(\phi) = 2\left(0.5^{51}\right)\int_{-1}^{1}\cdots\int_{-1}^{1}\theta_i 1_A(\theta_1,\cdots,\theta_{51})d\theta_1\cdots d\theta_{51} \qquad (10)$$

where $A = \left\{(\theta_1,\cdots,\theta_{51})\middle|\sum_{j=1}^{51} w_j\phi_j(\theta_j) > 0\right\}$.[23]

We consider four distinct profiles: ϕ^{WTA}, ϕ^{PR}, ϕ^a with $a = 102/538$, and ϕ^{CD} with coefficient $c = 2$, which are the winner-take-all profile, the proportional profile, a mixed profile, and a congressional district profile, respectively. The parameter $c = 2$ of the congressional district profile is the number used in Maine and Nebraska, corresponding to the two seats assigned to each state in the Senate. The parameter $a = 102/538$ of the mixed profile is chosen so that the proportion of electoral votes allocated on the winner-take-all basis is the same for all states, and the total number of electoral votes allocated in this way is the same as in the congressional district profile.

We compute (10) under these four profiles by a Monte Carlo simulation with 10^{10} iterations. The results are summarized in Tables 2 and 3. Table 2 shows the per capita payoff ($\pi_i(\phi)$) under the respective profiles. Table 3 shows the ratios of per capita payoff between different profiles ($\pi_i(\phi)/\pi_i(\psi)$). If the ratio is below 1, state i prefers ψ to ϕ.

It follows from Lemma 2 (ii) that as the number n of states increases, the ratios $\pi_i\left(\phi^{\text{WTA}}\right)/\pi_i\left(\phi^{\text{PR}}\right)$ and $\pi_i\left(\phi^a\right)/\pi_i\left(\phi^{\text{PR}}\right)$ converge to the respective correlations $\text{Corr}[\Theta, \phi^{\text{WTA}}(\Theta)] \approx 0.866$ and $\text{Corr}[\Theta, \phi^a(\Theta)] \approx 0.989$, where the values are computed for Θ uniformly distributed on $[-1, 1]$. Table 3 indicates that for the present example with 50 states plus DC, these ratios are indeed close to the respective correlations, which suggests that convergence of the π-ratios is fairly quick. In particular, as expected by Theorem 2, the proportional profile Pareto dominates the winner-take-all profile in the present case. As suggested by Proposition 4, all states prefer the mixed profile ϕ^a to the winner-take-all profile, and the proportional profile to ϕ^a.

The ratios $\pi_i(\phi^{\text{CD}})/\pi_i(\phi^{\text{PR}})$ in Table 3 are consistent with the result in Theorem 3. Small states prefer the congressional district profile to the

[23] It is easy to check that under the uniform distribution assumption, (10) is equivalent to the expression in Lemma 2 (i).

proportional one.

In addition, the values of $\pi_i(\phi^{\mathrm{CD}})/\pi_i(\phi^{\mathrm{WTA}})$ in the table show that the winner-take-all profile is Pareto dominated by the congressional district profile, and the welfare improvement by switching to the congressional district profile is greater for small states than for large states in terms of the ratio.

All of our numerical observations suggest the sensibility of the asymptotic results obtained in Section 4 in the example of the US Electoral College.

Table 2: Estimated payoffs in the US presidential election, based on the apportionment in 2016, via Monte Carlo simulation with 10^{10} iterations. The estimated standard errors are in the range between 3.9 and 4.1×10^{-6}.

electoral votes	number of states	$\pi(\phi^{\mathrm{WTA}})$	$\pi(\phi^{\mathrm{PR}})$	$\pi(\phi^{a})$	$\pi(\phi^{\mathrm{CD}})$
3	8	0.0113	0.0133	0.0130	0.0167
4	5	0.0151	0.0177	0.0174	0.0209
5	3	0.0189	0.0221	0.0217	0.0251
6	6	0.0226	0.0266	0.0261	0.0293
7	3	0.0264	0.0310	0.0305	0.0335
8	2	0.0302	0.0354	0.0348	0.0377
9	3	0.0340	0.0399	0.0392	0.0419
10	4	0.0378	0.0443	0.0436	0.0461
11	4	0.0416	0.0488	0.0479	0.0503
12	1	0.0454	0.0532	0.0523	0.0545
13	1	0.0492	0.0577	0.0567	0.0587
14	1	0.0531	0.0622	0.0611	0.0630
15	1	0.0569	0.0666	0.0655	0.0672
16	2	0.0607	0.0711	0.0699	0.0715
18	1	0.0684	0.0801	0.0788	0.0800
20	2	0.0762	0.0891	0.0877	0.0885
29	2	0.1120	0.1303	0.1284	0.1275
38	1	0.1494	0.1729	0.1706	0.1677
55	1	0.2356	0.2614	0.2615	0.2507

Table 3: Ratios between payoffs.

electoral votes	number of states	$\frac{\pi(\phi^{\mathrm{WTA}})}{\pi(\phi^{\mathrm{PR}})}$	$\frac{\pi(\phi^{a})}{\pi(\phi^{\mathrm{PR}})}$	$\frac{\pi(\phi^{\mathrm{CD}})}{\pi(\phi^{\mathrm{PR}})}$	$\frac{\pi(\phi^{\mathrm{CD}})}{\pi(\phi^{\mathrm{WTA}})}$
3	8	0.852	0.982	1.260	1.479
4	5	0.852	0.982	1.182	1.387
5	3	0.852	0.982	1.134	1.331
6	6	0.852	0.982	1.103	1.294
7	3	0.852	0.982	1.080	1.268
8	2	0.852	0.982	1.064	1.248
9	3	0.852	0.982	1.050	1.232
10	4	0.853	0.983	1.040	1.220
11	4	0.853	0.983	1.031	1.210
12	1	0.853	0.983	1.024	1.201
13	1	0.853	0.983	1.018	1.194
14	1	0.853	0.983	1.013	1.187
15	1	0.854	0.983	1.009	1.181
16	2	0.854	0.983	1.005	1.177
18	1	0.854	0.983	0.998	1.168
20	2	0.855	0.983	0.993	1.161
29	2	0.859	0.985	0.978	1.138
38	1	0.864	0.987	0.970	1.122
55	1	0.901	1.000	0.959	1.064

5 Concluding Remarks

This paper shows that the decentralized choice of the weight allocation rule in representative democracy constitutes a Prisoner's Dilemma: the winner-take-all rule is a dominant strategy for each group, whereas the Nash equilibrium is Pareto dominated. Each group has an incentive to put its entire weight on the alternative supported by the majority of its members in order to reflect their preferences in the social decision, although such a distortion by each group prevents efficient aggregation of the preferences of the society as a whole.

We also develop an asymptotic technique and show that the proportional rule Pareto dominates every other symmetric profile when the number of the groups is sufficiently large.

Our model may provide explanations for the phenomena that we observe in existing institutions of collective decision making. In the United States Electoral College, the rule used by the states varied in early elections until it converged by 1832 to the winner-take-all rule, which has remained dominantly employed by nearly all states since then. In many parliamentary voting situations, we often observe parties and/or factions forcing their members to align their votes in order to maximally reflect their preferences in the legislative decision, although some members may disagree with the party's alignment. The voting outcome obtained by the winner-take-all rule may fail to efficiently aggregate preferences, as observed in the discrepancy between the electoral result and the national popular vote winner in the US presidential elections in 2000 and 2016. Party discipline or factional voting may also cause welfare loss when each group pushes their votes maximally toward their ideological goals, failing to reflect all of their members' preferences in the legislative decision.

The Winner-Take-All Dilemma tells us that the society should call for some device other than each group's unilateral effort, in order to obtain a socially preferable outcome. As we see in the failure of various attempts to modify or abolish the winner-take-all rule, such as the ballot initiative for an amendment to the State Constitution in Colorado in 2004, each state has no incentive to unilaterally deviate from the equilibrium. The National Popular Vote Interstate Compact is a well-suited example of a coordination device (Koza et al. (2013)). As it comes into effect only when the number of electoral votes attains the majority, each state does not suffer from the payoff loss by a unilateral (or coalitional) deviation until a sufficient level of coordination is attained. The emergence of such an attempt is coherent with the insights obtained in this paper that the game is a Prisoner's Dilemma, and a coordination device is necessary for a Pareto improvement.

We have assumed that social decisions are binary. There are situations where this assumption may not fit. In the US presidential elections, third-party or independent candidates can, and do, have a non-negligible impact on the election outcome. It is not clear how the presence of such candidates alters the comparison of rules to allocate electoral votes. When the model is applied to legislative voting, the assumption of a binary decision might be justified on the grounds that choices are made between the status quo and a proposal. However, such an argument abstracts away the political process that gives rise to the particular choice of the proposal. Further analysis is necessary for the cases with more than two alternatives and is beyond the scope of this paper.

Appendix

(For Online Publication)

A1 Proof of Lemma 1

(i) Let D be the set of all SCFs, and $\pi(d) = \{(\pi_i(d))_{i=1}^n | d \in D\}$ the set of (ex ante) payoff vectors generated by SCFs. Then $\pi(d)$ is convex.[24] Let $\mathrm{Pa}(\pi(D))$ be the Pareto frontier of $\pi(D)$, i.e., the set of payoff profiles $u \in \pi(D)$ for which there is no $d' \in \pi(D)$ such that $u'_i \geq u_i$ for all i and $u'_i > u_i$ for some i.

We divide the proof of (i) into two steps.

Step 1. *Let $\lambda \in \mathbb{R}^n_+ \setminus \{\boldsymbol{0}\}$. Then the unique solution to the following maximization problem (11) is the payoff vector $u^\lambda := \left(\pi_i(d^\lambda)\right)_{i=1}^n$ under a cardinal λ-weighted majority rule d^λ.*[25]

$$\max_{u \in \pi(D)} \sum_{i=1}^n \lambda_i u_i. \tag{11}$$

Moreover, an SCF d satisfies $(\pi_i(d))_{i=1}^n = u^\lambda$ if and only if d is a λ-weighted majority rule.

Let $d \in D$ be any SCF. Then

$$\sum_{i=1}^n \lambda_i \pi_i(d) = \sum_{i=1}^n \lambda_i \mathbb{E}\left[\Theta_i d(\Theta)\right] = \mathbb{E}\left[d(\Theta) \sum_{i=1}^n \lambda_i \Theta_i\right]. \tag{12}$$

Since Θ is absolutely continuous, and so $\sum_{i=1}^n \lambda_i \Theta_i \neq 0$ almost surely, d

[24] This is because for any two SCFs d and d', any convex combination of the payoff vectors corresponding to d and d' can be realized as a compound SCF that randomizes between d and d'.

[25] Recall that a weighted majority rule with a given weight vector is unique only up to differences on a set of measure zero, inducing the same payoffs.

maximizes (12) if and only if $d(\Theta) = \text{sgn} \sum_{i=1}^{n} \lambda_i \Theta_i$ almost surely. That is,

$$d \text{ maximizes (12)} \Leftrightarrow d \text{ is a } \lambda\text{-weighted majority rule.} \tag{13}$$

This implies the first sentence of Step 1. Result (13) also implies that if d is not a λ-weighted majority rule, then $\pi_i(d) \neq \pi_i(d^\lambda)$ for some i, which proves the "only if" part of the second sentence of Step 1. The "if" part is trivial.

Step 2. *A payoff vector $u \in \pi(D)$ is in the Pareto frontier* $\text{Pa}(\pi(D))$ *if and only if there exists $\lambda \in \mathbb{R}_+^n \setminus \{\boldsymbol{0}\}$ such that $u = (\pi_i(d^\lambda))_{i=1}^n =: u^\lambda$, where d^λ is a λ-weighted majority rule.*

Since $\pi(D)$ is convex, we can apply Mas-Colell et al. (1995, Proposition 16.E.2) to show the "only if" part of Step 2.

To show the "if" part, suppose on the contrary that $u^\lambda \notin \text{Pa}(\pi(D))$ for some $\lambda \in \mathbb{R}_+^n \setminus \{\mathbf{0}\}$. Then there exists $u \in \pi(D)$ such that $u \neq u^\lambda$ and $u_i \geq u_i^\lambda$ for all i. Then $\sum_{i=1}^n \lambda_i u_i \geq \sum_{i=1}^n \lambda_i u_i^\lambda$. This contradicts the fact that u^λ is the unique solution to problem (11).

(ii) This follows from the trivial fact that the set of SCFs d_ϕ induced by profiles ϕ that are equivalent to a generalized proportional profile coincides with the set of all weighted majority rules.[26] □

A2 Proof of Part (i) of Lemma 2

We prove the statement for group 1. Let $\pi_1(\phi; n|\theta_1)$ be the conditional expected payoff for group 1 given that the group-wide margin is $\Theta_1 = \theta_1$, which by (3) is:

$$\pi_1(\phi; n|\theta_1) = \theta_1(\mathbb{P}\{w_1\phi(\theta_1) + S_{\phi_{-1}} > 0\} - \mathbb{P}\{w_1\phi(\theta_1) + S_{\phi_{-1}} < 0\}).$$

Since $S_{\phi_{-1}}$ is symmetrically distributed, the second probability can be written as $\mathbb{P}\{-w_1\phi(\theta_1) + S_{\phi_{-1}} > 0\}$. Thus, for $\theta_1 \in [0, 1]$, the above expression

[26]Indeed, if ϕ is equivalent to a generalized proportional profile with the vector of coefficients $\lambda \in [0,1]^n \setminus \{\mathbf{0}\}$, the induced SCF d_ϕ is a μ-weighted majority rule, where the weights are defined by $\mu_i := w_i\lambda_i$; conversely, if d is a μ-weighted majority rule, then $d = d_\phi$ for some profile ϕ that is equivalent to the generalized proportional profile with coefficients $\lambda_i := \frac{\mu_i}{w_i}$.

equals

$$\pi_1(\phi; n|\theta_1) = \theta_1 \mathbb{P}\{-w_1\phi(\theta_1) < S_{\phi_{-1}} \leq w_1\phi(\theta_1)\}.$$

By symmetry, twice the integral of this expression over $\theta_1 \in [0, 1]$ (instead of $[-1, 1]$) equals the unconditional expected payoff $\pi_1(\phi; n)$, which proves part (i) of Lemma 2. □

A3 Local Limit Theorem

We quote a version of the Local Limit Theorem shown in Mineka and Silverman (1970). We will use it in the proof of part (ii) of Lemma 2.

LLT. *(Mineka and Silverman (1970, Theorem 1)) Let (X_i) be a sequence of independent random variables with mean 0 and variances $0 < \sigma_i^2 < \infty$. Write F_i for the distribution of X_i. Write also $S_n = \sum_{i=1}^n X_i$ and $s_n^2 = \sum_{i=1}^n \sigma_i^2$. Suppose the sequence (X_i) satisfies the following conditions:*

(α) There exists $\bar{x} > 0$ and $c > 0$ such that for all i,

$$\frac{1}{\sigma_i^2} \int_{|x|<\bar{x}} x^2 dF_i(x) > c.$$

(β) Define the set

$$A(t, \varepsilon) = \{x| \ |x| < \bar{x} \text{ and } |xt - \pi m| > \varepsilon \text{ for all integers } m \text{ with } |m| < \bar{x}\}.$$

Then, for some bounded sequence (a_i) such that $\inf_i \mathbb{P}\{|X_i - a_i| < \delta\} > 0$ for all $\delta > 0$, and for any $t \neq 0$, there exists $\varepsilon > 0$ such that

$$\frac{1}{\log s_n} \sum_{i=1}^n \mathbb{P}\{X_i - a_i \in A(t, \varepsilon)\} \to \infty.$$

(γ) (Lindeberg's condition.) For any $\varepsilon > 0$,

$$\frac{1}{s_n^2} \sum_{i=1}^n \int_{|x|/s_n > \varepsilon} x^2 dF_i(x) \to 0.$$

Under conditions (α)-(γ), if $s_n^2 \to \infty$, we have

$$\sqrt{2\pi s_n^2}\mathbb{P}\{S_n \in (a,b]\} \to b - a.^{27} \tag{14}$$

A4 Proof of Lemma 3

Preliminaries. We prove the lemma for group 1. In the proof, we use the notation of LLT. Let

$$X_i := w_i \phi(\Theta_i, w_i),\ i = 1, 2, \cdots,$$

and $S_n := \sum_{i=1}^n X_i$. Then X_i has mean 0 and variance $\sigma_i^2 := w_i^2 \mathbb{E}[\phi(\Theta, w_i)^2]$, and so the partial sum of variances is $s_n^2 := \sum_{i=1}^n w_i^2 \mathbb{E}[\phi(\Theta, w_i)^2]$, where Θ represents a random variable that has the same distribution F as Θ_i.

Define the event

$$\Omega_n(\theta_1, w_1) = \left\{ -w_1\phi(\theta_1, w_1) < \sum_{i=2}^n X_i \le w_1\phi(\theta_1, w_1) \right\}.$$

We divide the proof into several claims. Claims 5.1-5.3 show that the sequence (X_i) defined above satisfies the conditions of the Local Limit Theorem (LLT) in Section A4. Claim 5.4 applies the LLT to complete the proof of Lemma 3.

Claim 5.1. $\frac{s_n^2}{n} \to \int_{\underline{w}}^{\bar{w}} w^2 \mathbb{E}[\phi(\Theta, w)^2] dG(w)$.

Proof of Claim 5.1. This holds since sequence (σ_i^2) is bounded and the statistical distribution G_n induced by $(w_i)_{i=1}^n$ converges weakly to G. □

Claim 5.2. *Conditions (α) and (γ) in the LLT hold.*

Proof of Claim 5.2. This immediately follows from the fact that sequence (X_i) is bounded and $s_n^2 \to \infty$. In particular, it is enough to define $\bar{x}$ to be any finite number greater than $\bar{w}$. □

[27] The original conclusion of Theorem 1 in Mineka and Silverman (1970) is stated in terms of the open interval (a, b). Applying the theorem to $(a, b+c)$ and $(b, b+c)$ and then taking the difference gives the result for $(a, b]$. In addition, the original statement allows for cases where s_n^2 does not go to infinity, and also mentions uniform convergence. These considerations are not necessary for our purpose, so we omit them.

Claim 5.3. *Condition (β) in LLT holds.*

Proof of Claim 5.3. Recall that ϕ has the form

$$w_i\phi(\theta_i, w_i) = h_1(w_i)h_2(\theta_i) + h_3(w_i)\operatorname{sgn}\theta_i.$$

Let $a_i = h_3(w_i)$. We first check that the sequence (a_i) satisfies the requirements in condition (β). First, (a_i) is bounded since h_3 is bounded. Now, for any i and any $\delta > 0$,

$$\begin{aligned}\mathbb{P}\{|X_i - a_i| < \delta\} &\geq \mathbb{P}\{|X_i - a_i| < \delta \text{ and } \Theta_i > 0\}\\ &\geq \mathbb{P}\left\{|w_i\phi(\Theta_i, w_i) - h_3(w_i)\operatorname{sgn}\Theta_i| < \delta \text{ and } \Theta_i > 0\right\}\\ &= \mathbb{P}\{|h_1(w_i)h_2(\Theta_i)| < \delta \text{ and } \Theta_i > 0\}.\end{aligned}$$

Letting $\bar{h}_1 > 0$ be an upper bound of $|h_1|$ and Θ a random variable distributed as Θ_i, the last expression has the following lower bound independent of i:

$$\mathbb{P}\{|h_2(\Theta)| < \delta/\bar{h}_1 \text{ and } \Theta > 0\} > 0,$$

which is positive by the assumptions on h_2 and on the distribution of Θ.

Next we check the limit condition in (β). Recall that $A(t, \varepsilon)$ is the union of intervals

$$\left(\frac{\pi m + \varepsilon}{|t|}, \frac{\pi(m+1) - \varepsilon}{|t|}\right), \; m = 0, \pm 1, \pm 2, \cdots,$$

restricted to $(-\bar{x}, \bar{x})$, where we can choose $\bar{x}$ to be any number greater than $\bar{w}$. To prove the limit condition in (β), it therefore suffices to verify that one such interval contains $X_i - a_i$ with probability bounded away from zero, for all groups i in some sufficiently large subset of groups. To do this, note that if $\Theta_i < 0$, then $X_i - a_i = h_1(w_i)h_2(\Theta_i) - 2h_3(w_i)$. The assumptions on h_2 and on the distribution of Θ imply that for any $\eta > 0$, there exists a set $O_\eta \subset [-1, 0]$ with $\mathbb{P}\{\Theta \in O_\eta\} > 0$ such that if $\Theta \in O_\eta$ then $|h_2(\Theta)| \leq \eta$. Therefore,

$$\Theta_i \in O_\eta \Rightarrow X_i - a_i \in T_{w_i,\eta},$$

where

$$T_{w_i,\eta} := [-2h_3(w_i) - \eta h_1(w_i), \; -2h_3(w_i) + \eta h_1(w_i)].$$

Since h_1 is bounded, we can make $T_{w_i,\eta}$ an arbitrarily small interval around $-2h_3(w_i)$ by letting $\eta > 0$ be sufficiently small. Moreover, since h_3 is continuous and not a constant, we can find a sufficiently small interval $[\underline{v}, \bar{v}] \subset [\underline{w}, \bar{w}]$ with $\underline{v} < \bar{v}$ such that if $w_i \in [\underline{v}, \bar{v}]$, then $-2h_3(w_i)$ is between, and bounded away from, $\frac{\pi m}{|t|}$ and $\frac{\pi(m+1)}{|t|}$ for some integer m. Fix such an interval $[\underline{v}, \bar{v}]$ and define

$$I := \{i | w_i \in [\underline{v}, \bar{v}]\}.$$

Then, for sufficiently small $\eta > 0$ and $\varepsilon > 0$, we have $T_{w_i,\eta} \subset A(t, \varepsilon)$ for all $i \in I$. Fixing such $\eta > 0$ and $\varepsilon > 0$, it follows that

$$\Theta_i \in O_\eta \text{ and } i \in I \Rightarrow X_i - a_i \in A(t, \varepsilon).$$

This implies that

$$\mathbb{P}\{X_i - a_i \in A(t, \varepsilon)\} \geq \mathbb{P}\{\Theta \in O_\eta\} =: p > 0 \text{ for all } i \in I,$$

and hence

$$\frac{1}{\log s_n} \sum_{i=1}^{n} \mathbb{P}\{X_i - a_i \in A(t, \varepsilon)\} \geq \frac{n}{\log s_n} \cdot \frac{\sharp\{i \in I | i \leq n\}}{n} \cdot p.$$

As $n \to \infty$, the first factor on the right-hand side tends to ∞ since s_n has an asymptotic order of $\sqrt{n}$. The second factor tends to $G(\bar{v}) - G(\underline{v}) > 0$, which is positive since G has full support on $[\underline{w}, \bar{w}]$. Therefore the left-hand side tends to ∞. $\square$

Claim 5.4. *As $n \to \infty$, uniformly in $w_1 \in [\underline{w}, \bar{w}]$,*

$$2 \int_0^1 \theta_1 \sqrt{2\pi n} \mathbb{P}\{\Omega_n(\theta_1, w_1)\} dF(\theta_1) \to \frac{2w_1 \mathbb{E}[\Theta \phi(\Theta, w_1)]}{\sqrt{\int_{\underline{w}}^{\bar{w}} w^2 \mathbb{E}[\phi(\Theta, w)^2] dG(w)}}. \tag{15}$$

By part (i) of Lemma 2,[28] *the left-hand side of (15) is $\sqrt{2\pi n}\pi_i(\phi; n)$, and therefore Lemma 3 holds.*

Proof of Claim 5.4. By Claims 5.2 and 5.3, we may apply the LLT to ob-

[28] It is easy to check that part (i) of Lemma 2 holds for rules $\phi(\cdot, w_i)$ that depend on weight w_i as well.

tain

$$\sqrt{2\pi s_n^2}\mathbb{P}\{\Omega_n(\theta_1, w_1)\} \to 2w_1\phi(\theta_1, w_1).$$

By Claim 5.1, this means that

$$\sqrt{2\pi n}\theta_1\mathbb{P}\{\Omega_n(\theta_1, w_1)\} \to \frac{2w_1\theta_1\phi(\theta_1, w_1)}{\sqrt{\int_{\underline{w}}^{\bar{w}} w^2\mathbb{E}[\phi(\Theta, w)^2]dG(w)}}. \tag{16}$$

Letting $\theta_1 = 1$ maximizes the left-hand side of (16) with the maximum value $\sqrt{2\pi n}\mathbb{P}\{\Omega_n(1, w_1)\}$. This maximum value itself converges to a finite limit. Hence the expression $\sqrt{2\pi n}\theta_1\mathbb{P}\{\Omega_n(\theta_1, w_1)\}$ is uniformly bounded for all n and $\theta_1 \in [0, 1]$. By the Bounded Convergence Theorem,

$$2\int_0^1 \theta_1\sqrt{2\pi n}\mathbb{P}\{\Omega_n(\theta_1, w_1)\}dF(\theta_1) \to 2 \cdot \frac{2w_1\int_0^1 \theta_1\phi(\theta_1, w_1)dF(\theta_1)}{\sqrt{\int_{\underline{w}}^{\bar{w}} w^2\mathbb{E}[\phi(\Theta, w)^2]dG(w)}}.$$

Since F is symmetric and ϕ is odd, this limit is exactly the one in (15).

To check the uniform convergence, note that for each n, the integral on the left-hand side of (15) is non-decreasing in w_1, since event $\Omega_n(\theta_1, w_1)$ weakly expands as w_1 increases.[29] We have shown that this integral converges pointwise to a limit that is proportional to the factor $w_1\mathbb{E}[\Theta\phi(\Theta, w_1)]$, which is continuous in w_1.[30] Therefore, the convergence in (15) is uniform in $w_1 \in [\underline{w}, \bar{w}]$.[31] □

A5 Proof of Part (ii) of Lemma 2

This follows immediately from Lemma 3, by noting that if ϕ is a symmetric profile, each group's rule can be written as $\phi(\theta_j, w_j) = \phi(\theta_j)$. □

[29] Let $\theta_1 \in [0, 1]$. If ϕ is a symmetric profile, i.e., if $\phi(\theta_1, w_1) = \phi(\theta_1)$, then $w_1\phi(\theta_1)$ is non-decreasing in w_1. If $\phi = \phi^{\text{CD}}$, then $w_1\phi^{\text{CD}}(\theta_1, w_1) = c\,\text{sgn}(\theta_1) + (w_1 - c)\theta_1$, which is non-decreasing in w_1 again. Thus event $\Omega_n(\theta_1, w_1)$ weakly expands as w_1 increases.

[30] If ϕ is a symmetric profile, this factor is linear in w_i. If $\phi = \phi^{\text{CD}}$, the factor equals $c\mathbb{E}(|\Theta|) + (w_i - c)\mathbb{E}(\Theta^2)$, which is affine in w_i.

[31] It is known that if (f_n) is a sequence of non-decreasing functions on a fixed finite interval and f_n converges pointwise to a continuous function, then the convergence is uniform. See Buchanan and Hildebrandt (1908).

A6 Proof of Proposition 2

By part (ii) of Lemma 2, we must show that $\mathrm{Corr}\,[\Theta, \phi^a(\Theta)]$ is decreasing in $a \in [0,1]$. By simple calculation,

$$\mathbb{E}(\Theta^2) \cdot \mathrm{Corr}\,[\Theta, \phi^a(\Theta)]^2 = \frac{a\mathbb{E}(|\Theta|) + (1-a)\mathbb{E}(\Theta^2)}{a^2 + 2a(1-a)\mathbb{E}(|\Theta|) + (1-a)^2\mathbb{E}(\Theta^2)}.$$

The derivative of this expression with respect to a has the same sign as

$$\left\{\tfrac{d}{da}(a\mathbb{E}(|\Theta|) + (1-a)\mathbb{E}(\Theta^2))^2\right\}\Big(a^2 + 2a(1-a)\mathbb{E}(|\Theta|) + (1-a)^2\mathbb{E}(\Theta^2)\Big)$$
$$- \Big(a\mathbb{E}(|\Theta|) + (1-a)\mathbb{E}(\Theta^2)\Big)^2\left\{\tfrac{d}{da}(a^2 + 2a(1-a)\mathbb{E}(|\Theta|) + (1-a)^2\mathbb{E}(\Theta^2))\right\}$$
$$= a(a\mathbb{E}(|\Theta|) + (1-a)\mathbb{E}(\Theta^2))(\mathbb{E}(|\Theta|)^2 - \mathbb{E}(\Theta^2)).$$

This is negative for any $a \in (0,1]$, since $\mathbb{E}(|\Theta|)^2 \leq \mathbb{E}(\Theta^2)$ in general, and the full-support assumption implies that this holds with strict inequality. □

A7 Proof of Theorem 4

Clearly, Lorenz dominance is invariant to linear transformations of payoffs. Thus, it suffices to prove that for large enough n, the payoff profile $\sqrt{2\pi n}\pi(\phi^{\mathrm{CD}}; n)$ Lorenz dominates the payoff profile $\sqrt{2\pi n}\pi(\phi; n)$. By equations (8) and (9) in the proof of Theorem 3, as $n \to \infty$ these amounts converge to $Bw_i + C$ and $A^\phi w_i$, respectively. A result by Moyes (1994, Proposition 2.3) implies that if f and g are continuous, nondecreasing, and positive-valued functions such that $f(w_i)/g(w_i)$ is decreasing in w_i, then the distribution of $f(w_i)$ Lorenz dominates that of $g(w_i)$. The ratio $(Bw_i+C)/(A^\phi w_i)$ is decreasing in w_i, and so the claimed Lorenz dominance holds in the limit as $n \to \infty$. Recalling that the convergences are uniform, the dominance holds for sufficiently large n. □

References

Atkinson, A. B. (1970): "On the measurement of inequality," *Journal of Economic Theory*, 2, 244–263.

Azrieli, Y. and S. Kim (2014): "PARETO EFFICIENCY AND WEIGHTED MAJORITY RULES," *International Economic Review*, 55, 1067–1088.

Banzhaf, J. F. (1965): "Weighted Voting Doesn't Work: A Mathematical Analysis," *Rutgers Law Review*, 19, 317–343.

——— (1968): "One man, 3.312 votes: A mathematical analysis of the Electoral College," *Villanova Law Review*, 13, 304–332.

Barberà, S. and M. O. Jackson (2006): "On the weights of nations: Assigning voting weights in a heterogeneous union," *Journal of Political Economy*, 114, 317–339.

Beisbart, C. and L. Bovens (2008): "A power analysis of Amendment 36 in Colorado," *Public Choice*, 134, 231–246.

Börgers, T. and P. Postl (2009): "Efficient compromising," *Journal of Economic Theory*, 144, 2057–2076.

Brams, S. J. and M. D. Davis (1974): "The 3/2's Rule in Presidential Campaigning," *Americal Political Science Review*, 68, 113–134.

Buchanan, H. and T. Hildebrandt (1908): "Note on Convergence of a Sequence of Functions of a Certain Type," *Annals of Mathematics*, 9, 123–126.

Bugh, G., ed. (2010): *Electoral College Reform: Challenges and Possibilities*, Ashgate Publishing.

Clolé ry, H. and Y. Koriyama (2020): "Trapped by the Prisoner's Dilemma, the United States Presidential Election Needs a Coordination Device," Institut des Politiques Publiques, Policy Brief 60, HAL.

De Mouzon, O., T. Laurent, M. Le Breton, and D. Lepelley (2019): "Exploring the effects of national and regional popular vote interstate compact on a toy symmetric version of the Electoral College: an electoral engineering perspective," *Public Choice*, 179, 51–95.

Edwards, G. C. (2004): *Why the Electoral College is Bad for America*, Yale University Press.

Eguia, J. X. (2011a): "Voting blocs, party discipline and party formation," *Games and Economic Behavior*, 73, 111–135.

——— (2011b): "Endogenous parties in an assembly," *American Journal of Political Science*, 55, 16–26.

Ehlers, L., D. Majumdar, D. Mishra, and A. Sen (2020): "Continuity and incentive compatibility in cardinal mechanisms," *Journal of Mathematical Economics*, 88, 31–41.

Feix, M. R., D. Lepelley, V. R. Merlin, and J.-L. Rouet (2004): "The probability of conflicts in a U.S. presidential type election," *Economic Theory*, 23, 227–257.

Felsenthal, D. S. and M. Machover (1998): *The measurement of voting power*, Northampton: Edward Edgar.

Fields, G. S. and J. C. H. Fei (1978): "On inequality comparisons," *Econometrica*, 46, 303–316.

Gelman, A. (2003): "Forming Voting Blocs and Coalitions as a Prisoner's Dilemma: A Possible Theoretical Explanation for Political Instability," *The B.E. Journal of Economic Analysis and Policy*, 2, 1–16.

Hummel, P. (2011): "Proportional versus winner-take-all electoral vote allocations," *Public Choice*, 148, 381–393.

Kartal, M. (2015): "A Comparative Welfare Analysis of Electoral Systems with Endogenous Turnout," *Economic Journal*, 125, 1369–1392.

Koriyama, Y., J.-F. Laslier, A. Macé, and R. Treibich (2013): "Optimal Apportionment," *Journal of Political Economy*, 121, 584–608.

Koza, J. R., B. F. Fadem, M. Grueskin, M. S. Mandell, R. Richie, and J. F. Zimmerman (2013): *Every Vote Equal: A State-Based Plan for Electing the President by National Popular Vote*, National Popular Vote Press, 4th ed.

KURZ, S., N. MAASER, AND S. NAPEL (2017): "On the Democratic Weights of Nations," *Journal of Political Economy*, 125, 1599–1634.

LARUELLE, A. AND F. VALENCIANO (2008): *Voting and Collective Decision-Making*, Cambridge University Press.

LE BRETON, M., D. LEPELLEY, AND H. SMAOUI (2016): "Correlation, partitioning and the probability of casting a decisive vote under the majority rule," *Journal of Mathematical Economics*, 64, 11–22.

LINDBECK, A. AND J. W. WEIBULL (1987): "Balanced-budget redistribution as the outcome of political competition," *Public Choice*, 52, 273–297.

MAS-COLELL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic Theory*, Oxford University Press.

MAY, K. (1948): "Probabilities of Certain Election Results," *American Mathematical Monthly*, 55, 203–209.

MINEKA, J. AND S. SILVERMAN (1970): "A Local Limit Theorem and Recurrence Conditions for Sums of Independent and Non-Lattice Random Variables," *Annals of Probability*, 41, 592–600.

MOYES, P. (1994): "Inequality reducing and inequality preserving transformations of incomes: symmetric and individualistic transformations," *Journal of Economic Theory*, 63, 271–298.

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

RAE, D. W. (1946): "Decision-Rules and Individual Values in Constitutional Choice," *American Political Science Review*, 63, 40–56.

SHAPLEY, L. S. AND M. SHUBIK (1954): "A method for evaluating the distribution of power in a committee system," *American Political Science Review*, 48, 787–792.

STRÖMBERG, D. (2008): "How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida," *American Economic Review*, 98, 769–807.

Wegman, J. (2020): *Let the People Pick the President*, St. Martin's Press.