Forming Voting Blocs and Coalitions as a Prisoner's Dilemma: A Possible Theoretical Explanation for Political Instability

Andrew Gelman*

*Department of Statistics, Columbia University, gelman@stat.columbia.edu

Forming Voting Blocs and Coalitions as a Prisoner's Dilemma: A Possible Theoretical Explanation for Political Instability[*]

Andrew Gelman

Abstract

Individuals in a committee or election can increase their voting power by forming coalitions. This behavior is shown here to yield a prisoner's dilemma, in which a subset of voters can increase their power, while reducing average voting power for the electorate as a whole. This is an unusual form of the prisoner's dilemma in that cooperation is the selfish act that hurts the larger group. Under a simple model, the privately optimal coalition size is approximately 1.4 times the square root of the number of voters. When voters' preferences are allowed to differ, coalitions form only if voters are approximately politically balanced. We propose a dynamic view of coalitions, in which groups of voters choose of their own free will to form and disband coalitions, in a continuing struggle to maintain their voting power. This is potentially an endogenous mechanism for political instability, even in a world where individuals' (probabilistic) preferences are fixed and known.

KEYWORDS: coalitions, cooperation, decisive vote, elections, legislatures, prisoner's dilemma, voting power

---

1   Introduction

Voting power is generally defined in terms of the possibilities that a given voter, or set of voters, can affect the outcome of an election. Ideas of voting power have been applied in a range of settings, including committee voting in legislatures, weighted voting (as in corporations), and hierarchical voting systems, such as the U.S. Electoral College and the European council of ministers.[1] In this paper, we consider costs and benefits, in terms of voting power, of forming coalitions within committees and legislatures.

The main results presented in this paper are: first, we show a connection between voting power and the formation of coalitions, which leads to a prisoner's dilemma in which coalitions are locally beneficial but have negative results for the legislature as a whole. Second, for the random voting model (in which all combinations of votes are counted equally), we derive results for optimal coalition sizes and the effects on voting power of various coalition structures. Third, we explore the benefits of coalition formation under the more general model in which voters can have different preferences.

Coalitions in legislatures and committees come in a wide range of sizes and durations, ranging from short-term vote trading, through agreements on a particular issue or set of issues, to more permanent factions and political parties. In multiparty systems, political parties themselves can split into factions or individual members or, conversely, form super-coalitions with other parties.

This paper builds a stylized model in which unstable and potentially nested coalitions arise naturally from the mathematics of voting power. Put briefly, if a group of legislators form a coalition, this increases their probability of affecting the outcome of the total vote. Similarly, coalitions themselves can ally to increase their voting power still further. However, forming a coalition decreases the voting power of those outside the group—enough so that the average voting power decreases for the legislature as a whole. As a result, there is always a motivation for out-group members to form coalitions of their own and to peel off members from existing coalitions. Paradoxically, the effectiveness of coalitions leads to their ultimate instability.

Because our results center on the voluntary formation of coalitions, we consider the framework of a legislative body or committee with n members that are free to group themselves into voting blocs. Each bloc represents a formal or informal binding coalition, so that all the votes for a bloc are assigned to the choice that wins a majority of the vote within the bloc. We assume two choices and no abstention throughout. The blocs can themselves be nested.

2   Definitions of voting power

Voting power can and has been defined in a variety of ways (see Saari and Sieberg, 1999, and Heard and Swartz, 1999); in this paper we shall use the definition based on the probability that an individual's vote affects the final decision of the legislature or committee.

Consider a committee with n voters. We use the notation i for an individual voter, $v_i = \pm 1$ for his or her vote, $v = (v_1, \ldots, v_n)$ for the entire vector of votes, and $R = R(v) = \pm 1$

---

[1]See, e.g., Shapley and Shubik (1954), Banzhaf (1965), Brams and Davis (1974), Colantoni, Levesque, and Ordeshook (1975), Felsenthal and Machover (1998), and Snyder, Ting, and Ansolabehere (2001), for various perspectives from political science and Heard and Swartz (1999) and Gelman, Katz, and Tuerlinckx (2002) for statistical reviews.

for the rule that aggregates the n individual votes into a single outcome. It is possible for R to be stochastic (because of possible ties or, even more generally, because of possible errors in vote counting). We shall assume various distributions on v and rules $R(v)$ which then together induce distributions on R.

The probability that the change of the individual vote v_i will change the outcome is,

$$\text{power}_i = \text{Power of voter } i = \Pr(R = +1|v_i = +1) - \Pr(R = +1|v_i = -1). \qquad (1)$$

If your voting power is zero, then changing your vote from -1 to $+1$ has no effect. We assume that tied votes at all levels are decided by coin flips.

A related measure that is sometimes considered is the probability of being satisfied with the outcome (Straffin, 1978, Heard and Swartz, 1999):

$$
\begin{aligned}
\Pr(\text{satisfied}) &= \text{Probability that voter } i \text{ is satisfied} \\
&= \Pr(R = v_i) \\
&= \Pr(R = +1|v_i = +1)\Pr(v_i = +1) + \Pr(R = -1|v_i = -1)\Pr(v_i = -1). \quad (2)
\end{aligned}
$$

If $\Pr(v_i = +1) = \frac{1}{2}$, then the measures (1) and (2) are essentially equivalent, because

$$
\begin{aligned}
\text{power}_i &= \Pr(R = +1|v_i = +1) + \Pr(R = -1|v_i = -1) - 1 \\
&= \frac{1}{2}\Pr(\text{satisfied}) - 1 \qquad (\text{if } \Pr(v_i = +1) = \frac{1}{2}). \qquad (3)
\end{aligned}
$$

In general, however, the power and satisfaction measures differ and address slightly different concerns. For example, consider a committee in which 90% of the members vote Yes and 10% vote No. Then 90% of the voters are satisfied with the outcome, but the voters are essentially powerless at the individual level. Politically, this means that the advocates of Yes or No do not need to woo the voters. Thus, getting the desired outcome is not the same as having power or influence.

In general there are many ways of measuring the fairness of an voting system (see, for example, Gelman, 2002); we recognize that the results on voting power will not directly apply to other measures. We are working in the tradition of game theory in which voting power summarizes one aspect of the multi-player decision process.

For most of this paper, we study the formation of coalitions among n voters with equal weights that are allowed to form arbitrary coalitions. The coalitions can be nested so that the entire electorate has a tree structure. At the bottom of the tree, votes are decided by majority rule, with a winner-take-all rule for each coalition. We shall also consider weighted voting, in which voter i has weight w_i, and the winner at each stage in the tree is determined by weighted majority.

3   Coalitions and the probability of casting a decisive vote

We work with the assumption that votes are determined by independent coin flips, which we call *random voting*. This is the standard model used to define voting power in coalitions, and it is important as a default model for understanding coalitions without respect to information about the preferences of individual voters.[2] The assumption is not that individuals

---

[2]In elections with large numbers of voters, correlations among voter preferences make the random voting model inappropriate (see Gelman, Katz, and Tuerlinckx, 2002), but this paper focuses on legislative voting, especially for discretionary issues in which temporary coalitions ("log-rolling") are possible.

themselves vote randomly, but that we have no information about how they will vote and so we assume 0.5 probability of voting yes or no. We discuss in Section 4.2 how coalitions could be studied under more general models.

Under random voting, all 2^n vote configurations are equally likely, and so the power of voter i is simply $2^{-(n-1)}$ times the number of configurations of the other $n-1$ voters for which voter i is decisive (and counting semi-decisive configurations, in which votes are exactly tied, as $1/2$). Voting power calculations under the random model can thus be seen as combinatorial.

3.1   Individual and average voting power

Voting power is strongly linked to the theory of coalitions in cooperative games (see Luce and Raiffa, 1957). The starting point and fundamental result is that a group of voters can increase their individual voting powers by forming a *coalition*, which we define here as a binding agreement to pool their votes so that they all go to the winner in the coalition. In practice, this agreement could be long-term, as with a political party, or short-term, if a group of legislators agree to vote together on a circumscribed set of issues.

Individual voting power.   Figure 1 illustrates with various systems of coalitions among 9 voters. The top level of the tree is itself a sort of coalition, in that the total vote is reduced to a simple $+1$ or -1 to determine a final winner. These trees (and the accompanying calculations) illustrate the benefits of joining a coalition—and they also illustrate the negative side: voters who are left out of a coalition tend to do worse than if no coalitions had been formed at all.

In scenario A—simple majority voting—a voter is decisive if the other 8 are split evenly, which occurs with probability 0.273 under random voting.

Scenario B shows the benefit of being in a large coalition—any voter within the coalition of 5 is decisive if he or she can sway that coalition, which occurs with probability 0.375. Voters outside the coalition have zero power, and so the average power of all the voters is $\frac{5}{9} \cdot 0.375 + \frac{4}{9} \cdot 0 = 0.208$, which is lower than under majority voting.

Scenario C illustrates a more complicated structure in which 3 voters are in a coalition and the other 6 vote independently. Then how likely is your vote to be decisive? If you are in the coalition, it is first necessary that the other 2 voters in the coalition be split; this happens with probability $1/2$. Next, your coalition's 3 votes are decisive in the entire vote, which occurs if the remaining 6 voters are divided 3-3 or 4-2; this has probability $\frac{50}{64}$. The voting power any of the 3 voters in the coalition is then $\frac{1}{2} \cdot \frac{50}{64} = 0.391$. What if you are not in the coalition? Then your vote will be decisive if the remaining votes are split 4-4, which occurs if the 5 unaffiliated voters (other than you) are split 4-1 in the direction opposite to the 3 voters in the coalition. The probability of this happening is $\binom{5}{1} 2^{-5} = 0.156$. Compared to the simple majority system (scenario A in the figure), you have more voting power if you are in the coalition and less if you are outside. The average voting power is 0.234, which is lower than under majority voting.

Average voting power.   One way to study the total effect of coalitions is to compute the average probability of decisiveness for all of the n voters. It has been proved (and we sketch a proof in the next paragraph) that, under the random voting model, this average voting power is maximized under simple popular vote (majority rule) and is lower under

A. No Coalitions

    A voter is decisive if the others are split 4-4:

$$\Pr(\text{Voter is decisive}) = \binom{8}{4} 2^{-8} = 0.273$$

  Average $\Pr(\text{Voter is decisive}) = 0.273$
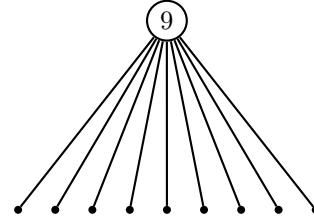
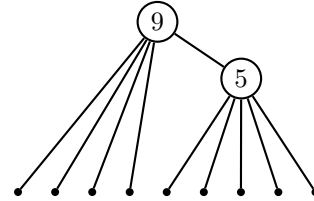**B. A Single Coalition of 5 Voters**

    A voter in the coalition is decisive if others in the coalition are split 2-2:

$$\Pr(\text{Voter is decisive}) = \binom{4}{2} 2^{-4} = 0.375$$

    A voter not in the coalition can never be decisive:

$$\Pr(\text{Voter is decisive}) = 0$$

  Average $\Pr(\text{Voter is decisive}) = \frac{5}{9}(0.375) + \frac{4}{9}(0) = 0.208$
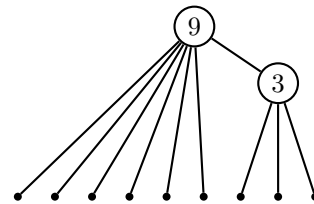
C. A Single Coalition of 3 Voters

    A voter in the coalition is decisive if others in the coalition are split 1-1 and the coalition is decisive:

$$\Pr(\text{Voter is decisive}) = \frac{1}{2} \cdot \frac{50}{64} = 0.391$$

    A voter not in the coalition is decisive with probability:

$$\Pr(\text{Voter is decisive}) = \binom{5}{1} 2^{-5} = 0.156$$

  Average $\Pr(\text{Voter is decisive}) = \frac{3}{9}(0.391) + \frac{6}{9}(0.156) = 0.234$

D. Three Coalitions of 3 Voters Each

    A voter is decisive if others in the coalition are split 1-1 and the other two coalitions are split 1-1:

$$\Pr(\text{Voter is decisive}) = \frac{1}{2} \cdot \frac{1}{2} = 0.250$$

  Average $\Pr(\text{Voter is decisive}) = 0.250$
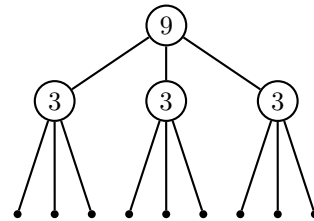
Figure 1: An example of four different systems of coalitions with 9 voters, with the probability of decisiveness of each voter computed under the random voting model. Each is a "one person, one vote" system, but they have different implications for probabilities of casting a decisive vote. Joining a coalition is generally beneficial to those inside the coalition but hurts those outside. The average voting power is maximized under A, the popular-vote rule with no coalitions. From Gelman, Katz, and King (2004).

Payoffs (in terms of your own voting power) from joining a coalition

Your option	Have other voters formed coalitions?	
	No	Yes
Stay alone	Moderate	Very low
Join a coalition	High	Low

Figure 2: Schematic of the prisoner's dilemma facing an individual voter when deciding whether to form a coalition. Joining a coalition increases your voting power, whatever the behavior of the other voters. However, if all the voters join coalitions, all of them have low voting power. Thus, a suboptimal outcome is assured if all players act rationally.

any coalition system. Figure 1 illustrates this point: the coalitions benefit their members but lower the average probability of decisiveness.

To prove the general result, we start with identity (3), which shows that, under random voting, maximizing average voting power is equivalent to maximizing the average probability of satisfaction. Only voters on the winning side will be satisfied, and so, conditional on the total vote, average satisfaction is maximized by choosing the option supported by more voters, which is simply majority rule. This theorem can also be seen as a corollary of more general results in graph theory (see Lemma 6.1 of Friedgut and Kalai, 1996). Specific results on average voting power can be obtained using combinatorics (see Kolpin, 2003).

One way to understand the result is that the winner-take-all rule in coalitions magnifies small differences (for example, a vote of 10-8 within a coalition is transformed to 18-0 at the next stage in the tree), which has the effect of amplifying noise (if the voting process is thought of as a system of communicating individual preferences up to the top of the tree). The least noisy system is majority rule, with no coalitions at all.

It may seem surprising that coalitions could cause *everybody's* voting power could decrease, but voting power is not a constant-sum quantity! To see this, consider an extreme case in which one committee member is chosen at random, and his or her vote determines the outcome. This system is symmetric with respect to the committee members and thus could be considered "fair," but it reduces any individual's voting power to $1/n$ (or, in the example of Figure 1, $1/9 = 0.111$), which is much lower than what would be obtained by simple majority voting (where the probability of casting a decisive vote is proportional to $1/\sqrt{n}$).

3.2 Forming coalitions as a prisoner's dilemma

Forming coalitions is beneficial to those who do it but is negative to "society" as a whole, at least in terms of average voting power. From a political perspective, this is reminiscent of the prisoner's dilemma (see, for example, Luce and Raiffa, 1957), a game in which the behavior that benefits each player in the game has negative consequences for all the players. The situation for voters is illustrated in Figure 2. To follow the usual terminology from game theory, to refuse to join a coalition is "cooperating," in the sense that this refusal is cooperative behavior with respect to the general population of voters. Conversely, joining

a coalition is "uncooperative" behavior relative to the general electorate, who are being excluded from the coalition. The prisoner's dilemma and equivalent collective action games have long been considered as fundamental models for political conflict and cooperation (see Coase, 1960, Axelrod, 1970, and Hardin, 1971).

As Figure 1 illustrates, it can be privately beneficial to join a coalition—especially if other voters have already done so. This raises the question, what sorts of coalition structures can arise spontaneously in a voting system? More formally, consider a set of n voters with no coalition structure (as in scenario A of Figure 1). Now allow groups of voters to join coalitions, with the rule that a set of voters will join a coalition only if the voting power increases for *each of the voters in the coalition.* Thus, we could move from scenario A to scenario C in Figure 1. From scenario C, we could then move to scenario D, since this transition benefits the six voters who would be joining the new coalitions.

We can think of this process as a walk in the space of trees. We shall refer to agreements as "locally beneficial" if the probability of decisiveness (voting power) increases for all the voters making the agreement. Possible agreements—that is, moves in tree space—include:

1. A set of separate voters forming a coalition

2. A coalition disbanding or dividing into sub-coalitions

3. A set of coalitions forming a super-coalition (without destroying their internal structure)

4. A set of coalitions merging into a single larger coalition.

In general, there can be more than one possible locally beneficial move from any given tree. For example, starting from scenario A in Figure 1, a move to either B or C is locally beneficial. And the move to B, for example, is not unique either, since any subset of 5 voters could form the coalition.

We can thus imagine rational voters moving through the space of trees, making locally beneficial agreements that lead to complex coalition structures that, as in scenario D of Figure 1, leave everyone worse off. These structures would not themselves be stable, however; for example, if the three coalitions in scenario D merge, they will return to scenario A. The actual behavior of the process depends on the moves that are allowed in tree space and the rules determining which locally beneficial agreements are made.

This nontransitivity of allowable moves between trees implies that there is no "objective function" that is increased by locally beneficial rules. This is related to the principle in economics that a group of agents can form a cartel that is Pareto-optimal within the group but has a negative utility for the larger economy of which it is a part.

3.3   Voting power with simple coalitions

To understand the potential benefits of coalitions for voting power, we perform closed-form and asymptotic computations for some relatively simple cases.

Forming a single coalition of size m.   Consider a system of n separate voters, and suppose m of these form a coalition. Under random voting, each of the m individuals in
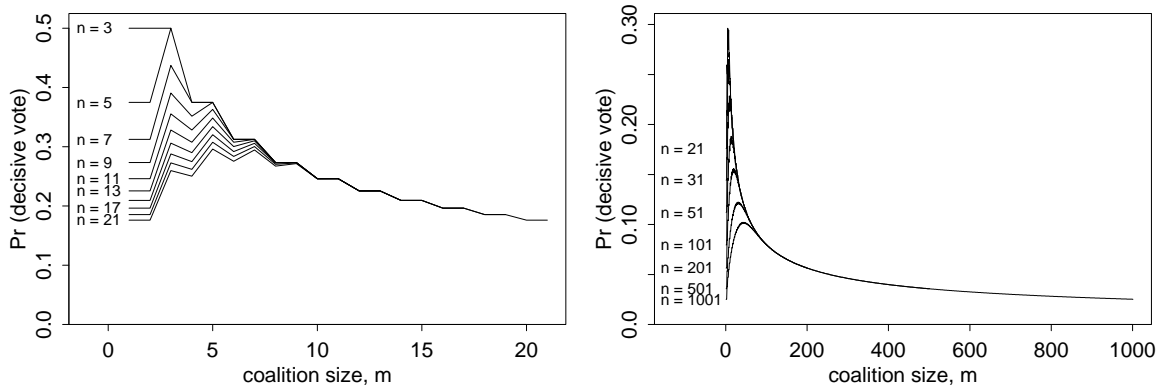
Figure 3: Probability of decisiveness (voting power) for a voter in a coalition of m, within a general population of n voters that are otherwise not divided into coalitions. Each line on the graphs shows the voting power as a function of m for a given value of n. The two plots are on different scales. The random voting model is assumed.

the coalition then has voting power,

$$\text{power}_i = \binom{m-1}{[(m-1)/2]} 2^{-(m-1)} \left(\Pr\left(\frac{n}{2} - m < x < \frac{n}{2}\right) + \frac{1}{2}\Pr\left(x = \frac{n}{2} - m\right) + \frac{1}{2}\Pr\left(x = \frac{n}{2}\right) \right), \tag{4}$$

where x has a binomial $(n-m, 1/2)$ distribution and represents the number of votes of "+1" among the $n-m$ voters not in the coalition. We in fact only need to evaluate (4) for odd values of n. If n is even, then voting power is unchanged if we evaluate at $n+1$, since the extra vote can only break a tie, which we are assuming would be done randomly anyway.

For any n, we can then evaluate (4) to see the potential gain in voting power from joining a coalition in an otherwise atomized electorate. Figure 3 shows voting power as a function of m, for each of several population sizes n. For clarity, low values of n are shown on the left plot and high values on the right plot. For any n, the minimum voting power is for $m = 1$ (no coalition), and, equivalently, $m = n$ (one large coalition). We can also see that it is never a good idea to have a coalition with an even number of members: if m is even, it is always as good or better to be in a coalition of size $m - 1$. Finally, voters only gain from forming relatively small coalitions—as the coalition size gets larger, voting power of the coalition members decreases, in the limit (where all the voters are in a single coalition) returning to the result with no coalitions at all.

Figure 3 reveals that the optimal coalition size increases slowly with n. To explore this more fully, we display in Figure 4 a plot of optimal coalition size vs. population size. The dotted line on the graph shows an asymptotic form for large n that we derive here. We begin by approximating the two factors in (4) using the normal distribution:

$$\text{for large } n: \quad \text{power}_i \approx \sqrt{\frac{2}{\pi m}} \left(2\Phi\left(\frac{m}{\sqrt{n-m}}\right) - 1 \right), \tag{5}$$

where $\pi = 3.14\ldots$ and Φ is the standard normal cumulative distribution function. Let $m_{\text{opt}}(n)$ be the value of m that maximizes (5) given n. To determine the behavior of m_{opt}
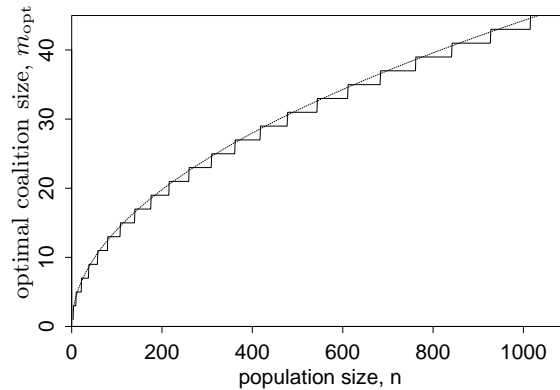
Figure 4: Optimal coalition size m_{opt}, for maximizing voting power under random voting, in a population of n voters that are otherwise not divided into coalitions. The dotted line shows the approximation $m_{\mathrm{opt}} = 1.4\sqrt{n}$. The solid line has jumps because the actual m_{opt} must be an integer and is always odd.

for large n, we first note that $m_{\mathrm{opt}}/n \to 0$ as $n \to \infty$ (because once m gets large compared to $\sqrt{n}$, the cumulative normal density saturates and the $\frac{1}{\sqrt{m}}$ factor causes power_i to decline as m increases further). We can then approximate (5) by,

$$
\begin{aligned}
\text{for large } n: \quad \mathrm{power}_i &\approx \sqrt{\frac{2}{\pi m}}\left(2\Phi\left(\frac{m}{\sqrt{n}}\right) - 1\right) \\
&= n^{-1/4}\sqrt{\frac{2}{\pi R}}\left(2\Phi(R) - 1\right), \tag{6}
\end{aligned}
$$

where $R = m/\sqrt{n}$. To maximize (6) given n, all we need to do is optimize the factor $\sqrt{\frac{2}{\pi R}}\left(2\Phi(R) - 1\right)$, which we can do numerically: the maximum is 0.565 and is achieved at $R = 1.40$.

Thus, for large n, the optimal m is approximately $1.4\sqrt{n}$, and the voting power from being in such a coalition is approximately $0.57\, n^{-1/4}$. This approximation is in fact good for small n also, as can be seen in Figure 4. The voting power for the coalition member can be compared to the approximate voting power of $\sqrt{\frac{2}{\pi n}} = 0.80\, n^{-1/2}$ if there is no coalition. Optimal coalition-formation allows an individual's voting power to decline in proportion to $n^{-1/4}$ rather than $n^{-1/2}$.

The relevance of these findings to our main point is that the prisoner's dilemma, which relies on the local benefits of coalition formation, applies for relatively small coalitions of size ranging from 3 to $1.4\sqrt{n}$. We shall now demonstrate the consequence of the prisoner's dilemma—thus if *all* voters follow the myopic but locally-beneficial policy of forming small coalitions, they all will suffer a loss of voting power.

All voters in coalitions. Now suppose that all n voters, knowing that forming coalitions increases their voting power, arrange themselves into coalitions of size m. Then there will be n/m such coalitions (assuming n is large enough that we can ignore that the numbers will not divide evenly). If we assume that both m and n/m increase with n, then

$$\text{with two levels of coalitions:} \quad \text{power}_i \approx \sqrt{\frac{2}{\pi m}}\sqrt{\frac{2}{\pi(n/m)}}$$
$$= \frac{2}{\pi\sqrt{n}}$$
$$= 0.65\, n^{-1/2}.$$

Thus, if all the voters form coalitions, they all become worse off than if they had stayed apart (in which case they would each have $\text{power}_i = 0.80\, n^{-1/2}$). This result generalizes and formalizes the examples and intuition discussed earlier.

Fully nested coalitions of size 3. In the most extreme case of coalition formation, consider an electorate of $n = 3^d$ voters that are arranged in coalitions of 3, that are themselves arranged in coalitions of 3, and so forth, to d levels. Then all the n voters are symmetrically-situated, and a given voter is decisive if the other 2 voters in his or her local coalition are split—this happens with probability $\frac{1}{2}$—and then the next two local coalitions must have opposite preferences—again, with a probability of $\frac{1}{2}$—and so on up to the top level. The probability that all these splits happen, and thus the individual voter is decisive, is $\left(\frac{1}{2}\right)^d = n^{-\log_3 2} = n^{-0.63}$, which is worse than simple majority voting, where voting power is proportional to $n^{-1/2}$.

3.4 More complex coalitions and weighted voting

In general, to evaluate the potential benefits of joining a coalition, a voter must know the configuration of the other voters in the committee. More specifically, if you are in a subtree that has n_{subtree} voters, and you are considering joining or leaving a coalition within that subtree, you must know the coalition structures of the others in the subtree. This will be necessary in order to determine the probability that your coalition is decisive within that subtree. At that point, the total vote of $\pm n_{\text{subtree}}$ propagates up the tree, and the structure of the remaining $n - n_{\text{subtree}}$ voters are irrelevant for the purpose of determining the proportional change in voting power from joining a coalition within the subtree.

Local calculations of changes in voting power. We assume that m voters will join a coalition only if it is locally beneficial, that is, if it increases the voting power for each of them (see Section 3.2). As noted just above, determining this increment of voting power in general requires calculation for all the other voters in the subtree. For a more tractable approximation, we consider the following local calculation of relative voting power.

Suppose that you are one of a small number m of voters in a subtree who are considering forming a coalition. Let V_{others} be the total vote in the rest of the subtree (applying winner-take-all rules within coalitions). The local calculation assumes that there are enough other distinct voters and voting blocs in the subtree that the distribution of V_{others} in the range

$[-m, m]$ is approximately uniform. Then we can approximate the change in voting power by counting your expected influence: the expected number of votes you will swing alone or in the coalition.

Coalitions of size 2 are *not* locally beneficial. For example, suppose you are considering forming a coalition with one other voter, so that $m = 2$. If you stay apart, your influence is 2 votes (the effect of changing from -1 to $+1$). If you join a coalition, then your combined vote will be ± 2, but there is only a $\frac{1}{2}$ chance that your vote will swing this (because your vote will either make or break a tie). So your expected influence is 2 votes, and joining the coalition gives no benefit.

Coalitions of size 3 *are* locally beneficial. Now consider your influence if you join with two other voters, so that $m = 3$. Your vote is decisive if the other two are split, which under random voting has a probability of $\frac{1}{2}$, and if this happens there will be a vote swing of 6. So your vote has an expected influence of 3. Thus, joining the coalition increases your voting power by an estimated factor of $3/2$.

Similar calculations show the estimated gains from joining larger coalitions. These calculations are valid as long as m is small compared to the standard deviation of V_{others}.

Weighted voting and coalitions of coalitions. Local calculations of approximate voting power can also be done for weighted voting. For example, consider again potential coalitions of size 2 or 3, but this time of weighted voters.

It is clear that it never makes sense to form a coalition of 2: if the voters have equal weights, the earlier calculation applies, and if the weights are unequal, then the voter with lower weight will always be outvoted and will have no power in the coalition.

For the same reason, three voters with weights w_1, w_2, w_3 should consider joining a coalition only if their weights satisfy the triangle inequality: that is, $w_1 < w_2 + w_3$, and so forth. In this case, we can locally approximate the potential benefit of joining. Suppose you are voter 1, so that staying apart gives you an influence of $2w_1$ on the total vote in your subtree. If you join the coalition, there is a $\frac{1}{2}$ chance your vote will be swing the entire group; your expected influence is thus $\frac{1}{2} \cdot 2(w_1 + w_2 + w_3)$. The gain in expected influence from joining is then $w_2 + w_3 - w_1$, which we already know is positive from the triangle inequality condition. The stability of coalitions of 3 is thus robust and holds under weighted voting (as long as neither of the three voters dominates the other two).

The same calculations apply when considering whether it is beneficial for a set of coalitions to form a super-coalition. In this case, each of coalition acts as a weighted voter in determining voting power. If the coalitions each gain voting power, then so do the individual voters within.

4   Discussion and extensions to the model

4.1   Formation and dissolution of coalitions

How can voters can spontaneously form structure through mutually-beneficial coalitions? One way to understand the behavior of coalitions would be through some sort of simulation, thinking of voters as cellular automata that form structures of local agreements.

In going beyond random voting, we need a model in which voters have unequal probabilities p_i of voting Yes and some sort of spatial structure (so that voters have neighbors with whom they can confer and consider joining coalitions). The model can be set up on a purely theoretical basis; for example, by using a Poisson process to place n voters on a two-dimensional space, assigning values of z_i from a Gaussian process, and then determining votes v_i with the rule, $\Pr(v_i = 1) = \Phi(z_i)$. Gelman, Katz, and Tuerlinckx (2002) consider some models of correlated votes within a tree structure. Another approach, more appropriate to models of voting in a legislature, is to set up probabilities and a correlation structure based on data from roll-call votes. In this case, the voting options of -1 and -1 should be assigned consistently across the issues being voted on, as in Poole and Rosenthal (1997).

We conjecture that the coalition-formation process is inherently unstable (hence the subtitle of this paper). By this we mean that if voters are in an ongoing process of joining and leaving coalitions, with each decision made myopically to immediately increase voting power, then there is no stable equilibrium coalition structure. Section 3.2 illustrates this general idea in the context of the example in Figure 1, where the different coalition structures form an intransitive cycle. This is consistent with the finding of Bernholz (1982) that cyclical preferences are possible in the presence of externalities—here, the externality is that joining a coalition can benefit all members of the in-group but yield a net loss in voting power for the entire legislature. To prove instability in general would require specification of possible moves in tree space, as discussed in Section 3.2, and assumptions about voters' preferences.

4.2   Heterogeneous voters

An important generalization to the voting model allows voters to have different preferences. Here we present some preliminary results that shed light on the conditions under which it is beneficial for these heterogenous voters to form coalitions.

For each voter i, we let p_i be the probability that he or she will vote for the + option. We assume that voters know each others' probabilities and can form binding coalitions with the goal of increasing their voting power. Compared to our analysis in Section 3 of random voting, our results on heterogeneous voters are more provisional, but we still find a prisoner's dilemma for coalition formation. The extent of the prisoner's dilemma and the stability of coalitions depends on the heterogeneity among the voters—in general, coalitions are more likely to be mutually beneficial among voters with probabilities p_i near $\frac{1}{2}$.

No benefit from joining a coalition of size 2. The next step is to generalize the results of Section 4.1 to go beyond the random voting model. This research can go in many directions; we illustrate here for the simple problem of evaluating the feasibility of coalitions of size 2 and 3. We shall evaluate the decisions based on the local calculations of expected influence, as described in Section 3.4.

Suppose you are a voter with probability p_1 deciding whether to join a coalition with another voter with probability p_2. If you stay apart, your influence is 2 votes (a change from -1 to $+1$). If you join the coalition, the expected total vote from the coalition is $2p_2$ if you vote $+1$ and $-2(1 - p_2)$ if you vote -1, and so the difference—the expected influence—is still 2. There is, once again, no benefit to forming a coalition of size 2.
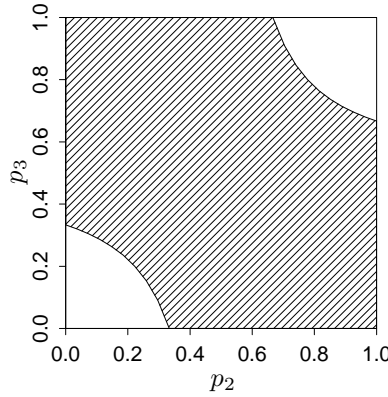
Figure 5: Suppose three voters, with equal weights and probabilities p_1, p_2, p_3 of voting $+1$, are considering forming a coalition. The shaded area on the graph shows the region of values of (p_2, p_3) for which it is beneficial for voter 1 to join the coalition. The coalition will form only if (p_1, p_2) and (p_1, p_3) are also in the shaded area.

Potential benefit from joining a coalition of size 3. Now suppose you are a voter with weight w_1 and probability p_1 deciding whether to join a coalition with two other voters with weights w_2, w_3 and probabilities p_2, p_3. To determine the benefit from joining a coalition, note that if the other two voters agree, then your vote will have no influence. Your expected influence through the coalition is thus,

$$
\begin{aligned}
\text{expected influence in coalition} \;=\;& (w_1 + w_2 + w_3)\Pr(\text{voters 2 and 3 disagree}) \\
=\;& 2(w_1 + w_2 + w_3)(p_2(1 - p_3) + p_3(1 - p_2)). \quad (7)
\end{aligned}
$$

Here we are assuming the voters are independent given the probabilities p_i. If not, the probability of disagreement can be calculated from the joint distribution.

In any case, the local calculation states that you should join the coalition if this expected influence (7) is greater than the expected gain of $2w_1$ if you were to vote alone; that is,

$$
\text{join if} \quad (w_1 + w_2 + w_3)(p_2(1 - p_3) + p_3(1 - p_2)) - w_1 > 0.
$$

After some algebra, this can be rewritten as,

$$
\text{join if} \quad (1 - 2p_2)(1 - 2p_3) < 1 - 2\frac{w_1}{w_1 + w_2 + w_3}. \quad (8)
$$

(The right side must be positive or else voter 1 is dominant and voters 2 and 3 would have no motivation to join the coalition.) For all three voters to be willing to join the coalition, condition (8) must also hold with the other two permutations of the indexes $\{1, 2, 3\}$.

For example, if $w_1 = w_2 = w_3$, then (8) becomes $(1 - 2p_2)(1 - 2p_3) < \frac{1}{3}$. The shaded region of Figure 5 shows the conditions on (p_2, p_3) where this condition holds. The points in the shaded region correspond to high probabilities that voters 2 and 3 will disagree, which is when it is beneficial for voter 1 to join the coalition. A sufficient (but not necessary)

condition for this condition to hold for all three voters is that p_1, p_2, and p_3 all be in the range $(0.5 \pm 1/\sqrt{12}) = (0.21, 0.79)$. This is a fairly broad range of probabilities, indicating the robust strength of coalitions of size 3 when weights are equal.

Larger coalitions. For larger coalitions, we can continue to use expected influence to approximately determine whether a potential coalition is locally beneficial. For example, consider m voters with equal weights, potentially unequal probabilities p_i, and independent votes given the p_i's. As discussed earlier, it is reasonable to assume independence if enough structure is built into the probabilities p_i. Under the local calculation, the coalition is possible if the expected influence of any vote within the coalition is larger than what could be gained by voting alone. This depends on the probability that the total vote is tied, which in turn is strongly dependent on the expected vote differential, $\mathrm{E}(\overline{V}) = \frac{1}{m} \sum_{i=1}^{m}(1 - 2p_i)$. If $\mathrm{E}(\overline{V})$ is close to 0, then it can be very beneficial to join (see Section 3.3). However, if $|\mathrm{E}(\overline{V})| >> \mathrm{sd}(\overline{V})$, then the probability of a tie will be too low for the coalition to be beneficial, compared to voting separately.

Thus, a large group of voters who agree will *not* increase their individual voting power by joining a coalition, but individuals in a more heterogeneous group may benefit (in terms of voting power) from joining together. These results can change if we study the probability of satisfaction rather than voting power, since identity (3) holds in general only if $p_i = \frac{1}{2}$.

4.3   Conclusions

In voting, formation of coalitions has features of the prisoner's dilemma game, in that a coalition increases voting power for its members but decreases the average voting power of the electorate. Under the random voting model (in which all vote combinations are equally likely), the optimal coalition size is approximately 1.4 times the square root of the number of voters. Going beyond the random voting model, one can explore the logic of coalitions among voters who differ in their preferences. Coalitions of size 2 never make sense, coalitions of size 3 make sense under a wide range of preferences, and large coalitions should only form among groups whose average preferences approximately cancel out. Coalitions among a group of like-minded voters are not beneficial in terms of voting power.

With coalition formation viewed as a prisoner's dilemma, one would expect an inherent instability—any structure of coalitions can be transformed (by forming new groupings or breaking old ones) into a new structure that increases the power for all the voters involved in making the change. Thus, one would expect coalitions to be fluid, even if voters' preferences remained fixed. The theory of voting power therefore potentially presents an endogenous mechanism for political instability, even in a world where individuals' (probabilistic) preferences are fixed and known.

References

Axelrod, R. (1970). *Conflict of Interest: A Theory of Divergent Goals with Applications to Politics*. Chicago: Markham.

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

Bernholz, P. (1982). Externalities as a necessary condition for cyclical social preferences. *Quarterly Journal of Economics* **97**, 699–705.

Brams, S. J., and Davis, M. D. (1974). The 3/2's rule in presidential campaigning. *American Political Science Review* **68**, 113–134.

Coase, R. (1960). The problem of social cost. *Journal of Law and Economics* **3**, 1–44.

Colantoni, C. S., Levesque, T. J., and Ordeshook, P. C. (1975). Campaign resource allocation under the electoral college (with discussion). *American Political Science Review* **69**, 141–161.

Felsenthal, D. S., and Machover, M. (1998). *The Measurement of Voting Power.* Edward Elgar Publishing.

Friedgut, E., and Kalai, G. (1996). Every monotone graph property has a sharp threshold. *Proceedings of the American Mathematical Society* **124**, 2993–3002.

Gelman, A. (2002). Voting, fairness, and political representation (with discussion). *Chance* **15**, 3, 22–26.

Gelman, A., Katz, J. N., and King, G. (2004). Empirically evaluating the electoral college. In *Rethinking the Vote: The Politics and Prospects of American Election Reform*, ed. A. N. Crigler, M. R. Just, and E. J. McCaffery, 75–88. Oxford University Press.

Gelman, A., Katz, J. N., and Tuerlinckx, F. (2002). The mathematics and statistics of voting power. *Statistical Science* **17**, 420–435.

Hardin, R. (1971). Collective action as an agreeable N-prisoners' dilemma. *Behavioral Science* **16**, 472–479.

Heard, A. D., and Swartz, T. B. (1999). Extended voting measures. *Canadian Journal of Statistics*, **27**, 177–186.

Kolpin, V. (2003). Voting power under uniform representation. *Economics Bulletin* **4** (2), 1–5.

Luce, R. D., and Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey*, chapter 12. Harvard University Press.

Poole, K. T., and Rosenthal, H. (1997). *Congress: A Political-Economic History of Roll Call Voting.* Oxford University Press.

Saari, D. G., and Sieberg, K. K. (1999). Some surprising properties of power indices. Northwestern University Center of Math Econ Discussion Paper #1271.

Shapley, L. S., and Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review* **48**, 787–792.

Snyder, J. M., Ting, M. M., and Ansolabehere, S. (2001). Legislative bargaining under weighted voting. Technical report, Department of Economics, Massachusetts Institute of Technology.

Straffin, P. D. (1978). Probability models for power indices. In *Game Theory and Political Science*, ed. P. Ordeshook. New York University Press, 477–510.