DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

HOW MUCH DOES A VOTE COUNT? VOTING POWER, COALITIONS, AND THE ELECTORAL COLLEGE

Andrew Gelman
Columbia University

Jonathan N. Katz
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 1121

May, 2001

How much does a vote count? Voting power, coalitions, and the Electoral College*

Andrew Gelman†　　Jonathan N. Katz‡

Abstract

In an election, the probability that a single voter is decisive is affected by the *electoral system*—that is, the rule for aggregating votes into a single outcome. Under the assumption that all votes are equally likely (i.e., random voting), we prove that the average probability of a vote being decisive is maximized under a popular-vote (or simple majority) rule and is lower under any coalition system, such as the U.S. Electoral College system, no matter how complicated. Forming a coalition increases the decisive vote probability for the voters within a coalition, but the aggregate effect of coalitions is to decrease the average decisiveness of the population of voters. We then review results on voting power in an electoral college system. Under the random voting assumption, it is well known that the voters with the highest probability of decisiveness are those in large states. However, we show using empirical estimates of the closeness of historical U.S. Presidential elections that voters in small states have been advantaged because the random voting model overestimates the frequencies of close elections in the larger states. Finally, we estimate the average probability of decisiveness for all U.S. Presidential elections from 1960 to 2000 under three possible electoral systems: popular vote, electoral vote, and winner-take-all within Congressional districts. We find that the average probability of decisiveness is about the same under all three systems.

Keywords: coalition, decisive vote, electoral college, popular vote, voting power

*We thank Stephen Ansolabehere, Robert Erikson, and Amanda O'Brien for helpful comments.

†Department of Statistics, Columbia University, New York, NY, 10027; gelman@stat.columbia.edu.

‡Division of the Humanities and Social Sciences, D.H.S.S. 228-77, Pasadena, CA 91125; jkatz@caltech.edu, http://jkatz.caltech.edu.

1 Introduction

The 2000 U.S. Presidential election has rekindled interest in exploring electoral reform, in particular the possible modification or outright elimination of the Electoral College. While we do not directly address the normative question of the value of the U.S. Electoral College, this paper does explore how coalitional behavior under various electoral rules affects the probability a given voter is decisive in an election, a natural measure to evaluate an electoral system. The Electoral College is just a particular coalitional voting system in which voters in a given state give all of their votes to the majority winner in that state.

The probability of a vote being decisive is important directly—it represents your influence on the electoral outcome, and this influence is crucial in a democracy—and also indirectly, because it could influence campaigning. For example, one might expect campaign efforts to be proportional to the probability of a vote being decisive, multiplied by the expected number of votes changed per unit of campaign expense, although there are likely strategic complications since both sides are making campaign decisions (c.f. Brams and Davis 1974). The probability that a single vote is decisive in an election is also relevant in determining the utility of voting, the responsiveness of an electoral system to voter preferences, the efficacy of campaigning efforts, and comparisons of voting power (Riker and Ordeshook 1968, Ferejohn and Fiorina 1974, Brams and Davis 1975, Aldrich 1993). Perhaps the simplest measure of decisiveness is the (absolute) Banzhaf (1965) index, which is the probability that an individual vote is decisive under the assumption that all voters are deciding their votes independently and at random, with probabilities 0.5 for each of two candidates. We shall refer to this assumption as the *random voting model*. While clearly an unrealistic assumption, it does provide a benchmark to evaluate competing electoral rules and make the problem theoretically tractable.

In a complicated electoral system, different voters have different probabilities of decisiveness, at which point it is natural to: (a) compare the probabilities for differently-situated voters, and (b) average the probabilities across all voters in the electorate. As we shall see, the results of (a) and (b) are closely related. Suppose that one is designing an electoral system, allowing for coalitions and winner-take-all subsets and even coalitions within coalitions (for example, subcommittees in a legislature or mini-electoral colleges within states). We show that the average probability of decisiveness under the random voting model is maximized under a popular-vote (or simple majority-rule) system. Any coalitional formation can only *reduce* the average probability of decisiveness, averaging across all the voters.

The random voting model is, of course, a gross oversimplification, and in fact its implications for voting power in U.S. elections have been extremely misleading in the political science literature, as has been discussed by Gelman, King, and Boscardin (1998). Under the random voting model it is easy to see that the Electoral College increases the voting power—that is, the probability of an individual's vote is decisive—for voters in larger states. However, this result is not relevant in practice as we will show by examining actual elections.

In calculating the probability that a voter is decisive allowing for coalitions there are two relevant components: the probability that that the voter is pivotal in determining the coalition's (i.e., the state's) choice and the the probability that the coalition is pivotal. Under the random voting model, the probability that an individual voter is pivotal in the coalitional choice decreases with size of the coalition, but this is more the compensated for because the larger coalitions (states) have larger influence on the final outcome. However, this result is artifactual when we examine actual U.S. Presidential returns state by state. We large states are only slightly more likely than small states to have close elections, and this difference is not enough to offset the benefit the small states get from having a guaranteed minimum of three electoral votes. In total, we find that voters in small states are actually advantaged by the Electoral College.

As second approach to the empirical analysis of U.S. Presidential elections, we use results from every since 1960 as the basis of a set of simulations to calculate the average probability that a given voter is decisive under the popular vote, the electoral college, and an alternative system in which each Congressional district is worth one electoral vote. We find that the average probability of decisiveness is similar under all three systems (although they differ as to how this probability is distributed among the voters).

We review the basic ideas of voting power and decisive votes, as well as present our basic notation, in Section 2. Then in Section 3 we present our main theoretical result along with some heuristic explanations and a discussion of how endogenous coalitions formation can arise that in the aggregate and make individual voters worse off. In Section 4 we consider deviations from random voting and how to estimate probabilities of decisive votes from actual elections. Section 5 applies these ideas to recent U.S. Presidential elections. The final section concludes.

2 Voting power under the random voting model: a review

We begin by reviewing the basic ideas of the voting-power literature and at the same time introducing the mathematical notation that we shall use in Section 3 to prove our main result.

We consider an election with two options, or candidates, which we shall refer to as $+$ and $-$; and n voters with votes v_i, $i = 1, \dots, n$, where each v_i is either $+1$, a vote for candidate $+$, or -1, a vote for candidate $-$. An *electoral system* is categorized by a rule, which we label R, transforming the vector $v = (v_1, \dots, v_n)$ to an electoral outcome: $R(v) = \{-1, 0, +1\}$, where 0 denotes the (presumably unlikely) event of a tie. Ultimately, a winner must be chosen, and so if $R(v) = 0$, a coin is flipped to see if the election goes to $+$ or to $-$.

The simplest electoral system is the *popular vote* or *majority rule*, under which $R(v) = \text{sign}(\sum_{i=1}^{n} v_i)$. But more general systems are possible; in fact, we consider all possible rules here (with some minor restrictions). A familiar example, from the U.S. Presidential system, is an *electoral vote* or *local winner-take-all* rule, in which the n voters are grouped into several coalitions, in which the winner in each coalition gets a fixed number of "electoral votes" (with these electoral votes split or randomly assigned in the event of an exact tie within the coalition), and then the candidate with the most electoral votes is declared the winner (or with a coin flipped if the electoral vote is tied).

What is the probability that a given voter is decisive? This means that a change in one vote from + to − could change the outcome of the election. We define

$$\text{voting power} = \text{probability that voter } i \text{ is decisive } = \frac{1}{2}\text{E}(R(v^+_{(i)}) - R(v^-_{(i)})), \quad (1)$$

where $v^+_{(i)}$ is the complete vector of n votes if voter i chooses candidate + and $v^-_{(i)}$ is the vector if he or she votes for −. Thus, $v^+_{(i)}$ and $v^-_{(i)}$ differ only in position i. Voting power, then, is an expectation of a quantity that equals 0 (in the most likely event that the vote makes no difference), 1/2 (if the vote can make or break a tie) or 1 (if the vote can singlehandedly determine the winner). Since a tie itself would be broken at random, this definition of voting power is equivalent to the probability that a change in v_i will change the election outcome.

Expression (1) is almost equivalent to the Banzhaf (1965) index, with the minor change of treating ties as 1/2. Other power indexes, which we do not discuss here, include that of Shapley and Shubik (1954) and the satisfaction index of Straffin (1978). For both these indexes, a voter receives points for being on the winning side of a vote even if the vote is not close. In this paper, we focus on the direct probability that a vote is decisive. As we shall discuss in Section 4, a key concern in applying these results is not the choice of "voting power" definition so much as the assumptions about the probability distribution of the n votes (See also Finkelstein and Levin 1990 and Heard and Schwartz 1999) for further discussion of these issues.

2.1 Some examples

Before getting to our main result in Section 3, we explore the power index defined in Equation (1) in some simple examples. For simplicity, we will just consider examples with an odd number of voters, so that your vote is decisive if the other voters are exactly split. We continue to assume random voting.

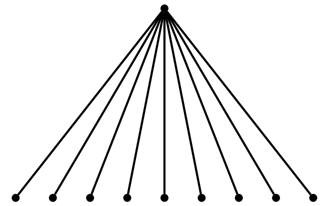
In a popular vote system with n voters, the probability that your vote is decisive is $\binom{n-1}{(n-1)/2} 2^{-(n-1)}$; that is, the probability that $x = \frac{n-1}{2}$ where x has a binomial distribution with parameters $n-1$ and $\frac{1}{2}$. For large (or even moderate) n, this can be well approximated using the normal distribution as $\sqrt{\frac{2}{\pi}} n^{-1/2}$, a standard result in probability (c.f., Woodroofe 1975).

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(Voter is decisive) = 0.273



B. A Single Coalition of 3 Voters.

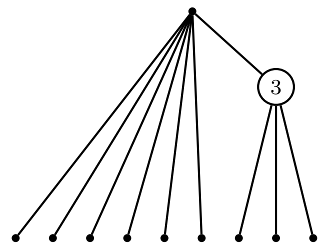
A voter in the coalition is decisive if others in the coalition are split 2-2 and the coalition is decisive:

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

A voter not in the coalition is decisive with probability:

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

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



C. A Single Coalition of 5 Voters.

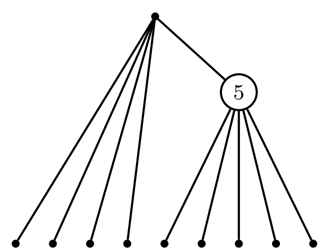
A voter in the coalition is decisive if others in the coalition are split 2-2:

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

A voter not in the coalition can never be decisive:

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

Average Pr(Voter is decisive) = $\frac{5}{9}(0.375) + \frac{4}{9}(0) = 0.208$



B. Three Coalitions of 3 Voters Each.

A voter is 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(Voter is decisive) = 0.250

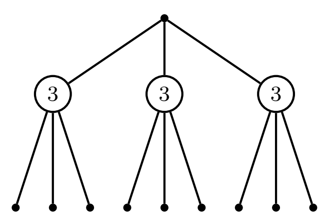


Figure 1: *An example of four different electoral systems with 9 voters. Each is a "one person, one vote" system, but they have different implications for probabilities of casting a decisive vote. The average probability of decisiveness is maximized under A, the popular-vote rule.*

Now consider *coalitions*, in which the members of a coalition with m members have the prior agreement that they will separately caucus, and then the winner of the vote in their coalition will get all of their m votes.

Under the random voting model, it is smart to join a coalition. At one extreme, suppose a single coalition has $\frac{n+1}{2}$ voters. Then this coalition determines the election outcome, and if you are in that coalition, your vote is decisive with approximate probability $\sqrt{\frac{2}{\pi}}\left(\frac{n}{2}\right)^{-1/2}$, which is approximately $\sqrt{2}$ times the probability of being decisive under the popular vote system. However, the $\frac{n-1}{2}$ voters not in the grand coalition have zero voting power; thus, the *average* probability of a decisive vote, averaging over all voters, is approximately $\sqrt{\frac{1}{\pi}}n^{-1/2}$, a factor of $\sqrt{2}$ *less* that under the popular vote system.

Voters benefit even under small coalitions. For a simple example, consider a election with 9 voters under different electoral rules as depicted in Figure 1. Under the popular vote system, any voter's chance of being decisive is $\binom{8}{4}2^{-8} = 0.273$. Now suppose that 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 election, 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 popular vote system, you have more voting power if you are in the coalition and less if you are outside. The average voting power is $\frac{3}{9}0.391 + \frac{6}{9}0.156 = 0.234$, which is lower than under the popular vote system (see panel A. Figure 1).

Finally, one can consider more elaborate arrangements. For example, suppose there are $n = 3^d$ voters, where d is some integer, who are divided into three equal-sized coalitions, each of which is itself divided into three coalitions, and so forth, in a tree structure. 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. The probability that all these splits happen, and thus the individual voter is decisive, is $\frac{1}{2^d} = n^{-\log_3 2} = n^{-0.63}$, which is lower than the probability under the popular vote system (for large n, that probability is approximately $0.8n^{-0.5}$). For example, if $n = 3^8 = 6561$, then the probability of a decisive vote is 1/256 with the tree-structure of coalitions, compared to about 1/102 with majority rule.

The above examples indicate that under the random voting model, it is to your benefit to be in coalitions, with larger coalitions generally being better. If you are in a coalition of size m, the probability that your coalition is tied is approximately proportional to $m^{-1/2}$, and the probability that your coalition is itself required to determine the election winner is approximately proportional to m; the product of these two probabilities thus increases with m, at least for $m << n$. A similar logic leads to large coalitions themselves fragmenting into sub-coalitions.

In a system such as the U.S. electoral college, the coalitions are set ahead of time rather than being subject to negotiation, and the mathematical analysis just presented suggests that voters in larger states have more voting power (Banzhaf 1967). This comparison is not valid in practice, however, given the observed departures from the random voting model, as we discuss in Section 5.

To return to the theoretical argument, however, it is sensible to be in a coalition, but the end result of coalitioning, at least in the examples we have considered, is to *decrease* the average probability of individual voters being decisive. We explore this paradoxical finding in the next section.

3 Popular-vote system maximizes average voting power under the random voting model

We prove here the main theoretical finding of our paper. Before proceeding to our main result we must assume the voting rules under consideration are *monotonic*:

Definition 1 (Monotonic) *A voting rule is monotonic if a single vote changed from $-$ to $+$ can never make the election of candidate $+$ less likely.*

This restriction can cause problems when several options or candidates are being considered (Arrow 1951) but is uncontroversial in a two-candidate election. Most common voting rules satisfy this restriction.

We can now state our main theoretical result:

Proposition 1 *Suppose that n voters vote according to the random voting model. Then, for any monotonic voting rule, the average probability of voters' being decisive is maximized under majority rule (i.e., popular vote); any other system for determining the election winner, no matter how elaborate, can only decrease (or leave unchanged) the average probability of decisiveness.*

In order to prove this result we need to define the following notation. Let v represent a vector of n votes, each of which is $+$ or $-$. We use $p(v)$ to index the $n!$ permutations of these votes. From Expression (1), the decisiveness of voter i is determined by the possibilities $v^+_{(i)}$ and $v^-_{(i)}$, which we call a *configuration pair*. If k is the number of voters, excluding voter i, who vote $+$, then we consider all permutations p that carry all the k voters for $+$ to the beginning of the list, the $n-k-1$ voters for $-$ to the end of the list, and put voter i in the k-th position. We call this an *ordered permutation*, and we use the notation $v(p)$ to indicate the vote vector for this permutation. For any given configuration pair with k pluses and $n-k-1$ minuses, there correspond $k!(n-k-1)!$ ordered permutations.

3.1 Proof

Under the random voting model, all configuration pairs are equally likely (each with probability $2^{-(n-1)}$), and the probability that voter i is decisive is simply $2^{-(n-1)}$ times the sum over all configuration pairs of the voting power number $\frac{1}{2}(R(v^+_{(i)}) - R(v^-_{(i)}))$ from Expression (1). The systemwide average probability of decisiveness is then just $\frac{1}{n}$ times the sum of this over all voters.

The key to our proof is a method of summing the voting power over all voters and all configuration pairs. To sum over configuration pairs, we can sum over all ordered permutations, giving each ordered permutation a weight of $\frac{1}{k!(n-k-1)!}$ to correct for the multiple counting and multiplying by $2^{-(n-1)}$ because that is the probability of any configuration pair under random voting. Then averaging the power index (1) yields,

$$\text{Average Pr(Voter is decisive)} = \frac{1}{n}2^{-(n-1)} \sum_{\text{permutations } p} \sum_{k=0}^{n-1} \frac{1}{k!(n-k-1)!}\frac{1}{2}\left[R(v(p)^+_{(i)}) - R(v(p)^-_{(i)})\right] \quad (2)$$

Each of the inner sums in (2) considers the n jumps between the $n+1$ ordered vote sequences (for example, if $n = 4$, these would be $----$, $+---$, $++--$, $+++-$, $++++$). For any permutation p, the n individual differences $\frac{1}{2}\left[R(v^+_{(i)}) - R(v^-_{(i)})\right]$ are each equal to 0, $\frac{1}{2}$, or 1 (because of the monotonicity assumption, they can never be negative), and they sum to at most 1.

The inner sum of Expression (2) is therefore maximized when the positive differences are attached to the highest values of the weight, $\frac{1}{k!(n-k-1)!}$. For an odd number of voters, this occurs when $k = \frac{n-1}{2}$, and under majority rule the difference $\frac{1}{2}\left[R(v(p)^+_{(i)}) - R(v(p)^-_{(i)})\right]$ equals 1 at this point. For an even number of voters, the weight is maximized when $k = \frac{n-2}{2}$ or $\frac{n}{2}$, and under majority rule the difference is $\frac{1}{2}$ at each of these points.

Thus, under majority rule, Expression (2) reduces, for odd n, to

$$\frac{1}{n}2^{-(n-1)} \sum_{p} \frac{1}{((n-1)/2)!((n-1)/2)!} \cdot 1 = \binom{n-1}{(n-1)/2} 2^{-(n-1)}$$

after summing over the $n!$ permutations. This is the maximum value of Expression (2) under any rule R for combining the n votes. Similarly, if n is even this expression is maximized at $\binom{n-1}{n/2}2^{-(n-1)}$ which also is realized under majority rule.

3.2 Heuristic explanations

As noted at the end of Section 2, it appears paradoxical that coalition formation, which increases the individual probability of decisiveness, uniformly decreases the probability when averaged across all voters. We have developed two heuristic arguments in order to better understand this result.

First we provide a technical explanation based on the direct form of the proof of the proposition. The switch from + to − has to occur at some point, and an equal split of the remaining $n - 1$ votes is the most likely outcome (under the random voting model). To put it another way, an electoral college system allows decisive votes when the votes are not split exactly evenly, but this comes at a cost of *not* allowing decisive votes in the more probable scenarios in which the votes are divided evenly.

Our second explanation considers the standpoint of the individual voter. Any aggregation adds noise to the system. For example, the winner-take-all system within U.S. states means that Florida's electoral votes will be divided 25-0, rather than 12.5-12.5. In a presidential system, a final aggregation is necessary to choose a single winner, but any coalitioning before this stage adds randomness, which reduces the probability of individual voters being decisive. For an extreme example, suppose that 270 electoral votes were chosen by a coin flip and the remaining 268 from the popular vote. In this case, all voters are treated symmetrically (so the voting system is "fair"), but any vote has a zero chance of being decisive. More generally, any randomness decreases the expected effect of a change in any individual vote.

4 Estimating voting power empirically

As has been noted by many researchers (e.g., Beck 1975, Margolis 1977, Merrill 1978, and Chamberlain and Rothchild 1981), there are theoretical and practical problems with a model that assumes votes are independent coin flips. The simplest model extension is to assume votes are independent but with probability p of going for the + outcome, with some uncertainty about p (for example, p could have a normal distribution with mean 0.50 and standard deviation 0.05). However, this model is still too limited to describe actual electoral systems. In particular, the parameter p must realistically be allowed to vary, and modeling this varying p is no easier than modeling vote outcomes directly. Following Gelman, King, and Boscardin (1998), one might try to construct a hierarchical model, as they did for U.S. Presidential elections with uncertainty at the national, regional, and state levels.

It is not our purpose here to come up with realistic models for voting; rather, we wish to understand the sensitivity of voting power results to the random voting assumption that all 2^n vote outcomes are equally likely. Here are some of the key implications of this assumption.

First, the vote differential, $\sum_{i=1}^{n} v_i$, is, under the model, unrealistically close to 0. For example, if $n = 1$ million, this difference has a mean of 0 and a standard deviation of 1000, and of course very few actual elections of this size are decided by less than 1000 votes. The result is that the random voting model drastically overestimates the probability of decisiveness, which has implications when considering the instrumental benefits of voting (see Beck 1975, and Gelman, King, and Boscardin 1988).

A second implication of random voting is that the vote differential within a coalition of size m has mean 0 and standard deviation proportional to $\sqrt{m}$. The probability the vote is exactly tied is thus approximately proportional to $1/\sqrt{m}$. At the next stage, the probability that the coalition is decisive for the final outcome is approximately proportional to m (if $m << n$). The product of the two probabilities indicates the voting power of an individual voter within the coalition, and is roughly proportional to $\sqrt{m}$, thus favoring large coalitions. As we discuss in Section 5, it is a mistake to try to apply this result to U.S. Presidential elections.

Third, the random voting model assumes that all voters are indistinguishable in their preferences, which obviously does not describe the real world in where voters identify themselves as Democrats and Republicans (and there are consistent differences among individuals within each party) in elections or committees. The relevance for the voting power results is that they assume coalitions of symmetric voters: for any individual voter, the goal is to be in a coalition that is otherwise split evenly. In reality, however, coalitions typically join like-minded voters, and so joining a coalition is probably less beneficial than under random voting.

A key variable affecting the probability of decisive vote is how likely the popular vote is to being tied. If the election is likely to be very close (as in the year 2000 U.S. Presidential election), then the theoretical result of Section 3 should hold: the average probability of a decisive vote should be highest under majority rule. If the most likely outcome is an uneven split in the vote, then it is possible (but not necessary) that a coalition system can increase the average probability of decisiveness. We explore this for U.S. Presidential elections in Section 5.1.

In summary, real elections and legislatures have far fewer exact or close-to-exact splits—among all n voters and within coalitions—than would be expected under the random voting model. Any voter is less likely to be decisive, and the advantage to joining a coalition, particularly a large coalition, is less would be estimated under random voting.

4.1 Using empirical models of votes to estimate the probability of decisiveness

We now discuss how to use empirical models of votes to compute the probability that a single vote is decisive. Consider a two-candidate election with majority rule in any given jurisdiction. Let V be the proportional vote differential (e.g., the difference between the Democrat's and Republican's vote totals, divided by the number of voters, n). If you vote, that will add $+1/n$ or $-1/n$ to V; the decisiveness of this vote is is 0 (if $|V| > 1/n$), 1/2 (if $|V| = 1/n$), or 1 (if $V = 0$).

Now suppose that that the proportional vote differential has an approximate continuous probability distribution, $p(V)$. This distribution can come from a theoretical model of voting (e.g., the random voting model discussed above) or empirical models based on election results or forecasts. Gelman, King, and Boscardin (1998) argue that, for modeling voting decisions, it is appropriate to use probabilities from forecasts, since these represent the information available to the voter before the election occurs. For retrospective analysis, it may also be interesting to use models based on perturbations of actual elections as in Gelman and King (1994). In any case, all that is needed here is some probability distribution.

For any reasonably-sized election, we can approximate the distribution $p(V)$ of the proportional vote differential by a continuous function. In that case, the expected probability of decisiveness is simply $2p(V)/n$ evaluated at the point $V = 0$. (If the number of voters n is odd, this approximates $\Pr(V = 0)$, and if n is even, it approximates $\frac{1}{2}\Pr(V = -1/n) + \frac{1}{2}\Pr(V = 1/n)$.) For example, in a two-candidate election with 10,000 voters, if one candidate is forecast to get 54% of the vote with a standard error of 3%, then the vote differential is forecast at 8% with a standard error of 6%. The probability that an individual vote is decisive is then $2\frac{1}{\sqrt{2\pi}(0.06)}\exp(-\frac{1}{2}(0.08/0.06)^2)/10000 = 0.0055$, using the statistical formula for the normal distribution.

The same ideas apply for more complicated elections, such as multicandidate contests, runoffs, and multistage systems (e.g., the Electoral College in the U.S. or the British parliamentary system in which the goal is to win a majority of individually-elected seats). In more complicated elections, it is simply necessary to specify a probability model for the entire range of possible outcomes, and then work out the probability of the requisite combination events under which a vote is decisive. For example, in the Electoral college, your vote is decisive if your state is tied (or within one vote of tied) and if, *conditional on your state being tied*, no candidate has a majority based solely on the other states. Estimating the probability of this event requires a model for the joint distribution of the vote outcomes in all the states (see Gelman, King, and Boscardin 1998).

4.2 What if an individual vote is never a decisive event?

As illustrated by the Presidential election in Florida in 2000, an election can be disputed even if the votes are not exactly tied. This may seem to call into question the very concept of a decisive vote. Given that elections can be contested and recounted, it seems naive to suppose that the difference between winning and losing is no more than the change in a vote margin from $+1$ to -1, which is what we have been assuming.

In fact, our decisive-vote calculations are reasonable, even for real elections with disputed votes, recounts, and so forth. We show this by setting up a more elaborate model that allows for a gray area in vote counting, and then demonstrating that the simpler model of decisive votes is a reasonable approximation.

As in the previous section, consider a two-candidate election and label V as the difference between the true number of votes received by each of the two candidates, divided by the number of votes, n. We model vote-count errors, disputes, etc., by defining $\pi(V)$ as the probability that the "+" option wins, given a true vote differential of V. With perfect voting, $\pi(V) = 0$ if $V < 0$, 1 if $V > 0$, or 0.5 if $V = 0$. More realistically, $\pi(V)$ is a function of V which equals 0 if V is clearly negative (e.g., $V < -0.001$), 1 if V is clearly positive (e.g., greater than one-tenth of one percent), and is between 0 and 1 if V is near 0.

In that case, the probability that your vote determines the outcome of the election, conditional on V, is $\pi(V + \frac{1}{n}) - \pi(V - \frac{1}{n})$. If your uncertainty about V is summarized by a probability distribution, $p(V)$, then your probability of decisiveness is,

$$\begin{aligned} \text{voting power} &= \mathrm{E}\left[\pi(V + \frac{1}{n}) - \pi(V - \frac{1}{n})\right] \\ &= \sum_V \left[\pi(V + \frac{1}{n}) - \pi(V - \frac{1}{n})\right] p(V). \qquad (3) \end{aligned}$$

At this point, we make two approximations, both of which are completely reasonable in practice. First, we assume that the election will only be contested for a small range of vote differentials, which will lie near 0: thus, there is some small ϵ such that $\pi(V) = 0$ for all $V < -\epsilon$ and $\pi(V) = 1$ for all $V > \epsilon$. Second, we assume that the probability density $p(V)$ for the election outcome has an uncertainty that is greater than ϵ (for example, perhaps $\epsilon = 0.001$ and V can be anticipated to an accuracy of 2%, or 0.02). Then we can approximate $p(V)$ in the range $0 \pm \epsilon$ by the constant $p(0)$. Expression (3) can then be written as,

$$\begin{aligned} \text{voting power} &= \int_{-\epsilon}^{\epsilon} \left[\pi(V + \frac{1}{n}) - \pi(V - \frac{1}{n})\right] p(0) dV \\ &= p(0) \int_{-\epsilon}^{\epsilon} \left[\pi(V + \frac{1}{n}) - \pi(V - \frac{1}{n})\right] dV \\ &= p(0) \left[\int_{-\epsilon+\frac{1}{n}}^{\epsilon+\frac{1}{n}} \pi(V) dV - \int_{-\epsilon-\frac{1}{n}}^{\epsilon-\frac{1}{n}} \pi(V) dV\right] \\ &= p(0) \left[\int_{\epsilon-\frac{1}{n}}^{\epsilon+\frac{1}{n}} \pi(V) dV - \int_{-\epsilon-\frac{1}{n}}^{-\epsilon+\frac{1}{n}} \pi(V) dV\right] \\ &= p(0) \left[\frac{2}{n} \cdot 1 - \frac{2}{n} \cdot 0\right] \\ &= 2p(0)/n, \end{aligned}$$

which is the same probability of decisiveness as calculated assuming all votes are recorded correctly.

5 Empirical results for the Electoral College

Perhaps the most frequently-considered example of voting power in elections (as distinguished from voting within a legislature) is the Electoral College system for the President of the United States. How does the probability of a decisive vote vary among states? For your vote to be decisive, two conditions must hold: (1) your state must be tied (or within one vote of a tie), and (2) the electoral votes of the other states must be close enough to a tie that your state's electoral votes are needed for a majority. One would expect the first condition to be more likely for small states and the second condition more likely for large states.

Gelman, King, and Boscardin (1998) computed these probabilities using a forecasting model for state-by-state Presidential election results from 1948 to 1992. The estimated probability of decisiveness varied based on several factors, including: (a) the closeness of the general election (the average probability of decisiveness was about 1 in a million in close election years such as 1960 and 1976, and about 1 in 100 million in landslides such as in 1964 and 1984); (b) the "median-ness" of the state (for example, in 1992, the probability of a vote being decisive was estimated to be about 10 times higher in Vermont, which was near the national median in vote preferences, than Utah, which is a strongly Republican state); and (c) the size of the state, with the probability of decisiveness being slightly higher, on average, in the smallest states. Figure 2 displays the estimated average probability of a vote being decisive versus the number of electoral votes in the individual's state, with each year from 1948 to 1992 shown by a different line.

In relation to the political science literature, the most important result in Figure 2 is the extremely low probability that a vote is decisive, even in the most favorable conditions of small states near the national median in a close election. The other important result is the relatively weak relation between the probability of a vote being decisive and the size of the state.

The result shown in Figure 2—that a voter in a small state is more likely to be decisive that a voter in a medium-sized or large state—contradicts a long-established claim in the political science literature that the Electoral College benefits voters in larger states. For example, Banzhaf (1968) claims to offer "a mathematical demonstration" that the Electoral College system "discriminates against voters in the small and middle-sized states by giving the citizens of the large states an excessive amount of voting power," and Brams and Davis (1974) claim that the voter in a large state "has on balance greater potential voting power ... than a voter in a small state." Owen (1975) and Rabinowitz and MacDonald (1986) come to similar conclusions.

Why do their findings differ from ours? Their calculations are based on the random-voter model or variants of it, and always with the assumption of independent and indistinguishable voters within each state, so that the standard deviation of the vote differential within any state j with m_j voters is proportional to $\sqrt{m_j}$. The probability that the state itself will be decisive in the national total is approximately proportional to e_j, the number

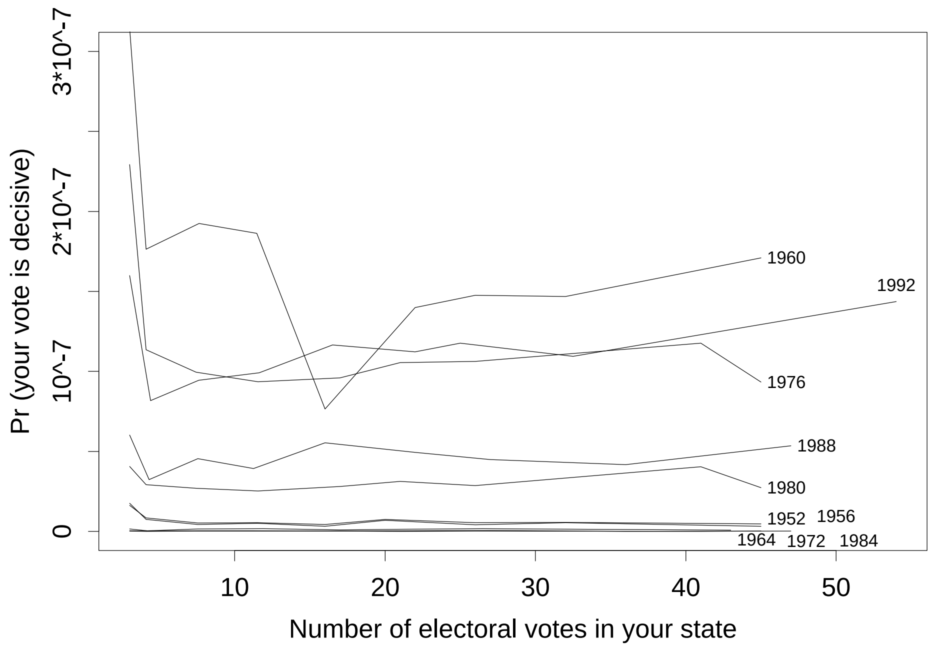


Figure 2: *The probability of a decisive vote as a function of the number of electoral votes in the voter's state, for each U.S. Presidential election from 1948–1992. The probabilities are calculated based on a forecasting model that uses information available two months before the election. This figure is adapted from Gelman, King, and Boscardin (1998). The most notable features of this figure are: first, that the probabilities are all very low; and second, that the probabilities vary little with state size, with the only consistent pattern being that voters in the very smallest states are a bit more likely to be decisive.*

of electoral votes in the state. The voting power of an individual voter in the state is then approximately proportional to $e_j/\sqrt{m_j}$, which is highest for the largest states.

The key assumption here is that the vote differential has a standard deviation of $\sqrt{m_j}$, which means that the *proportion* of the vote for either candidate within state j has a standard deviation of $1/\sqrt{m_j}$. Thus, as Banzhaf (1968), Brams and Davis (1974; 1975, p. 155), and Owen (1975, p. 953), make clear, the claim that voters in large states are more likely to be decisive is a consequence of the claim that elections in large states are more likely to be close.

In fact, however, this is not the case, or at least not to the extent implied by the $1/\sqrt{m_j}$ rule. For example, in the most recent Presidential election, none of the three largest states (California, Texas, and New York) was close, and it was in fact recognized before the election that the voters in these states had little chance of influencing the outcome. To be more systematic, we extend an analysis by Colantoni, Levesque, and Ordeshook (1975, pp. 144–145) and display in Figure 2 the vote differentials as a function of number of voters for all states (excluding the District of Columbia) for all elections from 1960 to

2000. On average, larger states have somewhat closer elections, but the pattern is very weak, as can be seen by the fitted lowess curve on the plot. By comparison, the graph also displays a curve proportional to $1/\sqrt{m_j}$, which shows how the absolute vote differential would be expected to decrease with state size if the voting power measures based on the random voting model were appropriate.

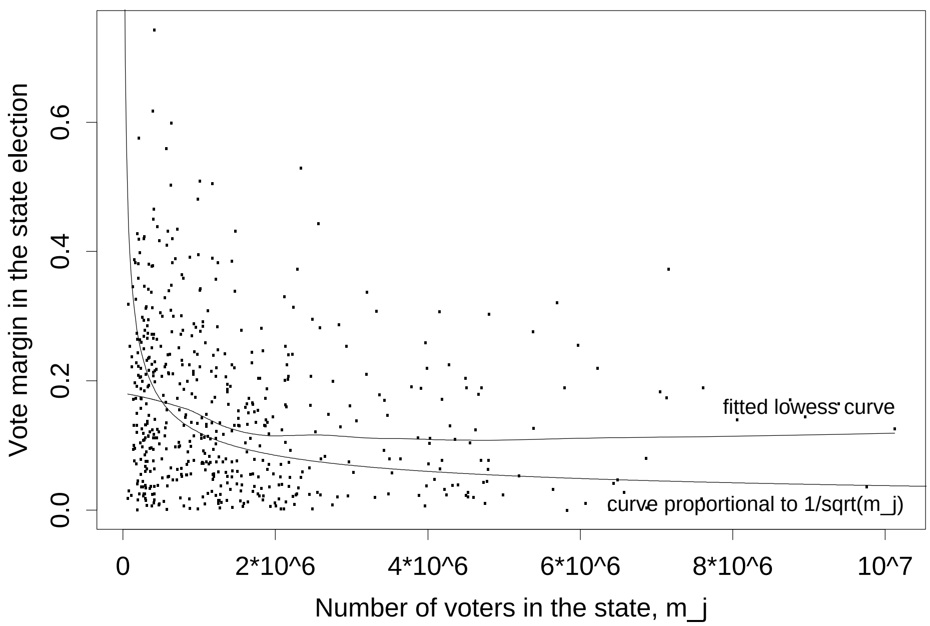


Figure 3: *The margin in state votes for President as a function of the number of voters m_j in the state: each dot represents a different state and election year from 1960–2000. The margins are proportional; for example, a state vote of 400,000 for the Democratic candidate and 600,000 for the Republican would be recorded as 0.2. Elections tend to be closer in larger states, but this pattern is very weak, as is shown by the nearly flat slope of the nonparametric (lowess) line displayed on the graph. By comparison, a line proportional to $1/\sqrt{m_j}$ is also shown; clearly, this theoretical curve does not fit the data.*

For another attack at the problem, Gelman, King, and Boscardin (1998) estimate a model with an error term for each state with a standard deviation proportional to $m_j^{-\theta/2}$, and the parameter θ is estimated to be about 0.2 (using data from 1948 to 1992), as compared to $\theta = 1$ which would arise from the random voting model.

Thus, the earlier claims that large states benefit from the Electoral College were mistaken because of their implicit assumption that elections in larger states would be much closer than those in small states. Although this $1/\sqrt{m_j}$ rule does *not* empirically apply to Presidential elections, it might hold in other elections or decision-making settings, in which case results such as Banzhaf (1965) could be reasonable. Empirical studies (such as that of Heard and Swartz (1999) for the Canadian Supreme Court) are needed to answer these questions.

5.1 Comparing the average probability of decisiveness under different electoral systems

We can also see if our theoretical finding that the average probability of decisiveness should be larger under straight popular vote than under the Electoral College system holds empirically by looking at past U.S. Presidential elections. That is, for each election since 1960, we compare the average probability of decisiveness for voters under the popular vote and electoral vote systems. We also compute the average probability of decisiveness under an alternative system in which each Congressional district is worth one electoral vote: that is, winner-take all in each of the 436 districts (counting D.C. as a district). In order to estimate probabilities of close elections and decisiveness, it is necessary to set up a probability model for vote outcomes. We want to go beyond the random voting model and set up a more realistic descriptor of vote outcomes. Gelman, King, and Boscardin (1998) fit a state-by-state election forecasting model, with probabilities corresponding to the predictive uncertainty two months before the election. Here, we use a simpler approach: we take the actual election outcome and perturb it, to represent possible alternative outcomes.

We label v_i as the observed outcome (the Democratic candidate's share of the two-party vote) in Congressional district i in a given election year and obtain a probability distribution of hypothetical election outcomes y_i by adding normally-distributed random errors at the national, regional, state, and Congressional-district levels, with a standard deviation of 2% at each level. We label n_i as the turnout in each district i and consider these as fixed—this is reasonable since uncertainty about election outcomes is driven by uncertainty about v, not n.

For any given election year, we use the multivariate normal distribution of the vector v of vote outcomes to compute the probability of a single vote being decisive in the election. For the popular vote system, we determine this probability for any voter; for the electoral-vote and congressional-district-vote systems, we determine the probability within each state or district and then compute national average probabilities, weighing by turnouts within states or districts. The actual probability calculations are done using the multivariate normal distribution as described by Gelman, King, and Boscardin (1998).

Our results appear in Figure 4. The most striking feature of the figure is that the average probability of decisiveness changes dramatically from year to year but is virtually unaffected by changes in the electoral system. This may come as a surprise—given the theoretical results from Section 3, one might expect the average probability of decisiveness to be much higher for the popular-vote system.

The results in Figure 4 are only approximate, not just because of the specific modeling choices made, but also because of the implicit assumption that the patterns of voting would not be affected by changes in the electoral system. For example, states such as California and Texas that were not close in the 2000 election might have had higher turnout under a popular vote system in which all votes counted equally. Thus, our

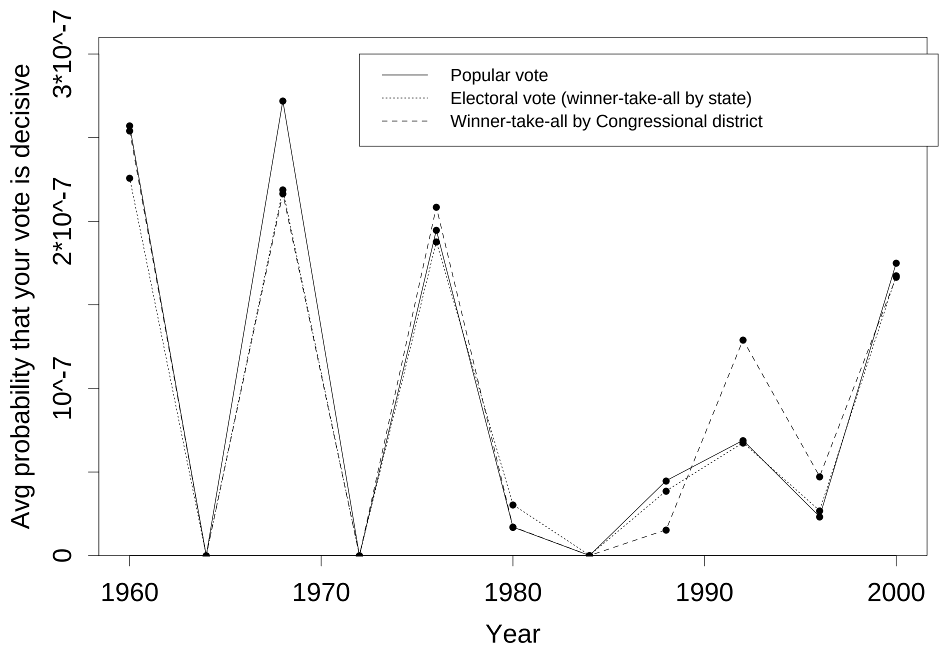


Figure 4: *The estimated average probability of a decisive vote, for each presidential election since 1960, calculated under the popular and electoral vote systems. The estimates are based on simulations from actual election outcomes rather than the forecasting model used in producing Figure 2. The average probability varies a lot by year but is not much affected by the electoral rule.*

results compare different electoral systems as applied to the actual observed votes and do not directly address counterfactual questions about what would happen if the electoral system were changed.

6 Discussion

As noted in Section 2, it makes sense for any individual to form a coalition, but as proved in Section 3, this lowers the average voting power for everybody. Coalitioning is thus like a prisoner's dilemma, which without constraints will lead to a stable equilibrium that is unfavorable to all (see, e.g., Luce and Raiffa 1957). These results hold under the random voting model, which is implicit in general power indexes such as those of Banzhaf (1965) and Shapley and Shubik (1954). One-person, one-vote does not mean that all voters have equal voting power.

For actual elections, the probability of decisiveness must be assessed using an empirical model of votes. In any such model, the key determinants of voting power are (1) the anticipated closeness and the uncertainty about the election outcome (which jointly determine the probability that the vote will be close to tied) and (2) the number of voters

m (because the probability of an exact tie is proportional to $1/m$). A careful analysis of data from U.S. Presidential elections shows that a vote from a small state is, on average, more likely to be decisive than a vote from larger states. Previous published results claiming the opposite suffered from inappropriate models (explicit or implicit) for the closeness of elections as a function of state size. In reality, voters in small states have disproportionately high voting power because of the extra two electoral votes received by each state.

The results for the Electoral College are interesting and suggest various lines of further research, including analysis of other possible Presidential electoral systems, state and local elections, and comparisons to other countries (including those with parliamentary systems) and other time periods. In a slightly different direction, it would be interesting to analyze the votes of legislators in subcommittees, committees, and roll-call votes, in order to study the empirical effects of structural coalitions on voting power in legislatures. Conversely, empirical results that do not follow directly from the theory (as in Figure 4) suggest ways in which a theory of voting in coalitions must be improved in order to be realistic.

Of course, our mathematical and empirical findings do not directly address normative questions such as: Which electoral system should be used? Or in a legislature, how should committees or subcommittees be assigned? Let alone more fundamental questions such as, is it desirable for the average voting power to be increased? After the 2000 election, some commentators suggested that it would be better if close elections were *less* likely, even though close elections are associated with decisiveness of individual votes, which seems like a good thing.

The issue of the desirability of close elections raises a conflict between two political principles: on one hand, *democratic process* would seem to require that every person's vote has a nonzero chance (and, ideally, an equal chance) of determining the election outcome. On the other hand, very close elections such as Florida's damage the *legitimacy* of the process, and so it might seem desirable to reduce the probability of ties or extremely close votes.

No amount of theorizing will resolve this difficulty, which also occurs in committees and leads to legitimacy-protecting moves such as voting with an informal straw poll. The official vote that follows is then often close to unanimous as the voters on the losing side switch to mask internal dissent. This paper's theoretical findings imply that such behavior is understandable but in a larger context can reduce the average voting power of individuals.

We conclude by recalling that individual measures of political choice, even if aggregated, cannot capture the structure of group power. For example, groups that can mobilize effectively are solving the coordination problem of voting and can thus express more power through the ballot box. For an extreme example, consider the case of Australia, where at one time Aboriginal citizens were allowed, but not required, to vote in national elections, while non-Aboriginal citizens were required to vote. Unsurprisingly, turnout was lower among Aboriginals. Who was benefiting here? From an individual-rights standpoint, the Aboriginals had the better deal, since they had the freedom to choose whether to vote. But, as a group, the Aboriginals' lower turnout would be expected to hurt their representation in the government and thus, probably, hurt them individually as well. Having voting power is most effective when you actually vote.

References

Aldrich, J. H. (1993). Rational choice and turnout. *American Journal of Political Science* **37**, 246–278.

Arrow, K. (1951). *Social Choice and Individual Values*. New York: Wiley.

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

Banzhaf, J. R. (1968). One man, 3.312 votes: a mathematical analysis of the Electoral College. *Villanova Law Review* **13**, 304–332.

Beck, N. (1975). A note on the probability of a tied election. *Public Choice* **23**, 75–79.

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

Brams, S. J., and Davis, M. D. (1975). Comment on "Campaign resource allocation under the electoral college," by Colantoni, C. S., Levesque, T. J., and Ordeshook, P. C.. *American Political Science Review* **69**, 155–156.

Chamberlain, G., and Rothchild, M. (1981). A note on the probability of casting a decisive vote. *Journal of Economic Theory* **25**, 152–162.

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.

Ferejohn, J., and Fiorina, M. (1974). The paradox of not voting: a decision theoretic analysis. *American Political Science Review* **68**, 525.

Finkelstein, M. O., and Levin, B. (1990). *Statistics for Lawyers*, 66–70. New York: Springer-Verlag.

Gelman, A., and King, G. (1994). A unified model for evaluating electoral systems and redistricting plans. *American Journal of Political Science* **38**, 514–554.

Gelman, A., King, G., and Boscardin, W. J. (1998). Estimating the probability of events that have never occurred: when is your vote decisive? *Journal of the American Statistical Association* **93**, 1–9.

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

Luce, R. D., and Raiffa, H. (1957). *Games and Decisions*. New York: Wiley.

Mann, I., Shapley, L. S. (1960). The a priori voting strength of the electoral college. RAND memo reprinted in Shubik, ed., *Game Theory and Related Approaches to Social Behavior*, New York: Wiley, 1964.

Margolis, H. (1977). Probability of a tie election. *Public Choice* **31**, 134–137.

Margolis, H. (1983). The Banzhaf fallacy. *American Journal of Political Science*, 321–326.

Merrill, S. (1978). Citizen voting power under the Electoral College: a stochastic model based on state voting patterns. *SIAM Journal of Applied Mathematics* **34**, 376–390.

Owen, G. (1975). Evaluation of a presidential election game. *American Political Science Review* **69**, 947–953.

Rabinowitz, G., and MacDonald, S. E. (1986). The power of the states in U.S. presidential elections. *American Political Science Review* **80**, 65–87.

Riker, W. H., and Ordeshook, P. C. (1968). A theory of the calculus of voting. *American Political Science Review* **62**, 25–42.

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.

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

Woodroofe, M. (1975). *Probability with Applications*. New York: McGraw Hill.