B.J.Pol.S. **34,** 657–674 *Copyright © 2004 Cambridge University Press*
DOI: 10.1017/S0007123404000237 *Printed in the United Kingdom*

Standard Voting Power Indexes Do Not Work: An Empirical Analysis

ANDREW GELMAN, JONATHAN N. KATZ AND JOSEPH BAFUMI*

Voting power indexes such as that of Banzhaf are derived, explicitly or implicitly, from the assumption that all votes are equally likely (i.e., random voting). That assumption implies that the probability of a vote being decisive in a jurisdiction with n voters is proportional to $1/\sqrt{n}$. In this article the authors show how this hypothesis has been empirically tested and rejected using data from various US and European elections. They find that the probability of a decisive vote is approximately proportional to $1/n$. The random voting model (and, more generally, the square-root rule) overestimates the probability of close elections in larger jurisdictions. As a result, classical voting power indexes make voters in large jurisdictions appear more powerful than they really are. The most important political implication of their result is that proportionally weighted voting systems (that is, each jurisdiction gets a number of votes proportional to n) are basically fair. This contradicts the claim in the voting power literature that weights should be approximately proportional to $\sqrt{n}$.

1 INTRODUCTION

1.1 Voting Power and Fairness in Weighted Voting Systems

Recent events such as the 2000 US presidential election and the expansion of the European Union (EU) have rekindled interest in evaluating electoral systems. Both the US Electoral College system for electing the president and the European Union's Council of Ministers, in which the representative from each country gets some specified number of votes, are examples of weighted or asymmetric voting systems. The US Senate can also be considered an asymmetric voting system, since the number of people represented by each senator varies greatly from state to state. In these asymmetric systems, voters have a potentially differential impact on electoral outcomes. A natural question that arises, therefore, is whether or not a particular system is politically fair.

In a weighted voting system, two aspects of voting power are of potential interest: (a) the voting power of a particular member of the legislature (or, of a particular state in the Electoral College, or a particular country represented in the EU), and (b) the power of an individual voter. The first aspect of voting power is relevant for understanding how the legislature works, and the second aspect relates to the fairness of the system with respect to the goal of representing people equally, an issue that is politically salient given current plans to expand the European Union and proposals to alter the voting structures of International Monetary Fund, the United Nations, and other international institutions.[1]

In this article we reconsider how to measure voting power in weighted voting systems.

* Gelman and Bafumi: Department of Political Science, Columbia University (Gelman is also in the Department of Statistics); Katz: Division of the Humanities and Social Sciences, California Institute of Technology. We thank David Park, the Editor of the Journal and several reviewers for helpful comments and the US National Science Foundation for grants SES-9987748, SES-0084368 and SES-0318115.

[1] For example, see D. S. Felsenthal and M. Machover, 'Enlargement of the EU and Weighted Voting in Its Council of Ministers', Report VPP 01/00 (2000), Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science; D. Leech, 'Designing the Voting System for the Council of the

The measurement of voting power and its consequences for constitutional design is an old question in political science.[2] Our contribution to this long debate is twofold. First, we show that the 'standard' measures of voting power make assumptions that are not supported empirically and, more importantly, lead to incorrect inferences about the distribution of voting power in actual electoral systems. Secondly, we *empirically* estimate voting power from observed election data. When such data are available, this empirical measure should be preferred since it does not require these problematic assumptions to be made.

1.2 A Priori *Measures of Voting Power*

Voting power can be defined and measured in many different ways,[3] with most measures based on the probability that a voter or a member of a legislative body is pivotal – that is, casts a vote that would change the outcome of the election. These 'standard' measures are considered *a priori* measures, in the sense that they calculate voting power under a given electoral system without reference to past or anticipated voting patterns in any particular elections. All the 'standard measures' of theoretical voting power yield the counterintuitive result that, in a proportional voting system, voters in large districts tend to have disproportionate power. Thus, it has been claimed that voters in large states have more power in the US Electoral College,[4] and that, if EU countries were to receive votes in the Council of Ministers proportional to their countries' populations, then voters in large countries would have disproportionate power.[5] These claims are controversial and are defended with mathematical arguments. However, these arguments are ultimately based on assumptions that can be checked, and falsified, with real data.[6]

1.3 Empirical Voting Patterns and Their Implications for A Priori *Measures of Voting Power*

This article presents empirical findings on the probability of a vote being decisive, for voters in jurisdictions of different sizes. The most important political implication of our findings is that proportional weighting systems are, in fact, basically fair to all voters, and alternative systems that have been recommended in the voting power literature – for example, giving each jurisdiction a vote proportional to the square root of the number of people represented – are unfair to voters in large jurisdictions. This is the same conclusion reached, from a game-theoretic argument, by Snyder, Ting and Ansolabehere.[7] This

(F'note continued)
European Union', *Public Choice*, 113 (2002), 437–64; and D. Leech, 'Voting Power in the Governance of the IMF', *Annals of Operations Research*, 109 (2002), 373–95.

[2] A complete history is beyond the scope of this article, but see D. S. Felsenthal and M. Machover, *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes* (Northampton, Mass.: Edward Elgar, 1998), for a brief history.

[3] See, for example, Felsenthal and Machover, *The Measurement of Voting Power*, and D. G. Saari and K. K. Sieberg, 'Some Surprising Properties of Power Indices', *Games and Economic Behavior*, 36 (2000), 241–63.

[4] J. R. Banzhaf, 'One Man, 3.312 Votes: A Mathematical Analysis of the Electoral College', *Villanova Law Review*, 16 (1968), 304–32.

[5] Felsenthal and Machover, 'Enlargement of the EU'. See also L. S. Penrose, 'The Elementary Statistics of Majority Voting', *Journal of the Royal Statistical Society* 109 (1946), 53–7.

[6] As is noted by A. D. Heard and T. B. Swartz, 'Extended Voting Measures', *Canadian Journal of Statistics*, 27 (1999), 177–86.

[7] J. M. Snyder, M. M. Ting and S. Ansolabehere, 'Legislative Bargaining Under Weighted Voting' (Technical report, Department of Economics, Massachusetts Institute of Technology, 2001).

conclusion is particularly important because large jurisdictions are already inherently underrepresented in political organizations such as the US Senate and the UN General Assembly that give the same number of votes to each jurisdiction.

Voting power indexes have been criticized before, largely from the direction that they do not capture idiosyncrasies in any given electoral system.[8] By their very nature, standard voting power measures rely only on the mathematical rules of a voting system and not on past or anticipated future patterns of voting within the system. What we report here is new in that it gathers data from a wide variety of elections to show the inappropriateness of the mathematical model underlying the usual measures of voting power. We also explain from a theoretical perspective why voting power measures are inconsistent with modern models of public opinion and electoral politics.

The article proceeds as follows. In the next section we review the mathematical model underlying standard voting power measures and derive the empirical implications of this model. In Section 3, we examine the predictions of the voting power model from data from the Electoral College system for electing the president of the United States. In Section 4, we examine data for other elections from the United States and Europe. Then in Section 5, we consider theoretical reasons why the voting power model fails in actual election data. We then conclude in Section 6, arguing for the relevance of empirical analysis to the evaluation of a priori voting power measures.

2 THE MATHEMATICAL MODEL UNDERLYING VOTING POWER MEASURES

In this section we briefly review the mathematical model underlying standard voting power measures.[9] We begin by defining what we mean by a standard voting measure.

First, a standard voting index or measure is one based on the probability of casting a decisive ballot in an election, what Felsenthal and Machover refer to as influence or I-power indexes. These differ from measures based on coalitional bargaining, such as the Shapely–Shubik index, which are referred to as P-power indexes by Felsenthal and Machover. The P-power indexes attempt to measure the amount of spoils a legislator or voter will receive from a given electoral system.[10] These P-power indexes have typically not been used to evaluate individual voters when electing representatives, such as for the US Senate or European Parliament. Secondly, standard indexes use the *random voting model* as their assumption of electoral behaviour under a given electoral rule. Under the random voting model, a voter is assumed to be equally likely to choose any of the alternatives on the ballot. We discuss the implications of this model in the rest of this section.

2.1 Defining the Power of an Individual Voter

More formally, in this article we shall consider elections with two parties (A and B) and voters in jurisdictions $j = 1, \dots J$. Each jurisdiction j has n_j voters and is given e_j 'electoral votes', the vote in each jurisdiction is chosen winner-take-all, and the total winner is the

[8] For example, see G. Garrett, and G. Tsebelis, 'Why Resist the Temptation to Apply Power Indices to the European Union?' *Journal of Theoretical Politics*, 11 (1999), 291–308.

[9] For a more thorough discussion, see Felsenthal and Machover, *The Measurement of Voting Power*.

[10] See Felsenthal and Machover, *The Measurement of Voting Power*, for complete discussion of the types of voting power indexes.

party with more electoral votes, with ties at all levels broken by coin tosses. We further define v_j as party A's share of the vote in jurisdiction j; E_{-j} as the total number of electoral votes, excluding those from jurisdiction j, that go for party A; and

$$E = \Sigma_{j=1}^{J} e_j$$

as the total number of electoral votes in the system.[11]

Voting power can be defined in various ways, but the definition most relevant to representation of voters is in terms of the probability that a voter is decisive.[12] At the top level, the voting power of jurisdiction j is the probability that party A wins if jurisdiction j goes for party A, minus the probability that party A wins if jurisdiction j goes for party B. Next, the power of a given *voter* in jurisdiction j is the probability that party A wins if that voter supports A, minus the probability that A wins if that voter supports B.

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. 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.[13] Perhaps the simplest measure of decisiveness is the (absolute) Banzhaf 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 parties.[14]

In general, the key step in defining voting power is assigning a probability distribution over all possible voting outcomes. The Banzhaf index and related measures are sometimes defined in terms of game theory and sometimes in terms of set theory, but they can all be interpreted in terms of probability models. For example, suppose that your voting power is defined as the number of coalitions of other voters for which your vote is decisive. This is simply proportional to the probability of decisiveness, under the assumption that all vote outcomes are equally likely – that is, the random voting model.

[11] We use the US Electoral College as a template for our analysis but our findings apply equally well to organizations such as the EU Council of Ministers in which each member represents a constituency j of size n_j. In such settings, we are ultimately interested in the power of the individual voters within the constituencies, which is determined by their power to determine their representative (that is, to affect the outcome of a vote within the jurisdiction) and the power of that representative (using his or her weighted vote w_j) in the larger council. The voting power of an individual citizen has the same two-stage mathematical structure (as described in this section) whether the aggregation is by representatives as in the Council of Ministers or (nearly) automatic as in the Electoral College.

[12] See Heard and Schwartz, 'Extended Voting Measures', and Felsenthal and Machover, *The Measurement of Voting Power*.

[13] W. H. Riker and P. C. Ordeshook, 'A Theory of the Calculus of Voting', *American Political Science Review*, 62 (1968), 25–42; J. Ferejohn and M. Fiorina, 'The Paradox of Not Voting: A Decision Theoretic Analysis', *American Political Science Review*, 68 (1974), 525–36; S. J. Brams and M. D. Davis, 'The 3/2's Rule in Presidential Campaigning', *American Political Science Review*, 68 (1974), 113–34; and J. H. Aldrich, 'Rational Choice and Turnout', *American Journal of Political Science*, 37 (1993), 246–78.

[14] J. R. Banzhaf, 'Weighted Voting Doesn't Work: A Mathematical Analysis', *Rutgers Law Review*, 19 (1965), 317–43. The index had also been proposed by L. S. Penrose, 'The Elementary Statistics of Majority Voting', *Journal of the Royal Statistical Society*, 109 (1946), 53–7, and others; see Felsenthal and Machover, 'Enlargement of the EU'. Section 1.2.3, for a historical overview.

As we discuss below, the random voting model has strong implications for voting power. It also has strong implications for actual votes – and these implications do not fit reality. The later parts of this article discuss the implications for voting power of that lack of fit.

2.2 What Does the Random Voting Model Imply about Voting Power?

For an individual in jurisdiction j, his or her vote is decisive if (a) the e_j electoral votes of jurisdiction j are decisive in the larger election, and (b) the individual's vote is decisive in the election within the jurisdiction. Using conditional probability notation, this can be written as,

$$\text{individual voting power} = \Pr\,(\text{jurisdiction } j\text{'s } e_j \text{ electoral votes are decisive}) \times$$

$$\times \Pr\,(\text{a given vote is decisive in jurisdiction } j \mid \text{jurisdiction } j\text{'s } e_j \text{ electoral votes are decisive}).$$

We can write each of these in our notation, keeping in mind that ties are decided by coin flips,

$$\Pr\,(\text{jurisdiction } j\text{'s } e_j \text{ electoral votes are decisive}) =$$

$$\Pr(E_{-j} \in (0.5E - e_j, 0.5E)) + \frac{1}{2}\Pr(E_{-j} = 0.5E - e_j) + \frac{1}{2}\Pr(E_{-j} = 0.5E) \quad (1)$$

$$\Pr\,(\text{a given vote is decisive in jurisdiction } j) = p_{v_j}\,(0.5)/n_j, \quad (2)$$

where p_{v_j} is the continuous probability density assigned to the vote for party A in jurisdiction j. (We are assuming that n_j in any district is large enough – greater than 100, say – so that the model for the v_j's can be approximated by a continuous distribution. A derivation of this approximation, under general conditions, appears in the Appendix.)

Under the random voting model, the two events, (1) and (2), are independent, so we can evaluate the probabilities separately, which we now do.

2.2.1 The probability that a jurisdiction is decisive, under the random voting model. Assuming random voting, one can directly calculate the probability that jurisdiction j's electoral votes are decisive by assigning a probability $1/2^{J-1}$ to each of the configurations of the other $J-1$ jurisdictions, which in turn induces a distribution on E_{-j}, so that Equation 1 can be calculated. In specific cases, the results can reveal important properties of the electoral system. We illustrate with a simple example with $J = 4$. Suppose $(e_1, e_2, e_3, e_4) = (12, 9, 6, 2)$. Then the fourth jurisdiction has zero voting power – its two electoral votes can never determine the winner. In addition, the first three jurisdictions each have equal voting power of 1/2 – any of these jurisdictions will be decisive if the other two are split. In this example, the relation between electoral votes and voting power is far from linear.[15]

If the number of jurisdictions is large, however, with no single jurisdiction dominating, and no unusual patterns (such as in the example above in which all but one of the e_j's are divisible by 3), then it is possible to approximate the probabilities in Factor 1 using a continuous distribution.[16] One can then approximate (1) by $\int_{0.5E-e_j}^{0.5E} p(E_{-j})dE_{-j}$. If the

[15] Some examples of this phenomenon in actual voting systems are provided by Felsenthal and Machover, *The Measurement of Voting Power*, and D. Leech, 'The Utility of the Voting Power Approach', part of a forum of responses by Felsenthal and others, 'In Defence of Voting Power Analysis', *European Union Politics* 4 (2003).

[16] This would be most simply done using the normal distribution, but other forms are possible, such as the scaled beta distribution used by A. Gelman, G. King, and W.J. Boscardin, 'Estimating the Probability of Events That Have Never Occurred: When Is Your Vote Decisive?' *Journal of the American Statistical Association*, 93 (1998), 1–9.

further assumption is made that e_j is small compared to the uncertainty in E_{-j}, then Pr (jurisdiction j's e_j electoral votes are decisive) will be approximately proportional to e_j.

In order to see that this approximation is reasonable, consider the Electoral College. In this case the values of e_j for the fifty states and the District of Columbia range from 3 to 54, with a total of 538. We can calculate directly Pr (jurisdiction j's e_j electoral votes are decisive) for each state assuming random voting and compare to the linear approximation we suggest. The linear fit has a relative error of less than 10 per cent for all states.

For the rest of this article we shall assume that Factor 1 is proportional to e_j. Computing Factor 1 more precisely in special cases is potentially important, but such details do not affect the main point of this article.

2.2.2 The probability that a voter is decisive within a jurisdiction, under the random voting model. The key way that the random voting model affects the calculation of voting power is through Factor 2, the probability that a vote is decisive within a jurisdiction.

Under random voting, the distribution p_{v_j} of the vote share among the n_j voters in jurisdiction j is approximately normally distributed with mean 0.5 and standard deviation $0.5/\sqrt{n_j}$, hence the approximation,

$$\text{Pr (a given vote is decisive in jurisdiction } j) = p_{v_j}\,(0.5)/n_j = \sqrt{\frac{2}{\pi}}\,\frac{1}{\sqrt{n_j}}.$$

What matters for comparing voters is not the constant factor $\sqrt{2/\pi}$ but the proportionality with $1/\sqrt{n_j}$.

2.2.3 The power of an individual voter under the random voting model. Now that Factors 1 and 2 have been approximated assuming random voting, they can be multiplied to yield a voting power approximately proportional to $e_j/\sqrt{n_j}$ for any individual in jurisdiction j.

Under the natural weighting system in which the electoral votes e_j are set proportional to n_j, an individual's voting power is then proportional to $\sqrt{n_j}$, and the sum of the voting powers of the n_j voters in the jurisdiction is proportional to $n_j^{3/2}$. Hence the titles of the papers by Banzhaf and by Brams and Davis.[17] Conversely, a suggested reform is to set e_j proportional to $\sqrt{n_j}$, so that individual voting power (assuming the random voting model) is approximately the same across jurisdictions.[18]

2.3 Probability Models for Voting: Going Beyond the Square-root Rule

As has been noted by many researchers, there are theoretical and practical problems with a model that models votes as independent coin flips (or, equivalently, that counts all possible arrangements of preferences equally).[19] The simplest model extension is to assume votes are independent but with probability p of going for party A, with some

[17] Banzhaf, 'One Man, 3.312 Votes'; Brams and Davis, 'The 3/2's Rule in Presidential Campaigning'.

[18] Penrose, 'The Elementary Statistics of Majority Voting', and Felsenthal and Machover, 'Enlargement of the EU'.

[19] See, for example, N. Beck, 'A Note on the Probability of a Tied Election', *Public Choice*, 23 (1975), 75–9; H. Margolis, 'Probability of a Tie Election', *Public Choice*, 31 (1977), 134–7; S. Merrill, 'Citizen Voting Power under the Electoral College: A Stochastic Model Based on State Voting Patterns', *SIAM Journal of Applied Mathematics*, 34 (1978), 376–90; and G. Chamberlain and M. Rothchild, 'A Note on the Probability of Casting a Decisive Vote', *Journal of Economic Theory*, 25 (1981), 152–62.

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 modelling this varying p is no easier than modelling vote outcomes directly. Actual election results can be modelled using regressions,[20] in which case the predictive distributions of the election outcomes can be used to estimate the probability of decisive votes. Gelman, King and Boscardin argue that, for modelling voting decisions, it is appropriate to use probabilities from forecasts, since these represent the information available to the voter before the election occurs.[21] For retrospective analysis, it may also be interesting to use models based on perturbations of actual elections.[22]

At this point, the idea of modelling vote outcomes seems daunting – there are so many different possible models and such a wealth of empirical data that it would seem impossible to make any general recommendations. Hence, researchers have argued in favour of the random voting model as a reasonable – or perhaps the only possible – *a priori* choice.

However, general *a priori* models other than random voting are possible. As discussed in the previous section, what is important for voting power is how p_{v_j} (0.5), the probability density of the vote proportion at 0.5, varies with n_j.

In particular, Good and Mayer and also Chamberlain and Rothchild derive a $1/n_j$ rule – that is, a model where the probability of decisiveness is inversely proportional to the number of voters.[23] This model arises by assuming that votes are binomially distributed, but with binomial probabilities p that themselves vary over the n_j voters in a jurisdiction. Then, for large or even moderate values of n_j, the distribution p_{v_j} of the vote proportion v_j is essentially fixed (not depending on n_j), so that p_{v_j} (0.5) is a constant, and the probability of a vote being decisive is proportional to $1/n_j$ (from (2)). The proportionality constant depends on the conditions of the election – but for comparing voting power across jurisdictions, all that matters is the proportionality with $1/n_j$.

A similar result arises if the probability distribution of votes is obtained from forecasts (whether from a regression-type model, subjective forecasts or some combination of the two). For example, in a two-party election with 10,000 voters, if one party is forecast to get 52 per cent of the vote with a standard error of 3 per cent, then the probability that an individual vote is decisive is approximate $[1/\sqrt{2\pi}(0.03)]\ \exp(-0.5(0.05/0.03)2)/n_j = 1.84/n_j$. We discuss some other models of votes at the end of Section 5.1.

2.4 Voting Power and the Closeness of Elections

To summarize, standard voting power measures are based on the assumption of random voting – or, more specifically, on the assumption that the probability of decisiveness within a jurisdiction is proportional to $1/\sqrt{n_j}$. This in turn corresponds to the assumption that p_{v_j} (0.5), the probability density function of the vote proportion v_j near 0.5, is proportional to

[20] See, for example, J. E. Campbell, 'Forecasting the Presidential Vote in the States', *American Journal of Political Science*, 36 (1992), 386–407.

[21] Gelman, King and Boscardin, 'Estimating the Probability of Events That Have Never Occurred'.

[22] See A. Gelman and G. King, 'A Unified Model for Evaluating Electoral Systems and Redistricting Plans', *American Journal of Political Science*, 38 (1994), 514–54.

[23] I. J. Good and L. S. Mayer, 'Estimating the Efficacy of a Vote', *Behavioral Science*, 20 (1975), 25–33; Chamberlain and Rothchild, 'A Note on the Probability of Casting a Decisive Vote'.

$\sqrt{n_j}$. Or, to say it another way, the key assumption is that elections are much more likely to be close (in percentage terms) when n_j is large.

In contrast, it is usual in forecasting elections to model the vote proportion v_j directly, with no dependence on n_j, in which case $p_{v_j}(0.5)$ does not depend on n_j and so the probability of decisiveness is proportional to $1/n_j$.

The models for votes have strong implications for voting power. The $1/\sqrt{n_j}$ model leads to the normative recommendation that, to equalize the voting power of all individuals, electoral votes e_j should be set roughly proportional to $\sqrt{n_j}$. In contrast, the model in which vote proportions do not depend on n_j implies that it is basically fair to set e_j roughly proportional to n_j.

Now that we have isolated the key question – how does the probability of a close election depend on n_j – we can explore it empirically in Sections 3 and 4, and theoretically in Section 5. The square-root rule has important practical implications. Such an assumption can and should be checked with actual data, not simply asserted.

2.5 Voting Power and Representation

Another way to look at voting power in a two-stage electoral system is in terms of the net number of voters whose opinions are carried by the representative. For example, if parties A and B receive 51 and 49 per cent of the vote, respectively, then party A has a net support of 2 per cent of the voters in that district. The number of electoral votes for jurisdiction j (or, more generally, its voting power) could then be set proportional to the absolute difference in votes between the two parties, which in our notation is $2n_j|v_j - 0.5|$. The vote v_j (and, to a lesser extent, the turnout, n_j) are random variables that are unknown after the election, and so it is natural to work with the expected net voters for the winning candidate in the jurisdiction, $\mathrm{E}(2n_j|v_j - 0.5|)$.

Under the random voting model, v_j has a mean of 0.5 and a standard deviation proportional to $1/\sqrt{n_j}$, and so the expected vote differential, $\mathrm{E}(2|v_j - 0.5|n_j)$, is proportional to $\sqrt{n_j}$. Penrose as well as Felsenthal and Machover use this reasoning to support the claim that large jurisdictions are overrepresented in proportional weighting.[24]

Conversely, if the proportional vote margin is independent of n_j, this supports proportional weighting and suggests a problem with the Banzhaf index and related voting power measures. We study the empirical relation between $\mathrm{E}(|v_j - 0.5|)$ and n_j in Sections 3 and 4.

3 DATA FROM THE ELECTORAL COLLEGE OF THE UNITED STATES

Perhaps the most frequently-considered example of voting power in elections (as distinguished from voting within a legislature) is for the president of the United States. The random voting model implies that the Electoral College benefits voters in large states, and this has been noted many times in the literature. For example, Banzhaf claims to offer 'a mathematical demonstration' that it 'discriminates against voters in the small and middle-sized states by giving the citizens of the large states an excessive amount of voting power',[25] and Brams and Davis claim that the voter in a large state 'has on balance greater

[24] Penrose, 'The Elementary Statistics of Majority Voting', and Felsenthal and Machover, 'Enlargement of the EU', Section 2.3.

[25] Banzhaf, 'One Man, 3.312 Votes'.

potential voting power … than a voter in a small state.'[26] Mann and Shapley, Owen, and Rabinowitz and Macdonald come to similar conclusions.[27] This impression has also made its way into the popular press; for example, Noah writes, 'the distortions of the Electoral College … favor big states more than they do little ones.'[28]

As discussed in the previous section, these claims – which are particularly counterintuitive given that small states are *overrepresented* in electoral votes, because of the extra two votes given to each state, no matter what size – arise directly from the square-root assumption embedded in the standard power indexes. In order to see whether this assumption is reasonable, it is necessary to look at the data.

There are four major factors affecting the probability of a decisive vote in the Electoral College. We have already discussed two of these factors: the number of voters in the state and the number of electoral votes. The third factor is the closeness of the national vote – if this is not close (as, for example, in 1984 or 1996), then the vote in any given state will be irrelevant. The fourth factor is the relative position of the state politically. For example, it is highly unlikely that voters in Utah will be decisive: if the national election is close enough that Utah's electoral votes will be relevant, then Utah will almost certainly go strongly towards the Republicans.

How can or should these factors be used to determine voting power? We consider two analyses. In Section 3.1, we look only at the number of voters and the number of electoral votes – that is, the 'structure' of the electoral system. As discussed in the previous section, the voting power will then depend on the dependence of $p_{v_j}(0.5)$ on n_j, which we can study empirically. Section 3.2 examines estimates of the probability of decisiveness from the political science literature that use a range of forecasting information, including the relative positions of the states, and then see empirically how voting power depends on the size of the state.

3.1 Closeness of the Election as a Function of the Number of Voters

As Banzhaf, Brams and Davis, and Owen make clear, the power-index results for the Electoral College are consequences of the claim that elections in large states are more likely to be close.[29] More precisely, the random voting model implies that the standard deviation of the difference in vote proportions between the two parties will be inversely proportional to the square root of the number of voters.

In fact, however, this is not the case, or at least not to the extent implied by the square-root rule. To analyse this systematically, we extend an analysis of Colantoni, Levesque

[26] Brahms and Davis, 'The 3/2's Rule in Presidential Campaigning'.

[27] I. Mann and L. S. Shapley, 'The A Priori Voting Strength of the Electoral College', RAND memo (1960) reprinted in M. Shubik, ed., *Game Theory and Related Approaches to Social Behavior* (New York: Wiley, 1964); G. Owen, 'Evaluation of a Presidential Election Game', *American Political Science Review*, 69 (1975), 947–53; G. Rabinowitz and S. E. Macdonald, 'The Power of the States in U.S. Presidential Elections', *American Political Science Review*, 80 (1986), 65–87.

[28] T. Noah, 'Faithless elector watch: gimme "equal protection" ', *Slate*, 13 December 2000, http://slate.msn.com/?id = 1006680

[29] Banzhaf, 'One Man, 3.312 Votes'; Brams and Davis, 'The 3/2's Rule in Presidential Campaigning', and S. J. Brams and M. D. Davis, 'Comment on "Campaign Resource Allocation under the Electoral College" ' by C. S. Colantoni, T. J. Levesque and P. C. Ordeshook, *American Political Science Review*, 69 (1975), 155–6; Owen, 'Evaluation of a Presidential Election Game'.

and Ordeshook and display in Figure1 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.[30]

We test the square-root hypothesis by fitting three different regression lines to $|v_j - 0.5|$ as a function of n_j. First, we use the lowess procedure to fit a nonparametric regression line.[31] Secondly, we fit a curve of the form $y = c/\sqrt{n_j}$, using least squares to find the best-fitting value of c. Thirdly, we find the best-fitting curve of the form $y = cn_j^{\alpha}$. If the best-fitting α is near -0.5, this would support the square-root rule and the claims of the voting power literature. But if the best-fitting α is near 0, then this provides evidence *rejecting* the random voting model in favour of the model in which the vote differential does *not* systematically vary with n_j As shown in Figure 1, the best-fitting α is -0.16, which is much closer to 0 than to -0.5.

We do not mean to imply by this analysis that state size is the only factor or even the most important factor determining the closeness of elections. Rather, we are giving insight into the fact that the power indexes used by Penrose, Banzhaf and others rely on an assumption that does not fit the data.

3.2 Using Election Forecasts to Estimate the Probability of a Decisive Vote

Another way to study voting power is to estimate the probability of casting a decisive vote in each state, using all available information, and then studying the dependence of this probability on state size. This was done by Gelman, King, and Boscardin using a hierarchical regression model with error terms at the national, regional, and state levels.[32] The model, based on that of Campbell,[33] was fairly accurate, with state-level errors of about 3.5 per cent. Figure 2 displays the resulting estimates of probabilities of decisive vote, for the presidential elections between 1948 and 1992. The relation between the probability of a vote being decisive and the size of the state is very weak.[34] The claims in the voting power literature that large states benefit from the Electoral College were mistaken because of the implicit or explicit assumption that elections in larger states would be much closer than those in small states.

4 DATA FROM OTHER ELECTORAL SYSTEMS

We have seen through two different analyses in Section 3 that the square-root rule does *not* empirically apply to presidential elections. But might it hold in other elections or decision-making settings, in which case claims such as Banzhaf's could be reasonable?

We examine the dependence of closeness of elections as a function of number of voters for various electoral systems in the United States and Europe. From Factor 2, the probability that a single vote is decisive is $p_{v_j}(0.5)/n_j$ or, more generally, $1/n_j$ times the

[30] C. S. Colantoni, T. J. Levesque and P. C. Ordeshook, 'Campaign Resource Allocation under the Electoral College' (with discussion), *American Political Science Review*, 69 (1975), 141–61.

[31] W. S. Cleveland, 'Robust Locally Weighted Regression and Smoothing Scatterplots', *Journal of the American Statistical Association*, 74 (1979), 829–36.

[32] Gelman, King and Boscardin, 'Estimating the Probability of Events That Have Never Occurred'.

[33] Campbell, 'Forecasting the Presidential Vote in the States'.

[34] The smallest states have slightly higher voting power, on average, which is attributable to the rule that all states, no matter how small, get at least three electoral votes.

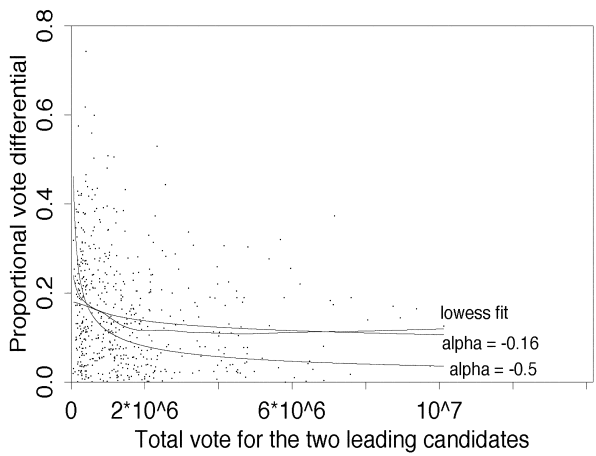

Fig. 1. The margin in state votes for president as a function of the number of voters n_j in the state: each dot represents a different state and election year 1960–2000. The margins are proportional; for example, a state vote of 400,000 for the Democratic candidate and 600,000 for the Republican would be recorded as 0.2. Lines show the lowess (non-parametric regression) fit, the best-fit line proportional to $1/\sqrt{n_j}$, and the best fit line of the form cn_j^{α}. As shown by the lowess line, the proportional vote differentials show only very weak dependence on n_j. The $1/\sqrt{n_j}$ line, implied by standard voting power measures, does not fit the data.

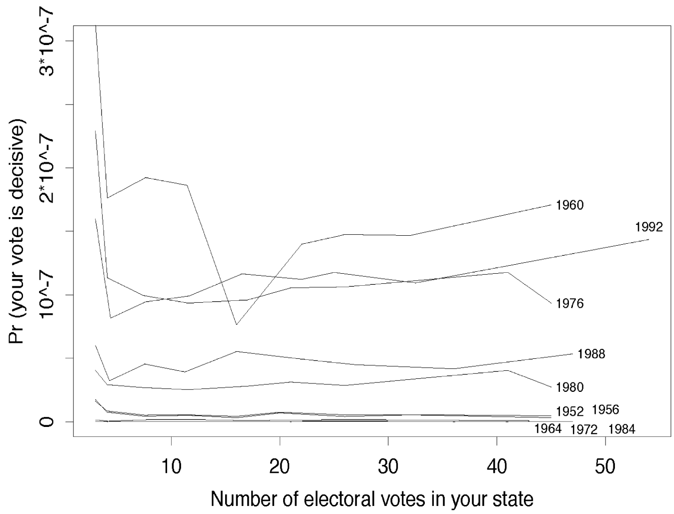

Fig. 2. The average probability of a decisive vote as a function of the number of electoral votes in the voter's state, for each US presidential election 1952–92 (excluding 1968, when a third party won in some states). The probabilities are calculated based on a forecasting model that uses information available two months before the election. This figure is adapted from Gelman, King and Boscardin, 'Estimating the Probability of Events That Never Occurred'. The most notable features of this figure are: first, that the probabilities are all very low; and secondly, that the probabilities vary little with state size, with the most notable pattern being that voters in the very smallest states are a bit more likely to be decisive.

probability that $2|v_j - 0.5|$, the proportional vote difference between the two leading parties, is within some specified distance of 0. Standard voting power indexes are based on a model that implies that the standard deviation of v_j is proportional to $1/\sqrt{n_j}$, so that as n_j increases, elections are more likely to be close.

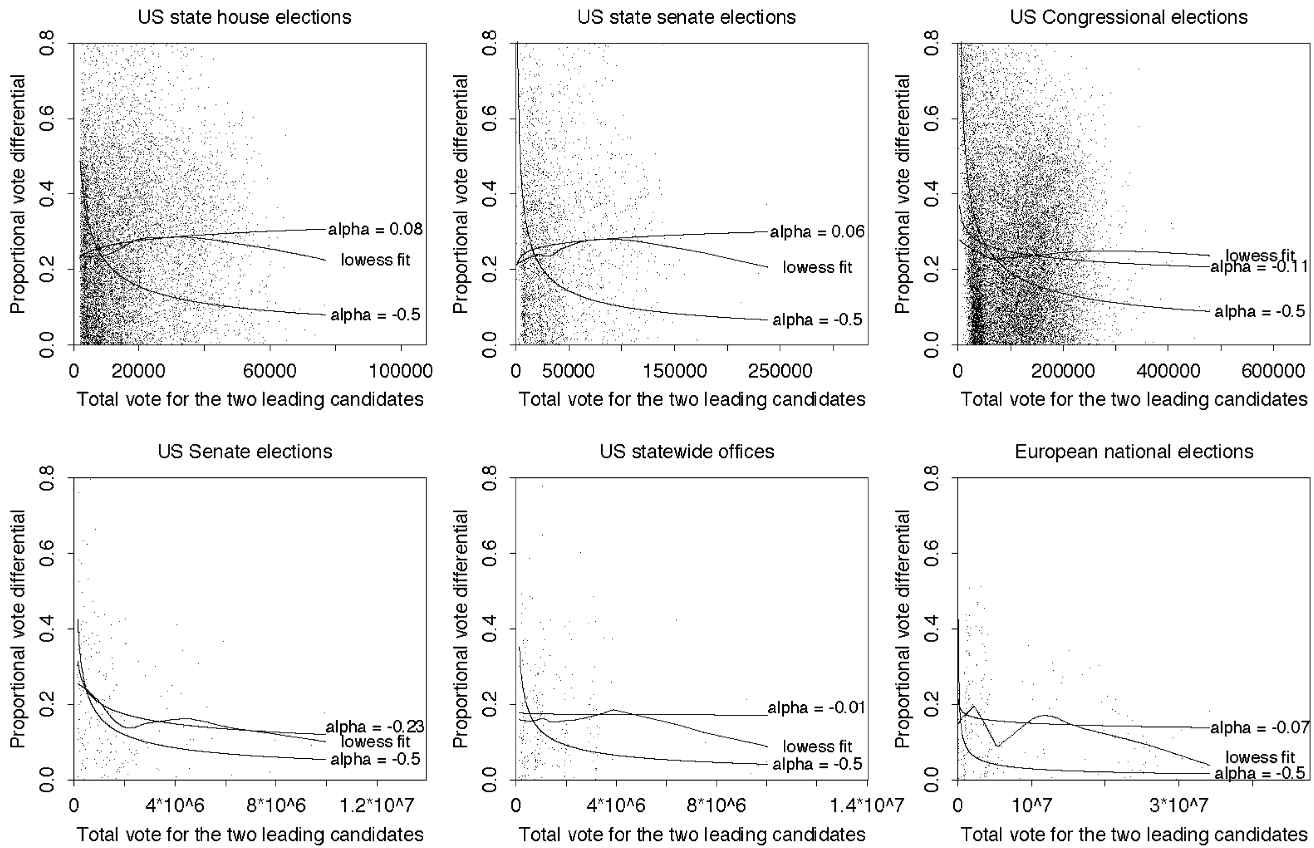

Fig. 3. Proportional vote differential v. number of major-party voters n_j, for contested elections in the following electoral systems: (a) lower houses of US state legislatures with single-member districts, 1984–90, (b) US state senates, 1984–90, (c) US House of Representatives, 1896–1992, (d) US Senate, 1988–96, (e) European national elections, 1950–98. Each plot includes lines showing the lowess (non-parametric regression) fit, the best-fit line proportional to $1/\sqrt{n_j}$, and the best fit line of the form cn_j^α. As shown by the lowess line, the proportional vote differentials show only very weak dependence on n_j. The $1/\sqrt{n_j}$ lines, implied by standard voting power measures, do not fit the data.

We replicate the analysis in Section 3.1 for these various electoral systems. The graphs in Figure 3 show, for each electoral system, the absolute value of the proportional vote differential v. the number of voters for the two leading parties in the election. We have excluded as uncontested any election in which the losing party received less than 10 per cent of the vote. As with the Electoral College (shown in Figure 1), we see in some cases a very slight decrease in the proportional vote differential as a function of the number of voters – but this decline is much less than predicted by the square-root rule. Each graph displays the lowess (non-parametric regression) line, the best-fit $1/\sqrt{n_j}$ line, and the best-fit line of the form cn_j^{α}. In each case, the lowess line is much closer to horizontal than to $1/\sqrt{n_j}$, and the best-fit parameters α are all closer to 0 than to -0.5, which would correspond to the square-root rule.

The lowess and best-fit power-law curves in Figures 1 and 3 are essentially flat. Or it could be said that they decline slightly with n, perhaps proportional to $n_j^{-0.1}$. This implies that the probability of a decisive vote is proportional to $n_j^{-0.9}$, which is far closer to $1/n$ as in the election forecasting literature than to $1/\sqrt{n}$ as in the voting power literature. Mulligan and Hunter find similar results in their study of the closeness of US House elections.[35]

5 THEORETICAL ARGUMENTS

Sections 3 and 4 give empirical evidence that the distribution of the vote share v_j is approximately independent of the number of voters, n_j, at least for reasonably large n_j. How can we understand this theoretically?

5.1 Understanding the Results Based on Local, Regional and National Swings

Politically, the reason why the square-root rule does *not* hold is that elections are affected by local, state, regional and national swings. Such swings have been found in election forecasting models and in studies of shifts in public opinion.[36]

Here, we can appeal to standard theories of public opinion, in which an individual voter's preference between two parties depends on the voter's ideological position, the parties' ideologies, and the voter's positive or negative impressions of the parties on non-ideological 'valence' issues (such as competence as a manager and personal character). To start with, the variation in ideology among the voters induces a spread in the distribution of voters' probabilities. Next, general changes in impressions about the valence issues – as caused, for example, by a recession or a scandal – shift the entire distribution of the probabilities, so that they will not necessarily be centred at 0.5 (as would otherwise be implied by a Downs-like theory of political competition).

Another way to understand the distribution of vote differentials is to compare to the square-root model, which implies that elections will be extremely close when n_j is large. This does not happen because of national swings which can shift the mean, in any particular election, away from 0.5.

[35] C. B. Mulligan and C. G. Hunter, 'The Empirical Frequency of a Pivotal Vote', National Bureau of Economic Research Working Paper 8590.

[36] See, for example, Gelman, King and Boscardin, 'Estimating the Probability of Events That Have Never Occurred'. See also B. I. Page and R. Y. Shapiro, *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences* (Chicago: University of Chicago Press, 1992).

More generally, swings in public opinion and votes can occur at many levels, from local to national and even internationally. Gelman, Katz and Tuerlinckx consider families of probability models in which voters are organized in a tree structure with correlation in their votes depending on their distance in the tree.[37] The result of this multilevel or fractal variation is that the standard deviation v_j will be expected to decline as a function of n_j, but at a slower rate than $1/\sqrt{n_j}$. Statistically, correlation structure affects the relation between sample size and standard deviation, with the $1/\sqrt{n}$ rule in general holding only for independent or weakly-correlated random variables.[38] In fact, the electoral data examined in this paper are roughly consistent with a power law with an exponent of $\alpha = -0.1$ (see Figures 1 and 3), which could be used to infer something about the fractal nature of voting patterns. For the purposes of this article, however, we are focusing on the fact that α is closer to 0 than to -0.5 for a wide variety of electoral systems, and this is consistent with modern models of public opinion, which consider large-scale vote swings, and on sociology, in which individuals are connected in complex networks.[39]

5.2 A Simulation Study Based on Presidential Votes by Congressional District

We can also understand the distribution of votes in terms of sums of random variables. For example, California in the 2000 presidential election had thirty-eight times as many voters as Vermont. If we could think of California as a sum of thirty-eight independent Vermont-sized pieces, then we would expect its v_j to have $1/\sqrt{38} = 1/6.2$ of the standard deviation of Vermont's. But this is not the case, either in terms of the closeness of the actual election result or in terms of forecasting. Uncertainties about v_j are not appreciably smaller in large than in small states.

To get a sense of what would happen if there were no state, regional or national opinion swings, we performed a simulation study using the Electoral College data displayed in Figure 1. For each election year, we took the votes for president by congressional district and subtracted a constant so that the national mean was 0.5, thus removing nationwide swings. We then permuted the 435 districts in each year and reallocated them to states based on the number of congressional districts in each 'state'; this random permutation removed all correlations associated with state and regional swings. Finally, we plotted the absolute vote differential v. turnout for these simulated elections and computed the best-fit line of the form cn_j^{α}.

Figure 4 shows the results. With these simulated data, the square-root rule fits very well. In fact, the best-fit power α was -0.494, essentially equal to a theoretical value of -0.5. This graph shows that classical voting power measures would be appropriate if elections had no state, regional or national swings. The comparison to Figure 1 shows the inappropriateness of the voting power model for actual presidential elections.

37 A. Gelman, J. N. Katz, and F. Tuerlinckx, 'The Mathematics and Statistics of Voting Power', *Statistical Science*, 17 (2002), 420–35.

38 See P. Whittle, 'On the Variation of Yield Variance with Plot Size', *Biometrika*, 43 (1956), 337–43, for a general discussion of this statistical issue.

39 See D. J. Watts, P. S. Dodds and M. E. J. Newman, 'Identity and Search in Small-world Networks', *Science*, 296 (2002), 1302–15.

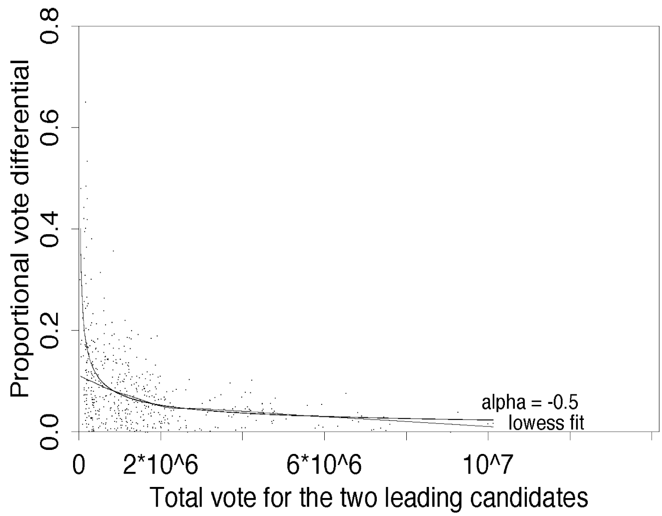

Fig. 4. Proportional vote differential v. number of voters n_j, for a random simulation of the electoral college based on presidential elections by congressional district 1960–2000. For each year, votes in all the districts were shifted so that the total vote was 50 per cent for each party. The districts were then permuted at random within each year so that, for example, 'Alabama in 1960' was constructed from nine randomly-chosen congressional districts in that election year, 'Alaska in 1960' was a different congressional district chosen at random, and so forth. Lines display the lowess fit and the best-fit line of the form cn_j^{α}. The best fit is $\alpha = -0.5$, which makes sense since the 'states' were formed by combining districts at random, eliminating state, regional and national swings.

6 DISCUSSION

6.1 Mathematical Results and Normative Claims

Proponents of voting power measures make strong claims – not just mathematical statements, but normative recommendations. For example, Penrose asserts, 'A nation of 400 million people should, therefore, have ten times as many votes (or members) on an international assembly as a nation of 4 million people', and later writers have made similarly confident pronouncements (see the titles of Banzhaf's two papers and the quotations at the beginning of Section 3).[40]

However, we have shown (in Sections 3 and 4) that the 'square-root rule' for closeness of elections, which underlies standard voting power measures, is inappropriate for data from a wide range of elections. Section 5 discussed theoretical reasons why the square-root rule does not hold.

One justification for voting power measures, even when they do not fit actual electoral data, is that they are *a priori* rules to be used in general, without reference to details of any particular elections.[41] We have no problem with the concept of *a priori* rules. After all, it seems quite reasonable for electoral votes to be assigned based on structural features such as the rules for voting and number of voters, and not on transient patterns of political preferences. For example, nobody is suggesting that Utah and Massachusetts get extra electoral votes to make up for their lost voting power due to being far from the national median.

[40] Penrose, 'The Elementary Statistics of Majority Voting'.
[41] See, for example, Felsenthal and Machover, *The Measurement of Voting Power*, p. 12.

However, it does seem reasonable to demand that an *a priori* rule be appropriate, on average, in the real world. At some point, the burden of proof has to be on the proposers of any rule to justify that it is empirically reasonable, in its general patterns if not in all details. As we have shown in Section 5, the random voting model is not consistent with accepted models of swings in public opinion. Snyder, Ting and Ansolabehere have demonstrated similar problems of voting power indexes with game theory.[42]

Voting power measures are based on an empirically falsified and theoretically unjustified model. A more realistic and reasonable model allows votes to be affected by local, regional and national swings,[43] with the result that large elections are *not* necessarily close, and that proportional weights in a legislature are approximately fair.

6.2 Constitutional Design and A Priori Voting Power

Standard voting power measures are based on considering all possible combinations of votes as equally likely and have been defended as reasonable on theoretical grounds *a priori*; for example, Leech characterizes the model as:

> a consideration of all possible outcomes that can theoretically occur, taking into account that each voter has the right to choose how to vote … Thus, voting power can be defined in terms of the rights of individual voters: we count up each outcome because each voter has the basic right, as a member of the institution, to exercise choice. There is no need, therefore, to invoke the principle of insufficient reason to justify simple random voting. *A priori* voting power can be defined on a more fundamental level in terms of voter sovereignty.[44]

Counting all possible combinations equally is not realistic for any particular electoral system, but is justified in the voting power literature as deriving from 'general philosophical principles that can be seen to apply equally to all countries and citizens'.[45] What this justification misses is that counting all voters equally is not the same as counting all *combinations* of votes as equal. As noted in Section 5.1, mathematical models exist in which all voters are treated equally and symmetrically, but with their votes being correlated (so that voters in the same area, state, city or country are more likely to vote in the same way), and these models yield results for the probability of a decisive vote that can be proportional to $1/\sqrt{n}$ (corresponding to standard voting power measures) or $1/n$ (leading to proportional weighting), or anything in between. The random voting model is simply a special case of this general family of models – and, as we have seen, it is a special case that does not fit any of the empirical data we have examined (and these cases include the US Electoral College and the European Union – two of the arenas where voting power measures have often been applied).

We believe it is acceptable and appropriate to use empirical data to construct and evaluate an *a priori* model – if, as in our examples, the empirical data cover the potential areas to which the model will be applied. To put it another way, we agree that voting power measures can be set up *a priori* ignoring the voting patterns in any particular set of elections – but we do *not* think it appropriate to ignore systematic patterns of voting that appear in all voting systems we have looked at. In designing a constitution and setting up a voting system, it would be inappropriate to assign weights based on the assumption that vote

[42] Snyder, Ting and Ansolabehere, 'Legislative Bargaining under Weighted Voting'.
[43] Or 'parallel publics' in the terminology of Page and Shapiro, *The Rational Public*.
[44] Leech, 'The Utility of the Voting Power Approach'.
[45] For example, Leech notes that France and Germany have common interests in some European Union issues (Leech, 'The Utility of the Voting Power Approach').

margins are inversely proportional to the square root of the number of voters, when empirically no such pattern appears.

6.3 Population and Turnout

Throughout this article we have made no distinction between the voters and the persons represented in an election. In reality, however, voter turnout varies dramatically between countries and between areas within a country. In addition, children and non-citizens are represented in a democracy even though they do not have the right to vote. Thus, it is standard for weights in voting systems to be set in terms of the population, rather than the number of voters, in each jurisdiction.

We agree that it is reasonable to set weights in terms of population but note that this inherently leads to variation in voting power: as the percentage voter turnout declines in a jurisdiction, the power of each remaining voter increases. We consider this acceptable because these voters represent that entire jurisdiction – but we recognize that this goes beyond the simple calculation of probability of decisiveness. The reasoning is closer to the representativeness argument of Section 2.5.

6.4 A 0.9-Power Rule?

As noted at the end of Section 4, our empirical analyses[46] are roughly consistent with a probability of decisive vote proportional to $n_j^{-0.1}$. This implies that a fair allocation of electoral votes is in proportion to the 0.9th power of the number of voters or, as discussed in Section 6.3, of the population in the jurisdiction.

We hesitate to make a recommendation of the 0.9-power rule since it lacks the Platonic appeal of the proportional and square-root rules. The proportional rule is close to the 0.9 power and is simpler to explain. However, if probabilistic assumptions were to be used in computing voting power of jurisdictions, as in Section 2.2.1, or to assign voting weights in a new or expanded legislature,[47] then it might be reasonable to use the 0.9 power to estimate the power of individual voters.

6.5 CONCLUSION

It is often claimed that, in a proportionally weighted electoral system, voters in large jurisdictions have disproportionate 'voting power'. This statement is only correct if elections in large jurisdictions are much closer than in small jurisdictions: the 'square root rule'. Empirically, this rule does not hold – in several electoral systems for which we have gathered data, the probability that an election is close is much more like a constant than proportional to $1/\sqrt{n_j}$.

From a theoretical perspective, our result – that large elections are not appreciably closer than small elections – makes sense because election results are characterized by national and regional swings. Voting power measures go wrong by assuming that the nj voters are acting independently (or, more generally, that they are divided into independent groups, where the number of groups is proportional to n_j).

We hope that this explication will allow researchers to understand the limits of current theoretical methods used in evaluating electoral systems better.

46 As are those of Mulligan and Hunter, 'The Empirical Frequency of a Pivotal Vote'.

47 As in Felsenthal and Machover, 'Enlargement of the EU' and Leech, 'Designing the Voting System for the Council of the European Union'.

APPENDIX: THE PROBABILITY OF AFFECTING THE ELECTION OUTCOME, 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 naïve to suppose that the difference between winning and losing is no more than the change in a vote margin from -1 to $+1$, as 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 grey area in vote counting, and then demonstrating that the simpler model of decisive votes is a reasonable approximation.

Consider a two-party election and label v as the proportion of the n votes received by party A. We model vote-count errors, disputes, etc., by defining $\pi(v)$ as the probability that party A wins, given a true proportion v. With perfect voting, $\pi(v) = 0$ if $v < 0.5$, 1 if $v > 0.5$, or 0.5 if $v = 0.5$. More realistically, $\pi(v)$ is a function of v that equals 0 if v is clearly less than 0.5 (e.g., $v < 0.495$), 1 if v is clearly greater than 0.5, and is between 0 and 1 if v is near 0.5.

In that case, the probability that your vote determines the outcome of the election, conditional on v (defined now as the proportion in favour of candidate A, excluding your potential vote), is $\pi(v + 1/n) - \pi(v)$. If your uncertainty about v is summarized by a probability distribution, $p(v)$, then your probability of decisiveness is:

$$\begin{aligned} \Pr\text{ (decisive vote)} &= \mathrm{E}[\pi(v+1/n) - \pi(v)] \\ &= \sum_v [\pi(v+1/n) - \pi(v)]\, p(v). \end{aligned} \tag{3}$$

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 proportions, which will lie near 0.5: thus, there is some small ε such that $\pi(v) = 0$ for all $v < 0.5 - \varepsilon$ and $\pi(v) = 1$ for all $v > 0.5 + \varepsilon$. Secondly, we assume that the probability density $p(v)$ for the election outcome has an uncertainty that is greater than ε (for example, perhaps $\varepsilon = 0.005$ and v can be anticipated to an accuracy of 2 per cent, or 0.02). Then we can approximate $p(v)$ in the range $0.5 \pm \varepsilon$ by the constant $p(0.5)$. Expression 3 can then be written as,

$$\begin{aligned} \Pr\text{ (decisive vote)} &= \int_{0.5-\varepsilon}^{0.5+\varepsilon} [\pi(v+1/n) - \pi(v)]\, p(0.5)dv \\ &= p(0.5) \int_{0.5-\varepsilon}^{0.5+\varepsilon} [\pi(v+1/n) - \pi(v)]dv \\ &= p(0.5) \left[\int_{0.5-\varepsilon+1/n}^{0.5+\varepsilon+1/n} \pi(v)dv - \int_{0.5-\varepsilon}^{0.5+\varepsilon} \pi(v)dv \right] \\ &= p(0.5) \left[\int_{0.5+\varepsilon}^{0.5+\varepsilon+1/n} \pi(v)dv - \int_{0.5-\varepsilon}^{0.5-\varepsilon+1/n} \pi(v)dv \right] \\ &= p(0.5)\, [1/n \cdot 1 - 1/n \cdot 0] \\ &= p(0.5)/n, \end{aligned}$$

which is the same probability of decisiveness as calculated assuming all votes are recorded correctly.[48]

[48] See also Good and Meyer, 'Estimating the Effcacy of a Vote', and H. Margolis, 'The Banzhaf Fallacy', *American Journal of Political Science*, 27 (1983), 321–6.