The Banzhaf Fallacy

Howard Margolis, Institute for Advanced Study, Princeton

Arguments have often been published claiming that the power of voters in presidential elections varies with the square root of the size of their state. This note points out that such results would apply only in a world in which elections were equivalent to experiments in tossing a perfectly fair coin. The technical argument is followed by an intuitive argument requiring no mathematics.

I

According to a frequently published argument (which I will label the Banzhaf argument) there is an intrinsic, large-state advantage in presidential elections, proportional to the square root of the relative size of the state. If the argument were empirically meaningful, as its proponents have always implied and usually asserted, then a voter in California would enjoy about twice the voting power of a voter in Virginia (reflecting the roughly fourfold ratio in population). This advantage, moreover, is presented as an intrinsic mathematical property of the electoral college scheme, not something deduced from demography, voting patterns, or some other appeal to empirical observations.

The earliest versions of the argument are Banzhaf (1968) and Riker and Shapley (1968). Subsequent variants of the argument can be found (among other places) in Brams and Davis, 1974, Owen, 1975, and Lake, 1979. With the obvious exception of Arrow's Impossibility Theorem, there cannot be many mathematical arguments more often cited by non-mathematical political scientists.

However, when mathematical analysis produces a counterintuitive result, it is prudent to allow for the possibility that what needs to be corrected is the mathematics. Brian Barry has made this point in a recent discussion of indices of voting power (Barry, 1980). As it happens, the particular voting power index we are considering here provides a very striking illustration of Barry's message.

A preliminary point of the discussion that follows is that the Banzhaf argument depends on a treatment of elections (for purpose of calculating a power index) "as if" they were like experiments in coin tossing, each voter tossing a coin to decide whether to vote Democratic or Republican. Of itself, the unrealism of the coin-tossing model does not make the Banzhaf result a fallacy, because all models use some more or less unrealistic simplifying assumptions (even in the natural sciences), and even extremely artificial assumptions occasionally produce adequate approximations of real situations.
But in this case, we will see, the result depends crucially on the "as if" assumption holding very exactly. A small step in the direction of greater realism is all that is needed to make the claimed power advantage of large-state voters disappear. So the problem is not directly that the Banzhaf sort of calculation treats elections as if they were experiments in coin tossing, but that unless elections were indeed very much like experiments in coin tossing, the Banzhaf argument gives a result that makes no sense empirically.

Section II gives a short version of the Banzhaf argument, providing the formulas we will need for the analysis of the fallacy in Section III. Section IV then concludes this note with an intuitive, nonmathematical version of the argument.

II

The Banzhaf proofs always define a voter's power to be effective proportional to the probability that the voter changes the outcome of the election within his district, times the probability that his district (his state in the electoral college context) change the overall outcome. The Shapley-Shubik analysis of the electoral college (which is not at issue here) finds the latter probability to be roughly proportional to the number of voters in the state. ¹ Therefore, we get

\[ P^*_{n} = nP(tie), \tag{1} \]

where \( P^*_{n} \) is the relative voting power of a voter in the state with \( n \) voters and \( P(tie) \) is the probability of a tie vote within the state.

From Equation 1, you can see that if \( P(tie) \) were the same in all states, then \( P^*_{n} \) would be just proportional to the number of voters. The voter in California would be four times more powerful than a voter in Virginia. On the other hand, other things equal, the more voters there are in a state, the less likely we are to have a tie vote. If \( P(tie) \) decreased proportionally to \( n \), then the greater swing of a voter in a big state in the event of a tie would be just offset by the reduced value of \( P(tie) \). The Californian and the Virginian would enjoy equal voting power. But what the Banzhaf argument claims to show is that while \( P(tie) \) does decrease as \( n \) increases, it does so only with \( \sqrt{n} \). Hence the power of an individual voter will be proportional to \( n / \sqrt{n} = \sqrt{n} \), and a Californian would be twice as powerful as a Virginian, as described at the outset of this note.

Mathematically, the crucial step in producing this inference is taken in treating \( P(tie) \) as proportional to the fraction of all possible combina-

¹ See the discussion of this point in Riker and Shapley (1968).
tions of voter choices that yield a tie. This gives the formula

\[ P(\text{tie}) = \frac{C_{n/2}}{2^n}, \tag{2} \]

where \( C_{n/2} \) is the binomial coefficient telling us the number of combinations that yield a tie, and the denominator, \( 2^n \), gives the number of all vote combinations, ties and nonties.

But the ratio in Equation 2 is just the result that arises—and only arises—for processes equivalent to coin tossing experiments (technically, for the case of Bernoulli trials with \( p = .5 \)). For this very simple case, a very simple formula (Margolis, 1977, Equation 4) gives \( P(\text{tie}) \), namely:

\[ P(\text{tie}) = 0.4/\sigma n = 0.8 / \sqrt{n} \tag{3} \]

where \( \sigma \) is the normalized standard deviation, which in this case is just \( \sqrt{p(1-p) / n} = 1/2 \sqrt{n} \).

Hence, the Banzhaf power ratio between voters in two states with \( n \) and \( kn \) voters is just

\[ \frac{P^*_{kn}}{P^*_n} = \frac{0.8kn}{\sqrt{kn}} / \frac{0.8n}{\sqrt{n}} = \sqrt{k}. \tag{4} \]

And if this argument is empirically reasonable, then a voter in California will indeed have twice the voting power of a voter in Virginia.

**III**

However, why should we suppose it reasonable to calculate the power ratio as if elections were experiments in tossing coins? Setting aside all other questions about the plausibility of treating elections as Bernoulli trials, what justifies giving exclusive attention to the special case where \( p = .5 \) exactly? If that were a reasonable way to model elections, then a plot of election results would show essentially all outcomes in a very tight spike peaking at exactly Republicans = Democrats = .5. In a statewide election with a million voters, the chance that the winner would get as much as 50.2 percent of the vote would be about 1 in 10,000. Obviously, a probability model that implies that is a model that does not even crudely approximate the empirical process to which we are applying the model.

Nevertheless, a power index based on the treatment of elections as experiments in coin tossing might still be reasonable on one or the other of two further arguments. First, we might imagine that although the case where \( p = .5 \) exactly is exceedingly unlikely, that is the only contingency in which a tie can arise, hence the only contingency that is relevant to the
Howard Margolis

power index. However, that argument is also obviously false. A tie can arise because \( p = .5 \) exactly, and the actual result hits the expected result right on the button; or because \( p \neq .5 \), but we happen to have just the right deviation to hit the tie outcome. (If we toss coins with slightly biased coins, that does not rule out the chance that we get a tie outcome.)

We come down, then, to this possibility: Although the special case with \( p = .5 \) exactly is not the only case in which we could get a tie, perhaps the power ratio for this special case is representative enough of the total expectation of getting a tie that the Banzhaf result remains essentially correct. However, this is not true either, though the argument is not quite so simple as for the earlier rejoinders.

Allowing for cases other than \( p = .5 \), we need the somewhat more complicated formula for \( P(\text{tie}) \) given by Margolis (1977, Equation 2):

\[
P(\text{tie}) = \Phi(x) / \sigma n = 2 \Phi(x) / \sqrt{n},
\]

where \( \Phi \) is the normal probability density function, and where for values close to \( p = .5 \) we can still use \( 1/2 \sqrt{n} \) as an adequate approximation for \( \sigma \). The Banzhaf ratio between two states with \( kn \) and \( n \) voters will then be

\[
P_{*_{kn}} / P_{*_{n}} = \sqrt{k} \hat{\phi}_{kn} / \hat{\phi}_{n},
\]

where \( \hat{\phi} \) is \( \phi(x) \) averaged across some reasonable sample of values for \( p \). With the general availability of computers, it is easy for a reader to run his own set of random trials, take the ratio, and observe that the Banzhaf ratio disappears once a modest step is taken in the direction of greater realism in evaluating \( P(\text{tie}) \). It is instructive, though, to do a hand calculation using a table of the normal probability function. Then you can see for yourself just what is happening. Let \( p \) vary over the modest range .498 to .502 by steps of .0001, assuming \( n = 1 \) million, \( k = 4 \). Using the average values for \( P(\text{tie}) \) that this yields, we find that the purported Banzhaf advantage of a factor of two for this case of a fourfold disparity in population declines to a mere 7 percent. But even that reflects only the crudity of the calculation.\(^2\)

\(^2\)Two unrelated points are worth mentioning here. First, Nicholas Miller has pointed out that if we let \( p \) vary, but assume that whatever value \( p \) takes is exactly the same in every state, then something close to the Banzhaf result will reappear. For then a tie at the electoral college stage is effectively restricted to just those cases where \( p \) is so very close to .5 that a substantial large-state advantage reappears. In other words, if we relax the assumption that \( p = .5 \) exactly, but then assume the variance among states is exactly 0, we reinstate Banzhaf. As with \( p = .5 \), a slight relaxation of this obviously unrealistic assumption leads to the disappearance of the Banzhaf effect. For example, in the calculation described here, it is sufficient to suppose that within the narrow range considered (.498 to
The following intuitive argument requires no mathematics.

We can think of a drunk with a rifle taking pot shots at a section of slatted fence ten feet wide. His shots tend to cluster toward the middle: most of them fall in the range of four to six feet from the ends. Once in a great while a shot might land exactly on the center slat of the fence, five feet from either end. But there is no bulls-eye or other feature of the setup that gives a special salience to the exact midpoint of the fence.

Now in addition to the scatter in results that would come from this process of the rifle itself were perfectly accurate (that is, in addition to the scatter due to the marksman's random pointing), we will have further scatter from the intrinsic inaccuracy in the rifle. Even if the rifle happens to be pointed at five feet exactly, there will remain a certain scatter in the impact points. In our rifle analogy it might vary with the age of the rifle, as in the Banzhaf sort of model for an election it varies with $a/\sqrt{n}$. But so long as this additional source of scatter is small compared with the rest of the randomness in the result, as would be the case either for our rifle analogy or for an empirically tenable probability model in the voting context, the second effect must be trivial. More formally, if we add consideration of a small source of scatter in a context in which the scatter is known (empirically) to be large compared with that added factor, the effect on the overall scatter (the effect on the root mean square standard error) must be small indeed.

If we think of the slats that make up the fence as representing the number of voters, then the probability of hitting the middle one, analogous to $P(\text{tie})$ in the election context, will be inversely proportional to the number of slats into which the ten feet of fence is divided. Similarly, in a plausibly realistic calculation $P(\text{tie})$ will vary inversely with $1/n$, not $1/\sqrt{n}$ as required for the Banzhof result.

.502) $p$ varies independently across states. Second, the text gives the easiest case, with $\sqrt{k} = 2$, and $\sqrt{n}$ also a convenient integer. More generally, let $p$ vary by steps on the order of $1/n$, and consider the values of $\phi(x)$ that result. Since $\sigma_x = 1/2 \sqrt{n}$, $\sigma_{kn} = 1/2 \sqrt{kn}$, the number of sample points for the small state ($n$ voters), in any interval $[0, x^*]$ of $\phi$ will be $\sqrt{k}$ times the number for the larger state ($kn$ voters), the latter being smoothly distributed among the former. The average value of $\phi$ among these sample points will then be essentially the same for the two states. But since there are $\sqrt{k}$ more points for the smaller state, we must have $\Sigma\phi_x / \Sigma\phi = 1/\sqrt{k}$. Of course the total number of sample points overall (in the range $[0, \infty]$) is the same for both states. But since values of $\phi$ decrease rapidly as $x$ increases, the contribution of points beyond, say, $x^* = 4$ is negligible. Further, for values of $n$ of interest in the electoral college context, the distribution of $p$ can be treated as flat over the range that will yield nonnegligible values of $x$. Therefore, the $1/\sqrt{k}$ ratio obtained for the sum of points within the range $[0, x^*]$ is effectively also the average for all points. In sum, then, for values of $k$ and $n$ of interest in the electoral college context, $\phi_x / \phi$ will not vary more than trivially from $1/\sqrt{k}$, and (from Equation 6) the Banzhof effect will be invisibly small.
For the special case in which the rifle happens to be pointing exactly at the midpoint (that is, for the only case considered in the Banzhaf argument), it will obviously be true that the rifle with a smaller deviation is more likely to have the shot land right at the midpoint. On the other hand, in every other case there will be a positive optimal error in the rifle itself. Over the whole range of outcomes, the Banzhaf claim makes no more political sense than a claim that the distribution of shots in the analog of the drunk shooting at the fence will change noticeably if we give the drunk a better rifle.

References

Barry, Brian. 1980. Is it better to be powerful or lucky? Political Studies, 28 (No. 2): 183–194.