Evenwel, Voting Power, and Dual Districting

Paul H. Edelman

ABSTRACT

I show that it is always possible to draw legislative districts that would be close in both total population and citizen voting-age population (or, indeed, any pair of populations that is desired). Thus, the Supreme Court need not choose between equalizing representation and equalizing voting power as it was asked to do in *Evenwel v. Abbott*. By example, I show that requiring equality of both total population and citizen voting-age population may, however, force the dilution of minority votes. Some of my analysis depends on how the Court chooses to assess the deviation in voting power. I derive the relationship between the deviation of voting power and the deviation of voting populations and show that the standard of 10 percent deviation in voting populations leads to a deviation of less than 10 percent in voting power over a broad range of models.

1. INTRODUCTION

Since the beginning of "one person, one vote" (OPOV) jurisprudence, the Supreme Court has remained vague about what, exactly, is to be equal among the political districts. Most of the cases have been resolved by creating districts so that total population (TP) is the same in each district—that is, the Court has sought to have equal amounts of representation in each district. Yet in these same opinions the Court has justified its decisions by appealing to the importance of every voter having equal weight in the voting process. But this suggests that the districts should have equal populations of voters, something quite different than equal populations.

While academic commentary on this conflict has been common (for example, Bennett 2000; Levinson 2002; Fishkin 2012; Grofman and Scarrow 1981; Gaddie, Wert, and Bullock 2012), the courts have been

PAUL H. EDELMAN is Professor of Mathematics and Professor of Law at Vanderbilt University. The author thanks Ed Cheng, Richard Hasen, William Hubbard, Nicholas Miller, Suzanna Sherry, and an anonymous referee for their helpful comments.

[*Journal of Legal Studies*, vol. 45 (January 2016)]
 0047-2530/2016/4501-0007$10.00

reluctant to address the question. The Supreme Court has heard only one case challenging the use of some measure other than TP in districting, *Burns v. Richardson* (384 U.S. 73 [1966]). In that case, the Court held that Hawaii could use resident voting-age population (RVAP) as the basis for districting its state senate, although the Court found it acceptable only because the result was not too dissimilar from what would have resulted had the state used TP instead.

There has been more discussion at the appeals court level. The best-known example is Judge Alex Kozinski's dissent in *Garza v. County of Los Angeles* (918 F.2d 763 [1990]), in which he laid out the case for requiring equal numbers of voters in the districts as the correct meaning of OPOV. Unfortunately, the Supreme Court refused to grant certiorari, and the debate languished. A few other cases raised the issue, but they failed at the appeals court level, with the Supreme Court refusing to take them up.

Finally, the Supreme Court agreed to hear *Evenwel v. Abbott* (135 S. Ct. 2349 [2015]), in which the appellants asked that the Court find that Texas, which had used TP as the basis for drawing districts, must instead use citizen voting-age population (CVAP) to equalize voting power. They framed the matter this way: "The question presented is whether the one-person, one-vote principle of the Fourteenth Amendment creates a judicially enforceable right ensuring that the districting process does not deny voters an equal vote" (Brief for Appellants, *Evenwel v. Abbott,* No. 14-940, p. i [S. Ct. filed July 31, 2015]).

By an 8–0 vote the Court held in *Evenwel* that Texas need not use CVAP and could continue to use TP for districting purposes. It left to another time the issue of whether Texas was required to use TP or whether it had the option to choose a different population measure at some time in the future.

Most of the debate on the question presented in *Evenwel* has focused on the relative merits of using TP or CVAP as the appropriate guide for districting. The commentary has assumed either tacitly or explicitly that these alternatives are exclusive—that it is not possible to achieve both equality of total population and equality of voting population. So Levinson (2002, p. 1287) says that the choice of TP is a "rule that, as a practical matter, assures *unequal* power of voters in different districts with different rates of alien population," and Gaddie, Wert, and Bullock (2012, p. 446) opine that "[i]t will likely be necessary to demonstrate that citizen apportionment is a general interest policy that does not violate the individual voting rights in order to justify the total population deviations that will inevitably result."

Because of these tacit assumptions, no commentary has focused on a secondary claim in the case—that it is possible to draw the Texas state senate districts so as to simultaneously achieve equality in both TP and CVAP in each district. This claim by the petitioners was bolstered by an affidavit from an expert witness, but no further support was provided other than an appeal to the increasing sophistication and power of districting software.[1] Justice Ruth Bader Ginsburg, writing for the majority in *Evenwel,* noted this claim when she opined, "Insofar as appellants suggest that Texas could have roughly equalized both total population and eligible-voter population, this Court has never required jurisdictions to use multiple population baselines" (slip op. at n. 15). Justice Clarence Thomas noted this in his concurrence as well (slip op. at 17 [Thomas, J., concurring]).

If it is indeed always possible to create districts that all have both equal TP and equal CVAP, then potentially the Supreme Court can have its cake and eat it too. Justice Ginsburg seemed skeptical when she expanded on her previous observation by saying, "In any event, appellants have never presented a map that manages to equalize both measures, perhaps because such a map does not exist, or because such a map would necessarily ignore other traditional principles" (slip op. at n. 15). Justice Ginsburg's skepticism notwithstanding, it remains legal for states to avail themselves of such an option, which raises the question of whether, in fact, such a districting option would always exist. No matter how sophisticated a piece of software is, if it is asked to do the impossible, it will fail. And even if such a districting is achievable in the case of the Texas state senate, as averred in *Evenwel,* how does one know that such a districting is always achievable? This question has never been addressed before.

The purpose of this paper is to show that one should, indeed, expect that districts can be drawn that are simultaneously equal in two different populations no matter how those populations are distributed. I call this the dual-districting theorem. The basis for this claim is the pancake theorem, a theorem from topology. In particular, I show that if the population densities are reasonable and if one is allowed to cut the districts by straight lines, then districts can always be cut that are equal in two different populations simultaneously. As I discuss later, these conditions

1. This reference is buried on page 46 of the Brief for Appellants, *Evenwel v. Abbott,* No. 14-940 (S. Ct. filed July 31, 2015).

are close enough to being met by real-world problems as to be no real restriction.

While the dual-districting theorem is a positive theorem, ensuring the ability to simultaneously achieve equal representation and equal voting power, there is, potentially, a very large price to pay for it. Simultaneously achieving two goals may mean that other goals must go by the wayside, and that is the situation here as well. In this case, the goal of achieving majority-minority districts may be in fundamental conflict with the other two. I give an elementary example to illustrate this problem. The conclusion may well be that the cost to minority voting rights is too high to make a dual districting justifiable.

At the state and local levels, courts have not required perfect equality of populations, and other political considerations may encourage some deviation from perfection. This raises the question of how much deviation the courts should permit. For TP the courts have already established a threshold of 10 percent in state districtings. But what about for CVAP (or some other population related to the number of voters)? It turns out that how much deviation the court should allow is intimately related to the question of how, exactly, to model voting power. Different models suggest different amounts of deviation. Section 3 of this paper details the connection between deviations and voting-power models. By a careful analysis of voting-power models, I show that there is a principled choice for the amount of population deviation that is consistent with all of the standard models of voting power.

2. DISTRICTING WITH TWO POPULATIONS

Here I give a theoretical defense of the ability to district in such a way as to achieve equality of two different populations simultaneously. That is, I give a theoretical reason to expect that it would always be possible to district a state so that every district would have the same TP and also the same CVAP. None of the existing districting literature addresses this problem. Such a claim was made in the complaint of *Evenwel* in the case of the Texas state senate districts. The argument I outline shows that Texas is not unique in this regard.

The theoretical argument is based on a topological theorem known as the pancake theorem. It has many formulations, both discrete and continuous. It is the two-dimensional version of a much more general theorem conjectured by Hugo Steinhaus and first proved by Stefan Banach, and it follows from a now-standard application of the Borsuk-Ulam theorem (Beyer and Zardecki 2004). For ease of presentation, I give a simplified continuous version. See Beyer and Zardecki (2004) for a statement and proof of a more general formulation.

Let *S* be a region in the plane (think of *S* as being a state). Two continuous population densities, μ and σ (which I think of as being people per square mile), are defined on *S*. Let μ_T (and, respectively, σ_T) be the total population of type μ, which I can write as $\mu_T = \iint_S dA$ (and similarly for σ). A directed line *l* has an orientation (an arrow) so that the left and right sides of the line are well defined. The positive side of *l* is the half-plane to the left, and the negative side of *l* is the half-plane to the right.

Theorem 1: The Pancake Theorem. There exists a directed line *l* passing through *S* so that exactly half of the μ-population is on each side of *l* and exactly half of the σ-population is on each side of *l*. Formally, if I let S^+ be the intersection of *S* with the left side of *l* and S^- be the intersection of *S* with the right side, I have

$$\frac{1}{2}\mu_T = \iint\limits_{S^+} \mu dA = \iint\limits_{S^-} \mu dA \tag{1}$$

and

$$\frac{1}{2}\sigma_T = \iint\limits_{S^+} \sigma dA = \iint\limits_{S^-} \sigma dA \tag{2}$$

Proof. This particular version of the pancake theorem follows from the same argument as that presented in Sieradski (1992, theorem 3.1.6). Q.E.D.

Suppose that μ is the density of the TP and σ is the density of the CVAP. If I am unconcerned about local governmental boundaries and such, it follows from the pancake theorem that it is always possible to divide a state into two districts so that each district would have half of the TP and half of the CVAP. Before proceeding further it should be noted that the pancake theorem does not extend to three populations. A proof of this fact appears in the Appendix.

If I want to district the state into four districts with the same property (a quarter of each population appears in each district), I can iterate the procedure, dividing S^+ and S^- each by a line guaranteed by the theorem. Indeed, this method works for any number of districts that is a power of 2.

Of course I may want to divide the state into some other number of districts. For instance, could I get five districts each with $\frac{1}{5}$ of the TP and the CVAP? The answer is yes, although the argument is somewhat more involved. Fix a number n and by repeated use of the pancake theorem produce 2^n districts each with $1/2^n$ of the two populations. Glue these small districts together into five districts each containing either $\lfloor 2^n/5 \rfloor$ or $\lceil 2^n/5 \rceil$ of the small districts (where $\lfloor x \rfloor$ means round x down and $\lceil x \rceil$ means round x up). The five resulting districts would differ by an amount of at most $1/2^n$ in each population, because of the rounding. But if n is sufficiently large, then in the limit I can get a set of five districts where each would contain the same amount of TP and CVAP. Of course there is nothing special about the number five here, so I arrive at the following theorem:

Theorem 2: Dual Districting Theorem. With S, σ, and μ as described above, and for any positive integer k, it is possible to divide S into k pieces so that each piece would contain $(1/k)$th of both σ-population and μ-population.

The dual-districting theorem then gives theoretical support to the intuition that one can district a state to achieve equal TP and equal CVAP simultaneously, as was asserted in *Evenwel* for the case of the Texas state senate. This theory ignores two important considerations. The first is that the populations in question are really discrete and not continuous. The second is that districting is not (and under some state laws may not be) accomplished by purely straight lines. These objections are much less important than they may seem, and I deal with them in more detail in the Appendix. For now let me say that since districting is done at the census block level, and these blocks are quite small, the effective population distribution can reasonably be approximated by a continuous function. And while districting is not accomplished with straight lines, the lines do tend to follow geographic borders that are close enough to make little difference.

It is easy to convince oneself that one can construct districts using the theorem in such a way that the districts are connected and compact. Since any way of gluing the small districts together would result in equal populations, one can choose a way that achieves reasonably shaped districts as well.

But reasonably shaped and equipopulous districts are not all that are desired in a districting plan. There may be other considerations as well.

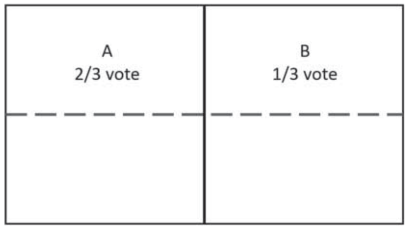

Figure 1. Impossibility of majority-minority districts under dual districting

Some may be mandated by state law, such as respect for political subdivisions and respect for communities of interest. It is plausible that these requirements are compatible with dual districting, given the flexibility one has in drawing some of the boundaries.

Other goals might include more political aims such as incumbent protection, partisan advantage, and majority-minority districts. Incumbent protection, the drawing of districts so that incumbents would not be forced to run against each other, should be manageable since one need only pay attention to the domiciles of the candidates and avoid too much alteration in the boundaries of the current districts. Partisan advantage and majority-minority districts, on the other hand, may be difficult and in some cases impossible. In particular, majority-minority districts may become impossible, as the following example (see Figure 1) illustrates.

Consider a square city with a uniform total population that is to be divided into two districts. Suppose that in the left half (A) of the city $\frac{2}{3}$ of the population are citizen voters and in the right half (B) only $\frac{1}{3}$ of the population are citizen voters (where the citizen voters are uniformly distributed throughout). In addition, suppose that the left half of the city is all Anglo and the right half is all Hispanic. As guaranteed by the dual-districting theorem, there is a way to draw two districts so that each has exactly half of both the total population and the voters. It is easy to see that essentially the only way to do that is by cutting the city into a top half and a bottom half, as illustrated in Figure 1.[2] In this (essentially) unique dual districting, each district has twice the number of Anglo voters as Hispanic voters. It follows, then, that the only way to achieve equality in both TP and CVAP results in a districting that divides the Hispanic vote, making it impossible to have a majority-minority district.

2. One can show that each district must contain exactly half of A and half of B. Any way to achieve that would be fine, but cutting the city in half horizontally is the most natural.

As this example makes clear, the requirement that a districting plan achieve equality in both TP and CVAP is such a strong condition that there may be little opportunity to achieve other goals. This could be a positive, if the goals forgone are partisan or incumbent gerrymanders. But it also could be a negative, if majority-minority districts are desired. The trade-off obviously depends on the polities that are being districted, and so the desirability of a dual districting may be very region specific.

It is also worth noting that, at the state and local level, districts are allowed to deviate from perfect equality of population. This flexibility might allow for more accommodation of a third goal beyond equality of TP and CVAP. So, depending on the actual population distributions, there may be a way to accommodate majority-minority districts while achieving near equality in both TP and CVAP.

3. THE CORRECT WAY TO MEASURE VOTING POWER

As I have shown above, the possibility of simultaneously districting for equality of TP and CVAP is in reach. The ability to realize that goal may well depend on how much leeway the districts have to fall short of actual equality. The courts have already established thresholds of disparity when dealing with TP. How should those thresholds change when districting according to CVAP? The first aspect to this inquiry is to decide what measure captures the weight of a citizen's vote, and the second is to choose a measure for the deviation of that weight between voters. Neither of these choices is self-evident, and I deal with each in turn.

Finally, I can use the deviation in voting populations as an upper bound on the deviation of voting power no matter which measure of voting power is chosen. As a result, courts can choose an acceptable threshold for the deviation of voting power without having to endorse a particular model of voting power.

3.1. The Weight of a Vote

In the cases challenging the equal-representation model of OPOV, it has been asserted as obvious that the appropriate way to weight a vote is by the size of the CVAP or by the RVAP. For this paper I set aside the question of exactly which population of eligible voters is appropriate and use CVAP. More difficult than the question of which voting-age population is correct is what exactly to do with the number that is chosen.

Quantifying voting power is a difficult endeavor. While there are many ways to do it, I focus on the formulation that is most widely accepted in the political science literature: the power of an individual's vote is the likelihood that the individual's vote is decisive; that is, if the voter were to change her vote, the outcome would change as well.

A small example may help to illustrate the problem. Consider a district with only one voter. That voter clearly has all of the power under any reasonable measure of power. Now consider another district with three voters, a, b, and c. Suppose they are casting a ballot to decide between candidates, X and Y. The eight different possible outcomes of this vote

$$\emptyset|abc,\ a|bc,\ b|ac,\ c|ab,\ ab|c,\ ac|b,\ bc|a,\ abc|\emptyset,$$

where by a|b I mean that the voters in the set a vote for X and those in b vote for Y. Look at voter a. In half of these eight outcomes, if a were to change her vote, the outcome would change. For example, if the vote was ab|c, then X would win, but if a were to change her vote the result would be b|ac, resulting in Y's victory. If all of a's votes were equally likely, would it be unreasonable to assign a power of $\frac{1}{2}$ to a since, in half of her possible votes, her vote is determinative? And if one accepts this logic, then the power of a voter in a district with three voters is not one-third of that of a voter constituting a single district, but actually one-half.

This analysis, which goes back to Penrose (1946) and was rediscovered by Banzhaf (1965), depends on the probabilistic assumption that all of a's votes were equally likely and leads to a model where the voting power of a voter in a voting population of size q is proportional to $1/q^{1/2}$. The idea behind this model is that it gives a baseline measure of the power of a voter in the absence of any further information about coalitions, abstentions, ideological biases, and so forth. The model was presented to the Supreme Court on two occasions (*Whitcomb v. Chavis*, 403 U.S. 124 [1971], and *Board of Estimate of City of New York v. Morris*, 489 U.S. 688 [1989]), where it fared poorly. It was, more or less, accepted in one state high-court case, however (*Ianucci v. Board of Supervisors of Washington County, NY*, 282 N.Y.S.2d 502 [1967]). For more details about the judicial reaction to Banzhaf's measure, see Felsenthal and Machover (1998, ch. 4).

A different theoretical approach is due to Good and Mayer (1975), who approach the question from the mind of the voter. In modeling why a voter would cast a ballot at all, they assume that the voter would act like a rational Bayesian. In order to decide whether or not to vote, the voter needs to know the likelihood that he would be casting a decisive vote, for otherwise there is no reason to bear the cost of voting at all. It then turns out that under most any Bayesian prior for the voting probabilities of the other voters, the probability of being decisive is concentrated, and this leads to a voting-power model proportional to $1/q$.

An empirical approach to voting power is taken by Gelman, Katz, and Bafumi (2004). They reject Banzhaf's a priori assumption that voters are equally likely to cast their ballot for one side as for another. Instead, they statistically test a number of different models where the form of voting power is proportional to $1/q^{\alpha}$. They get mixed results depending on the type of election involved, sometimes seeing that there is little dependence at all on q—that is, $\alpha = 0$; the voting power is the same no matter the size of the polity (Gelman, Katz, and Bafumi 2004, p. 666)—and in other elections seeing a better fit to $\alpha = .9$ (p. 669).

It is important to observe that judicial rejection of the voting-power model was in the context of replacing the goal of equal representation with that of equal voting power. As noted earlier, the courts had been unwilling to consider the goals as distinct. But because the Court agreed to hear *Evenwel,* it is acknowledging that voting power is distinct from equal representation. How the Court would reconcile these two views is yet to be seen, but to do so it would have to confront these models of voting power.

What one sees from these models is that a voter's power has the form $1/q^{\alpha}$ for some value $0 \leq \alpha \leq 1$, where q is the size of the population of voters. But why does the model matter? If it gives results that are proportional to a function of the relevant population, and one wants power to be equal, then all one needs do is make the populations in the districts equal. If one could, and was required to, design the districts so that they had exactly the same relevant population, then the choice of model would be irrelevant. But the courts allow the states to have some deviation of the populations when they district, and one would expect that districting according to CVAP is no different. It is when one measures the deviation from a perfect districting that the choice of voting-power model makes a difference, as I now show.

3.2. Measuring Deviation from One Person, One Vote

To understand how the choice of voting-power model affects the deviation measure, one must understand how the courts measure deviation in

Table 1. Total Population Deviation

District	CVAP	% Deviation
1	707,651	–27.8
2	922,180	–5.9
3	1,098,663	12.1
4	1,081,089	10.4
5	1,088,388	11.1
Average	979,594.2	
Total		39.9

Note. CVAP = citizen voting-age population.

the voting context. While there are any number of different ways to do this, the courts have settled on the measure of total deviation as the appropriate one for the purposes of OPOV. Here I discuss this statistic and show how it interacts with voting-power measures.

In the case of traditional OPOV inquiries, where one is measuring representation and wishes to equalize the number of people in each district, total deviation works like this: For each district one computes the number of people per representative as well as the size of the ideal district—that is, the total population divided by the number of districts. Then for each district one computes the percentage deviation from the ideal district size and takes the difference between the most overrepresented and the most underrepresented district. That difference is the total deviation.

As an example, consider the districting in question in *Garza,* as described in Table 1. The population in district 1 is 27.8 percent smaller than the average district, and the population in district 3 is 12.1 percent larger than the average. Thus the total deviation, the difference between the smallest and largest districts, is 39.9 percent. The Supreme Court has found that a total deviation of less than 10 percent would be considered de minimis for legislatively enacted apportionments of state legislatures, and it allows somewhat larger deviations for local governments (*Connor v. Finch*, 431 U.S. 407 [1977]). There is little doubt that a total deviation of 39.9 percent would be a violation of OPOV.

If the courts are concerned with voting power, then it would be natural for them to compute the total deviation of voting power between the districts. What if one applies the total-deviation measure to the Banzhaf voting power (where $\alpha = \frac{1}{2}$) instead of the population numbers themselves? With the population data from Table 1, the results in Table 2 are obtained.

Table 2. Voting Power Deviation

District	CVAP	Voting Power $\alpha = \frac{1}{2}$	% Deviation
1	707,651	.00134	17.7
2	922,180	.00118	3.1
3	1,098,663	.00108	−5.6
4	1,081,089	.00109	−4.8
5	1,088,388	.00108	−5.1
Average	979,594.2	.00114	
Total			23.3

Note. CVAP = citizen voting-age population.

I show here that when voting power is used, instead of the population numbers themselves, the total deviation drops by almost half. This is not special to this example. As I show in the Appendix, under the assumption that the district sizes do not vary too far from average, the total deviation of the Banzhaf voting power is roughly half of the total deviation of the populations.

To be consistent, the Court should be interested in limiting the deviation between the voting power of the voters in the different districts. If the Court were to set a limit on that deviation, then if one knew which voting model was correct, one could solve backward to figure out what deviation in the populations would be acceptable. This requires a choice of a voting model, however, and the Court may well be reluctant to make such a choice. How much does the choice of the voting model affect the outcome, though? Perhaps the difference in the conversion from deviation of voting power to deviation in populations is not really affected by which model is chosen. Maybe the Court can get by without overtly selecting such a model.

To be precise, I am interested in how the relationship between the total deviation of the voting power and the total deviation of the populations changes as a function of α. To analyze this, I proceed as follows. Let q_a be the average district size and q_B and q_L be the two districts with the biggest and smallest populations, respectively. Define

$$\varepsilon = \frac{q_B - q_a}{q_a} \quad \text{and} \quad \delta = \frac{q_a - q_L}{q_a}.$$

If I let TVP(α, ε, δ) be the total deviation of the voting power using the $1/q^{\alpha}$ voting model, and TD(ε, δ) be the total deviation of just the populations, then I can compute

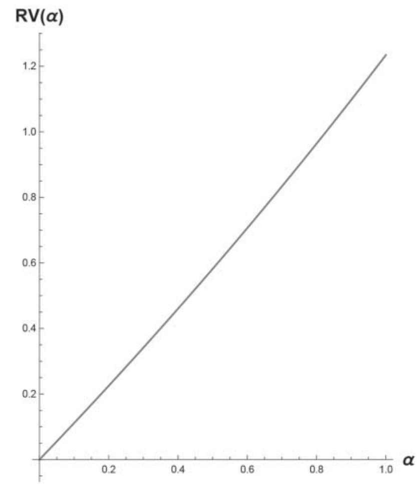

Figure 2. Ratio of total deviation of voting power to total deviation of populations

$$\mathrm{RV}(\alpha,\ \varepsilon,\ \delta) = \frac{\mathrm{TVP}(\alpha,\ \varepsilon,\ \delta)}{\mathrm{TD}(\varepsilon,\ \delta)},$$

which is the ratio of the total deviation of the voting power to the total deviation of the underlying populations as a function of the voting model.

Theorem 3: Total-Deviation Theorem. For ε and δ suitably small and $0 \leq \alpha \leq 1$, I have that

$$R(\alpha,\ \varepsilon,\ \delta) \approx \left(1 - \frac{\varepsilon - \delta}{2}\right)\alpha.$$

From the example in Table 2, where $\varepsilon \approx .11$ and $\delta \approx .27$, I get the graph in Figure 2. Note that the total-deviation theorem implies that

$$\mathrm{RV}(\alpha,\ .11,\ .27) \approx \left(1 - \frac{.11 - .27}{2}\right)\alpha \approx 1.08\alpha.$$

It follows from the total-deviation theorem that if the largest deviations from the average are roughly symmetric—that is, if $\varepsilon \approx \delta$—then RV $\approx \alpha$. That is to say that if I use the Banzhaf voting-power model ($\alpha = \frac{1}{2}$), then the total deviation of voting power would be roughly half of the total deviation of population, whereas if I use the empirical model of Gelman, Katz, and Bafumi (2004) ($\alpha = .9$), the total deviation of voting power is nine-tenths that of the total deviation of the populations. A full proof of the total-deviation theorem appears in the Appendix.

The implication of this analysis is that if the Court were to decide that a total deviation of 10 percent in the CVAP was acceptable, this would allow for a 5 percent total deviation in the Banzhaf voting power and a total deviation of 9 percent in the Gelman voting power. On the other hand, if the Court were persuaded by the Banzhaf model, it might choose a total deviation of 20 percent in the population to get a 10 percent total deviation in the Banzhaf power and an 18 percent deviation in the Gelman power. In this way, the total-deviation theorem can give some guidance to the Court in deciding what threshold of deviation might be appropriate. In particular, a total deviation of 10 percent in CVAP implies a total deviation in voting power of less than 10 percent over the whole range of models that I consider.

4. CONCLUSION

In *Evenwel,* the Court found that states need not equalize voting power in their legislative districts. It has left unresolved whether equalizing representation is the sole measure to be used or whether equalizing voting power may be used in its stead. This paper demonstrates that it is very likely that one can achieve a districting that achieves both equal representation and equal voting power if the Court would also allow enough deviation from perfect equality. Given previous decisions, the Court may not follow that route for congressional districts. On the other hand, if states are given more flexibility in how they district, they may wish to draw districts that are equal in both population and voting power. The dual-districting theorem suggests that they can.

The choice of an acceptable level of deviation is intimately tied to which theory of voting power is adopted, either explicitly or implicitly. The more power the Court, or legislature, feels that an individual voter has—that is, the smaller the value of α—the larger the acceptable disparity in population, in an essentially linear fashion. On the other hand, it is possible to choose a level of deviation that is broadly consistent with any of the standard theories of power.

Of course, by adding an extra requirement to the districting plan, one necessarily loses some flexibility. In one sense this could be a good thing—it would limit the opportunity to gerrymander for either partisan or incumbent advantage. On the other hand, it would also impair the ability to draw districts that are majority minority. Indeed, as shown by the earlier example, dual districting may well be antithetical to achieving majority-minority districts. If the Court mandates, or a legislature decides to include, considerations of equal voting power in its districting, it may well have to confront how the Voting Rights Act interacts with this goal.

APPENDIX: PROOFS AND CAVEATS

A1. Pancake Theorems

The version of the pancake theorem presented in Section 2 is not the most general one available. Two ways exist to extend it that may apply to the districting context, but both extensions are at the cost of more technical detail; I discuss these possibilities here. I also explain why the pancake theorem is sharp—that is, the theorem fails if there are three different populations to district.

In my presentation of the pancake theorem I assumed that the population densities were continuous. As noted in Section 2, while this is not a perfect description, it is a more than reasonable approximation. In fact, some generalizations of the pancake theorem assume weaker conditions on the densities. For example, in Živaljevič (2004, theorem 14.2.1) the only requirement is that the densities are σ-additive Borel measures. As explaining this approach is likely to tax the nonspecialist reader, I chose not to develop it.

The other direction I might go is to a fully discrete version. Imagine that there are scattered some number of blue points and perhaps a different number of red points in the plane. Ideally one would like to conclude that one can draw a line so that half of each color appears on each side of the line. Alas, that cannot always be done, as can easily be seen: if there are an odd number of red points, then a line cannot be drawn that misses all of the points but divides them into equal numbers on either side, since that would imply that the number of points is even. So one knows that sometimes the points would have to appear on the line. However, if some of the points have to appear on the line, and if a large number of points are collinear, one can easily find oneself with a large number of points on the dividing line—maybe even all of them.

I can avoid this pathology by assuming that the points are all in general position—that is, no three of the points appear on a line. But such an assumption is as bad as, and arguably worse than, the assumption about the continuity of the density functions. So, all things considered, it is better to follow the development as given in Section 2, even if all of the above considerations represent theoretical exceptions that are irrelevant in the real world.

The complaint that the pancake theorem relies on continuous densities whereas districting is relative to discrete individuals is not really accurate. Districting in practice is done at the census block level, where one works with, essentially, densities—the number of people (or voters) in the district. So one can think of the densities μ and σ as being built by gluing together the densities of each census block. These blocks are small enough that the assumption that they are continuous is a reasonable one. There are also discrete versions of the pancake theorem that give similar results.

But what about the problem that in districting in real life one is not allowed to just draw straight lines? The lines are required (either by law or at least by custom) to respect local government boundaries. Would those requirements be enough to undercut the ability to simultaneously district by TP and CVAP?

At first blush one might think that one has more flexibility if the district lines need not be straight, but that is a bit of an illusion. The lines of the district are generally not allowed to divide census blocks, and so the correct mathematical parallel would be to use point masses for the populations and allow only a limited set of topological lines that avoid going through any of the points. Making precise what topological lines would be acceptable would add another level of difficulty to the analysis that I do not believe would be helpful. And, again, this is a purely mathematical concern not likely to be relevant to actual problems.

Having said that, one does have to worry that the actual lines may not divide blocks, and so the populations necessarily have a lumpy character. But the lumpiness is so small compared to the total population that one can mostly ignore it in real life. Since *Karcher v. Daggett* (462 U.S. 725 [1983]), congressional districts in a state have had to be essentially identical in TP, and so for dual districting all of the adjustment would have to fall in CVAP. Only if the courts required the same level of exactitude in both TP and CVAP would there be any possible concern.

For state and local districts the law is different. States are allowed a total deviation of up to 10 percent in TP among the districts in their districting plans. In the case of Texas state senate districts, that amounts to roughly 80,000 people. With that amount of flexibility in the system, if one starts from a perfect districting à la the dual-districting theorem and then pushes the lines a small amount to conform with the required boundaries, one can easily stay within the 10 percent window that states are allowed.

Finally, I explain why the pancake theorem does not work with three different populations. Imagine that the three populations were each uniform in separate discs of radius 1 and the three centers of the discs do not lie on a line. The only way to divide a uniform disc into equal parts is to cut through the center of the disc. But that means that for each pair of populations there is a unique line that can divide both simultaneously—namely, the unique line defined as going through the two centers of those discs. By construction, such a line cannot go through the third center, and hence the third population is not divided in two equal parts. Hence, one cannot guarantee a division that would simultaneously cut all three sets.

A2. Proof of the Total-Deviation Theorem

Here I establish the relationship between the ratio of total deviation of voting power to total deviation of population and the exponent α that characterizes the voting power. I assume that voting power is proportional to $1/q^{\alpha}$, where q is the relevant population and $0 \leq \alpha \leq 1$. Assume that the polity is divided into n single-representative districts with populations of q_i, where $1 \leq i \leq n$, respectively. Let q_a be the average district size—that is, $q_a = (1/n)\sum_i q_i$. The total deviation for this districting, TD, is given by

$$\begin{aligned} \text{TD} &= \max_{i,j}\left(\frac{q_i - q_a}{q_a}\right) - \left(\frac{q_j - q_a}{q_a}\right) \\ &= \frac{1}{q_a}\max_{i,j}(q_i - q_j) \\ &= \frac{1}{q_a}(\max_i q_i) - (\min_j q_j). \end{aligned} \tag{A1}$$

As discussed in the text, I model the voting power of an individual voter in a binary election among q voters as proportional to $V(q) = 1/q^{\alpha}$. It follows that the total deviation of voting power, TDVP(α), is given by

$$\begin{aligned} \text{TDVP}(\alpha) &= \max_{i,j}\left[\frac{V(q_i) - V(q_a)}{V(q_a)}\right] - \left[\frac{V(q_j) - V(q_a)}{V(q_a)}\right] \\ &= \frac{1}{V(q_a)}\max_{i,j}[V(q_i) - V(q_j)] \\ &= \frac{1}{V(q_a)}[\max_i V(q_i)] - [\min_j V(q_j)]. \end{aligned} \tag{A2}$$

If I let $q_B = \max_i q_i$ and $q_L = \min_j q_j$, it follows that $V(q_B) = \min_i V(q_i)$ and $V(q_L) = \max_i V(q_i)$, and thus

$$\text{TD} = \frac{1}{q_a}(q_B - q_L) \tag{A3}$$

and

$$\text{TDVP}(\alpha) = \frac{1}{V(q_a)}[V(q_L) - V(q_B)]. \tag{A4}$$

Substituting the formula for V and simplifying, I obtain

$$\begin{aligned} \text{TDVP}(\alpha) &= q_a^{\alpha}(q_L^{-\alpha} - q_B^{-\alpha}) \\ &= q_a^{\alpha}\left(\frac{q_B^{\alpha} - q_L^{\alpha}}{q_B^{\alpha} q_L^{\alpha}}\right). \end{aligned} \tag{A5}$$

My concern is with how a change in the total deviation of the districting affects the total deviation of the voting power. To analyze this, I consider the ratio $R(\alpha)$ given by

$$R(\alpha) = \frac{\text{TDVP}(\alpha)}{\text{TD}} = q_{\text{a}}^{\alpha}\left(\frac{q_{\text{B}}^{\alpha} - q_{\text{L}}^{\alpha}}{q_{\text{B}}^{\alpha} q_{\text{L}}^{\alpha}}\right) \Bigg/ \frac{q_{\text{B}} - q_{\text{L}}}{q_{\text{a}}} = q_{\text{a}}^{1+\alpha}\left[\frac{q_{\text{B}}^{\alpha} - q_{\text{L}}^{\alpha}}{q_{\text{B}}^{\alpha} q_{\text{L}}^{\alpha}(q_{\text{B}} - q_{\text{L}})}\right].$$

To analyze this function I express q_{B} and q_{L} in terms of q_{a}. Suppose that $q_{\text{B}} = q_{\text{a}} \times (1 + \varepsilon)$ and $q_{\text{L}} = q_{\text{a}} \times (1 - \delta)$, which is to say that ε is the percentage by which q_{B} exceeds the average and δ is the percentage by which q_{L} is less than the average. Substituting these values into $R(\alpha)$, I show that the q_{a} terms cancel out, and the result is

$$R(\alpha,\ \varepsilon,\ \delta) = \frac{(1+\varepsilon)^{\alpha} - (1-\delta)^{\alpha}}{(1+\varepsilon)^{\alpha}(1-\delta)^{\alpha}(\varepsilon+\delta)}.$$

If I expand R as a Maclaurin series in α, the first two terms are

$$R(\alpha,\ \varepsilon,\ \delta) \approx 0 + \frac{\ln(1+\varepsilon) - \ln(1-\delta)}{(\varepsilon+\delta)}\alpha + O(\alpha^2).$$

Assuming that both ε and δ are small enough, I can further expand this second term as a power series in ε and δ to get

$$\begin{aligned} R(\alpha,\ \varepsilon,\ \delta) &\approx \frac{(\varepsilon - \varepsilon^2/2 + \cdots) - (-\delta - \delta^2/2 - \cdots)}{\varepsilon + \delta}\alpha \\ &\approx \left(1 - \frac{\varepsilon - \delta}{2}\right)\alpha. \end{aligned} \tag{A6}$$

REFERENCES

Banzhaf, John F., III. 1965. Weighted Voting Doesn't Work: A Mathematical Analysis. *Rutgers Law Review* 19:317–43.

Bennett, Robert W. 2000. Should Parents Be Given Extra Votes on Account of Their Children? Toward a Conversational Understanding of American Democracy. *Northwestern University Law Review* 94:503–65.

Beyer, W. A., and Andrew Zardecki. 2004. The Early History of the Ham Sandwich Theorem. *American Mathematical Monthly* 111:58–61.

Felsenthal, Dan S., and Moshé Machover. 1998. *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes*. Cheltenham: Edward Elgar.

Fishkin, Joseph. 2012. Weightless Votes. *Yale Law Journal* 121:1888–1910.

Gaddie, Ronald Keith, Justin J. Wert, and Charles S. Bullock III. 2012. Seats, Votes, Citizens, and the One Person, One Vote Problem. *Stanford Law and Policy Review* 23:431–54.

Gelman, Andrew, Jonathan N. Katz, and Joseph Bafumi. 2004. Standard Voting Power Indexes Do Not Work: An Empirical Analysis. *British Journal of Political Science* 34:657–74.

Good, I. J., and Lawrence S. Mayer. 1975. Estimating the Efficacy of a Vote. *Behavioral Science* 20:25–33.

Goodman, Jacob E., and Joseph O'Rourke, eds. 2004. *Handbook of Discrete and Computational Geometry*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC.

Grofman, Bernard, and Howard A. Scarrow. 1981. The Riddle of Apportionment: Equality of What? *National Civic Review* 70:242–54.

Levinson, Sanford. 2002. One Person, One Vote: A Mantra in Need of Meaning. *North Carolina Law Review* 80:1269–98.

Penrose, L. S. 1946. The Elementary Statistics of Majority Voting. *Journal of the Royal Statistical Society* 109:53–57.

Sieradski, Allan J. 1992. *An Introduction to Topology and Homotopy*. Boston: PWS-Kent.

Stone, A. H., and J. W. Tukey. 1942. Generalized "Sandwich" Theorems. *Duke Mathematical Journal* 9:356–59.

Živaljević, R. T. 2004. Topological Methods. Pp. 305–29 in *Handbook of Discrete and Computational Geometry,* edited by Jacob E. Goodman and Joseph O'Rourke. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC.