Chapter 6. Do Traditional Districting Principles Matter?

> So geographers, in Afric maps,
>
> With savage pictures fill their gaps,
>
> And o'er uninhabitable downs
>
> Place elephants for want of towns.
>
> from *Poetry, a Rhapsody* by Jonathan Swift

6.1. Why might *traditional districting principles matter?*

Since its beginning, drawing maps has been an integral part of political life in America. Traditionally, it was politicians (sometimes with scholarly assistance) who shaped the lines that would, in turn, shape their own elections. The courts seldom intervened. Even when called upon, they would do no more than invalidate a single election (Cortner 1970). The current Court, in a spate of untraditional activism, has invalidated congressional redistricting plans, ironically claiming t.d.p.'s as its compass. In a series of cases, starting with *Shaw* v. *Reno* (1993, henceforth "Shaw I") and continuing in *Abrams* v. *Miller* (1997), it has all but required that states follow t.d.p.'s such as contiguity, compactness, respect for political boundaries, and population equality — if they create majority-minority districts. In a reversal of roles, scholars and politicians trail the court asking: Do traditional districting principles matter?

Is there justification for concern with gerrymanders? Will traditional districting principles "solve" the "problem" of racial and partisan gerrymandering? Early research failed to find much of a connection between intentional gerrymandering and electoral

outcomes. Recently, racial gerrymandering has been blamed for many of the heavy Democratic losses in the 1992 and 1994 congressional elections (Hill 1995; Swain 1994). Furthermore, new studies have shown that redistricting has significant influence on representation. (Cameron, Epstein and O'Halloran 1996; Gelman and King 1994; Kousser 1995; Cain, 1985 #96 but see O'Rourke 1980; Rush 1993 for an opposing view.; Squire 1995) We know less about how to effectively control gerrymanders. While some political scientists have studied the effects of redistricting, other political and legal scholars have proposed to limit its abuses by imposing rules on redistricting (Polsby and Popper 1991; Stern 1974). These two bodies of work, one studying rules for redistricting, the other redistricting's consequences, have remained separate, and the efficacy of the latter's proposals remains largely untested.

In this study, I bridge the gap between the (primarily) empirical work on how redistricting affects elections and the (primarily) theoretical work on redistricting principles. Many redistricting principles (See Lijphart 1989, for a survey of both traditional and non traditional principles.) have been proposed on the basis of theory, under the supposition that they would control gerrymandering, reduce bias, and improve elections; few, however, have been subjected to rigorous empirical analysis. For example, even the most recent and detailed studies of district compactness (Niemi et al. 1990; Niemi and Wilkerson 1991) have not attempted to analyze the connections between this principle and electoral outcomes. Many empirical questions about redistricting principles remain unanswered: How much do redistricting principles affect gerrymandering? Do compact districts change the results of elections by increasing

partisan fairness, or by making them more competitive? Do "ugly" districts decrease turnout, or change the attitudes of voters in those districts?

In this chapter I answer these questions by building on previous statistical maximum-likelihood analyses of redistricting and by applying these statistical models to a novel set of data. I measure the compactness, contiguity, "respect for political boundaries," and malapportionment of all United States congressional districts from 1789 through 1993[234] — creating a set of data that has not previously, to my knowledge, appeared in print.

6.1.1. How Do T.D.P.'S Matter In The Court?

In the courts, many types of districts have been under attack, but congressional districts have undergone particularly close recent scrutiny by the Supreme Court. In all the recent cases in which the Court has particularly emphasized compactness and other t.d.p.'s, the Court has been looking at congressional districts.

Compactness, contiguity and malapportionment have long been factors in redistricting jurisprudence. "Traditional districting principles," however, first entered the Court's opinions in *Shaw I* and the Court continued to emphasize them in *Miller* v.

---

[234] At the time of writing, this data series is incomplete. The current series omits compactness scores for districts from 1913-1960 and omits intra-decadal redistrictings for compactness and malapportionment scores after 1913.

Johnson (1995) and in *Bush* v. *Vera* (1996). The way that the Court uses traditional districting principles, however, has changed even through these recent cases. (For the remainder of this paper, I will refer to "traditional districting principles" as t.d.p.'s and to districts that violate them as "ugly" districts, whether or not the districts actually appear irregular on a map.)

The Court, legal scholars, and political scientists, have long claimed that lack of compactness, contiguity and population inequality were the evidence and/or instruments of partisan, incumbent and minority vote dilution. (See Chapter 2.) In Chapter 1, I argued that the current Court has departed from previous jurisprudence and now claims two additional roles for t.d.p.'s: First, in *Shaw I* and in *Bush* v. *Vera* (1996), the Court claims that violations of t.d.p.'s cause symbolic and "expressive" harm by sending "pernicious messages" to both voters and politicians. Second, in *Miller* v. *Johnson* (1995) and in *Shaw II*, the court treats violation of traditional districting principles in majority-minority districts as circumstantial evidence of "racial classification."

There is no need here to recreate the arguments of previous chapters, but a brief review of the Court's use of t.d.p.'s is in order. In *Shaw I* the Court stated that redistricting legislation that is "so bizarre on its face that it is unexplainable on grounds other than race" demands close scrutiny. Although the Court stressed that compactness, contiguity and other criteria were not constitutionally required, they indicated that these are "objective factors that may serve to defeat a claim that a district has been gerrymandered on racial lines." For the majority in this case, reapportionment was an area where "appearances do matter" because, in the majority's view, districts that

separate people by race while disregarding political and geographic boundaries reinforce the *perception* that members of the same race necessarily share political views. Such districts also send the "pernicious" message to politicians that they should only represent the majority voting group in the district. In *Shaw,* violation of t.d.p.'s is an integral part of the harm perceived by the Court — violation of these principles *actively* cause harm by sending a pernicious message to politicians. (Pildes and Niemi**,** in their oft-cited 1993 article, both give shape to the implicit and inchoate theory of the Court and name it *expressive harm* (Pildes and Niemi 1993). Adopting this terminology, O'Connor explicitly refers to "expressive harm," in the plurality decision for *Vera*.)

In *Miller,* t.d.p.'s are still important, but they are no longer an integral part of the harm caused by redistricting. Instead, violations of these principles act as merely circumstantial evidence of intent:

> The plaintiff's burden is to show, either through circumstantial evidence of a district's shape and demographics *or more direct evidence going to legislative purpose,* that race was the predominant factor motivating the legislature's decision to place a significant number of voters within or without a particular district (emphasis added).

Similarly, in *ShawII* Justice Rehnquist maintains this use of traditional criteria as circumstantial evidence, writing that "The plaintiff bears the burden of proving the race-based motive and may do so either through circumstantial evidence of a district's shape and demographics or through more direct evidence going to legislative purpose." (116 S.

Ct. 1900) Justice Rehnquist explicitly denies Justice Stevens's dissenting claim that adherence to traditional districting principle isolates a case from strict scrutiny:

> In his dissent, Justice Stevens argues that strict scrutiny does not apply where a State respects or compl(ies) with traditional districting principles … That, however, is not the standard announced and applied in *Miller*, where we held that strict scrutiny applies when race is the Apredominant≅ consideration in drawing the district lines… such that Athe legislature subordinate(s) race-neutral districting principles... to racial considerations.≅

In *Vera*, which was delivered concurrently with *Shaw II*, the role of t.d.p.'s changed yet again — compliance with t.d.p.'s is not merely one piece in a body of circumstantial evidence, but is used as a threshold requirement for strict scrutiny. We can avoid strict scrutiny altogether, even if we are motivated by race, if we pay reasonable[235] attention to "traditional districting criteria": "We do not hold that any one of these factors is independently sufficient to require strict scrutiny. The Constitution does not mandate

---

[235] The court does not require that districts be drawn strictly to follow these criteria: "We thus reject, as impossibly stringent, the District Court's view of the narrow tailoring requirement, that a district must have the least possible amount of irregularity in shape, making allowances for traditional districting criteria."

regularity of district shape... and *the neglect of traditional districting criteria is merely necessary*, not sufficient" (116 S. Ct. 1953, emphasis added).[236] O'Connor, delivering the judgment of the court, stresses this point repeatedly, the same point that Rehnquist disavows :"Under our cases, the States retain a flexibility that federal courts enforcing 2 lack, both insofar as they *may avoid strict scrutiny altogether* by respecting their own t.d.p.'s"

6.1.2. Why Might Traditional Districting Principles Matter (To Political Scientists)?

Tradition may sometimes bear its own weight in the courts, but political scientists, as a rule, are interested in results. The Court posits two connections between t.d.p.'s and politics: They claim that violation of t.d.p.'s cause expressive harm that changes the behavior of voters and politicians, and that these violations are good proxies of gerrymanders. Not surprisingly, many political scientists and other scholars have made additional claims.

Many advocates of the control of gerrymandering through formal principles (traditional and otherwise) have singled out the "traditional" principle of compactness as

---

[236] Again, in *Abrams* v. *Johnson* (1997), the Court stressed the importance of traditional districting principles. Although it dealt with t.d.p.'s in less detail than the other cases citted, it weighed t.d.p's and especially the principle of following traditional county boundaries, in upholding the districts drawn to replace those invalidated by *Miller*.

especially potent. Many different compactness standards have been developed for this purpose; Niemi (Niemi, et al. 1990) provides a thorough survey. Advocates of compactness commonly claim that compactness will reduce or prevent partisan gerrymandering, and some claim that compactness standards will end gerrymandering, without harming representation (Stern 1974). Other advocates of compactness have made similar, but weaker, claims.

The effects that are claimed for various t.d.p.'s include, but are not limited to, the prevention of gerrymandering: Stern (1974) argues that compact districts, by limiting gerrymandering, will increase district competitiveness. Polsby and Popper, in an article in the *Yale Law & Policy Review* in 1991, claim that compactness, contiguity and population equality will together make "the gerrymanderer's life a living hell" (Polsby and Popper 1991). In a later article, they claim in addition that compactness will benefit geographically concentrated minorities (Polsby and Popper 1993).

O'Rourke argues that, in practice, districts that split county boundaries raise campaign costs and that districts that split precincts confuse both election officials and voters (O'Rourke 1995). Related to this is a hypothesis, originally put forward by Bernard Grofman, that violations of traditional districting criteria indicate severe violations of "cognizability" (Grofman 1985; Grofman 1993)[237]. Grofman's claims that

---

[237] Cognizability is, roughly, the ability to easily describe and identify the district. "Egregious violations of the cognizability principle can be identified by making use of

violations of cognizability can weaken voters' attachment to districts and legislators, strengthen incumbency advantage, and decrease turnout. Opponents of compactness and other t.d.p.'s dispute the claim that these restrictions will reduce gerrymandering without negative consequences. Cain shows that "esthetic" geographical criteria can interfere with "good government" criteria such as unbiasedness and competitiveness, and argues against the former on those grounds (Cain 1984, ch. 3). Lowenstein contends that compactness is not a fair standard because he believes that geographically compact districts systematically bias district plans in favor of Republicans (Lowenstein and Steinberg 1985). Karlan, in a pre-*Shaw* article argues in general that compactness is a limited principle and that violations of compactness may be necessary to prevent minority vote dilution (Karlan 1989); post-*Shaw* she specifically disputes the Court's claims that violation of t.d.p.'s cause "expressive" harms (Karlan 1993).

Pro and con, much has been claimed about the relationships between t.d.p.'s and political harms, but little has been empirically tested. Erikson searched for, but did not find, a relationship between electoral bias and malapportionment in the 1960's

---

standard criteria of districting, such as violation of natural geographic boundaries, grossly unnecessary divisions of local sub-unit boundaries (such as city and county lines), and sundering of proximate and contiguous natural communities of interest." (Grofman 1993, 1263)

congressional elections[238] (Erikson 1972), but few other attempts have been made to connect districting principles with electoral effects. If any proponent of, or even some opponents of, t.d.p.'s are right, then districting plans that follow t.d.p.'s should be politically different from districting plans that violate t.d.p.'s. In general, are districting plans that are compact, contiguous, and well-apportioned better? Do such plans produce more competitive races? Are such plans, overall, more responsive to swings in partisan voting? Are they less biased? Do voters in such districts find their districts more "cognizable" — do such voters turn out to vote at a higher rate, vote less frequently for the incumbent, or have more trust in government or their representative? Do Representatives from such districts act differently?

To answer these questions required that I confront several hurdles. In previous chapters, I have developed ways of measuring t.d.p.'s quantitatively, obtained data sources that would allow me to assess t.d.p.'s, and use simulation to make prediction about the effects of compactness on partisan outcomes. In this final chapter, I put these measurements together with electoral and demographic data and use statistical models to analyze the effects of t.d.p.'s on electoral outcomes.

---

[238] Erikson found that reapportionment after the *Baker* v. *Carr* unintentionally advantaged the Republicans, but found no connection between this advantage and t.d.p.'s.

6.2. Data Sources

In Chapter 3, I measured the extent to which all U.S. decadal redistricting plans followed t.d.p.'s. This data is necessary, but not sufficient. In this chapter, I have combined my data on t.d.p.'s with numerous sets of data describing the other half of the relationship, political outcomes. Fortunately, these sets of data are more widely available and more readily accessible than data on t.d.p.'s.

For the political data used in this paper, I turned to a number of standard publicly available sources, in addition to those documented in Chapter 3. For records of voter turnout, and to compute the competitiveness, bias and responsiveness of district plans, I used election returns from I.C.P.S.R. [Inter-university Consortium for Political and Social Research, 1972 #82; King, 1994 #81. I relied on the latter source for electoral data after 1900, as it has undergone more extensive checking and error correction than the former source. The disadvantage of this source, however, is that it does not include third-party turnout in races which were primarily bipartisan. To evaluate voter trust I used survey data from the *American National Election Studies* (Miller and Studies. 1995). For detailed demographic and electoral data in the 1990's, I relied on U.S. Census STF3A data and *Politics in America* (Congressional Quarterly Staff 1996; U.S. Dept. of Commerce 1992).

6.3. Geographic and Population Measures

In this section I review the measures of t.d.p.'s used in the main analysis. These measures are the same as were used in the historical analysis of Chapter 3.[239]

For this data, most of the methods for measuring malapportionment give very similar evaluations of districts. In the main part of this analysis, I focus on the common and understandable measure, the population coefficient of variation.[240]

I adopt the same methods used to categorize contiguity and violations of traditional districting principles, as I have used in previous chapters. Roughly, contiguity refers to mathematical non-single-point contiguity, and "traditional boundaries" are equated with town, county and city lines. As I showed in Chapter 3, however measuring contiguity and breaches of "traditional boundaries" is more complicated in practice than in theory – Chapter 3 documents the (relatively small) number of instances where classification was questionable.

---

[239] In addition to these measures, a number of other less popular measures are used in the section comparing geographic and population measures (Section 4.1), so that the comparison between types of measures is broader.

[240] This is the standard deviation divided by the mean. Think of it as a measure of the average deviation.

6.4. Geography, Population and Politics

6.4.1. Can Compactness and Malapportionment be Measured Consistently?

As I showed in Chapter 3, trends in the numbers of non-contiguous districts and in the number of districts that split traditional boundaries are dramatic, even when difficulties in classifying such difficulties are accounted for. This section compares measurements of malapportionment and compactness. In this section we will see that malapportionment can be measured by a number of different methods, yet still yield the same rankings over districts. Compactness, however, is different. Compactness measures do not seem to be measuring the same thing.

For the purposes of this section, in order to make the comparisons more broadly applicable, I have added a number of additional popular measures of compactness and several popular measures of malapportionment. In order to compare the compactness and malapportionment scores over the same set of districts, this section focuses on the 1st through 62nd congresses, as both compactness and malapportionment data were available for this period.

There are a number of different methods to measure malapportionment used in the scholarly literature and by the courts. The most popular early measures of population variation were the difference between the largest and smallest districts (divided by the mean), the population variance ratio, which is the ratio of the largest to the smallest district, the maximum (or average) percent deviation from the mean, and the “electoral percentage,” which is the minimum percentage of the population represented by a bare

majority of seats.[241] Later court cases have tended to stress the difference between the largest and smallest districts, and this measure was emphasized in *Karcher* v. *Daggett* (1983). These measures have a number of theoretical faults, and Foster (1985) argues cogently that such measures as the Gini index, Theil's measure of entropy, and the coefficient of variation have more desirable theoretical properties.

I included the "electoral percentage" for the sake of completeness, not because it is a particularly good measure. There are a number of problems that result from applying the "electoral percentage" to congressional district plans of individual states. First, as Dixon (1968) argues, this measure runs into numerical problems when a plan has few districts — in a state with two congressional districts the "electoral percentage" is necessarily 100 percent. Dixon also argues (See Chapter 7, Section 7.a of Dixon.) that this measure does not adequately reflect the realities of controlling an election, and I agree with him. It is also true that the other population measures fail to directly assess political power; but unlike the electoral percentage, the other measures can be argued to capture the formal, intrinsic value of an individual vote (Lowenstein 1990). Last, the electoral percentage was defined with legislative districts in mind, and when we apply it instead to congressional districts it loses much of its political meaning, since capturing a bare majority of a state's congressional seats does not have the same direct political implications as does capturing a legislature. Consequently, I expected that the electoral

---

[241] See Dixon (1968) Chapter 17, Section 4 and Chapter 18, Section 2.

percentage would perform poorly in these circumstances and would not be consistent with the other measures of malapportionment.

In the rest of this study, I have used three measures of compactness. For comparison in this section, I include in this section a fourth, and very common, measure, the direct area-perimeter (AP) comparison, defined as A/0.282P (Flaherty and Crumplin 1992).

	(Max-Min) /Mean	*Pop. Variance Ratio*	*Coeff. of Variation*	*Gini Coeff*	*Thiel's Entropy*
Pop. Variance Ratio	0.96				
Coeff. of Variation	0.77	0.76			
Gini Coeff.	0.81	0.80	0.89		
Theil's Entropy	0.80	0.79	0.91	0.94	
Electoral Percentage	-0.28	-0.28	-0.14	-0.18	-0.16

Table 6-1. Kendall's τ_B between population equality scores for multi-district plans: 1-62nd Congresses (349 observations).

Surprisingly, despite theoretical qualms about how best to measure malapportionment, all but one of these measures seem to be measuring the same underlying concept (Table 6-1). The courts can use any popular measure of malapportionment and they will come to the same conclusions in the vast majority of cases. This should give us confidence that these standards are being applied consistently, and that these standards are judicially manageable.

Unfortunately, we cannot be so confident about measuring compactness. Compactness measures do not seem to be measuring the same thing. Many of the measures of compactness in this study disagree more often than not. Table 6-2 and Table

6-3 show the level of agreement (τ_B) over scores for different compactness plans, while Table 6-4 shows the level of agreement among scores for individual districts; neither shows the same level of consistency that was shown by malapportionment measures.[242]

	AC	*AP*	*LW*
AP	0.08		
LW	0.56	0.09	
NORM	0.53	0.07	0.27

Table 6-2. Kendall's τ_B between mean compactness scores for multi-district plans: 1-62nd Congresses (349 observations).

[242] Here, my results disagree with those in (Niemi and Wilkerson 1991). There are five reasons for this disagreement: Niemi, et al. compare some other compactness measures that come from the same class of general measures, they use a much more limited range of district data, their district plans have more districts (15–100 districts) than the average congressional plan, and they use two measures of correlation, Pearson's p and Spearman's rho, that measure agreement less directly than Kendall's Tau. See Liebtrau (1983) for a discussion of the relative merits of these measures.

	AC	*AP*	*LW*
AP	0.23		
LW	0.51	0.29	
NORM	0.45	0.26	0.16

Table 6-3. Kendall's τ_B between minimum compactness scores for multi-district plans: 1-62nd Congresses (349 observations). Here, the compactness of a plan is defined by its least compact district.

	AC	*AP*	*LW*
AP	0.11		
LW	0.54	0.08	
NORM	0.51	0.09	0.28

Table 6-4. Kendall's τ_B between compactness scores for individual districts: 1-62nd Congresses (3390 observations).

In Chapter 2, I had noted a number of theoretical objections to the AP measure. We see, in practice, that the AP measure has little in common with any other measure, as might be expected from a standard that is entirely arbitrary.

Furthermore, although in theory sufficiently compact districts will be contiguous,[243] in practice compactness and contiguity are often at odds. Table 6-5 shows two different

---

[243] Compactness implies contiguity only for "well-behaved" measures of compactness,

measures of concordance between contiguity and compactness scores.[244] While I have included Kendall's measure for consistency in presentation, all but 58 districts are contiguous, which causes an extreme percentage of ties across the contiguity category and distorts Kendall's measure. A more appropriate measure in this case is Somer's d which is very similar to Kendall's measure except that it looks only at pairs of districts which differ in their levels of contiguity.

	τ_B	*Somer's d*
AC	0.06	0.32
AP	0.08	0.44
LW	0.05	0.28
NORM	0.06	0.34

Table 6-5. Comparison between level of contiguity (contiguous, questionable, and non-contiguous) and compactness scores for individual districts: 1-62nd Congresses (3390 observations). All but 58 of these districts were practically contiguous.

Since both Somer's d and Kendall's Tau are positive, the more compact a set of districts, the more likely they are to be contiguous. This relationship is weak, however.

---

and even then only at some threshold level. See Altman 1995.

[244] I divided districts into three categories of contiguity, as described in Chapter 3.

Using Somer's d, we can see that when we compare two districts with different levels of contiguity, the least contiguous district will often be the more compact of the two.[245]

6.4.2. Relationships

Clearly, things have changed since the Court imposed, and strengthened, equal-population requirements. Fewer districts are contiguous, more districts split traditional boundaries, and districts are, at least by perimeter scores, less compact than they once were. Does it matter? Do violations of t.d.p.'s, which have clearly become more common, commonly affect politics? In this section, I test several of the most popular hypotheses regarding the political consequences of t.d.p.'s.[246]

---

[245] Both the Tau B measure and Somer's d use all 3390 observations to generate comparison pairs. Somer's d, however, effectively discards *pairs* where both members of the pair are contiguous (or both are equally non-contiguous).

[246] I have not tested all hypotheses related to the electoral effects of traditional district criteria. I would like to note three untested hypotheses here: First, I have not tested the effects of t.d.p.'s on incumbency advantage. Although I find that t.d.p.'s do not affect electoral responsiveness on the whole, it is possible that they affect the propensity for a voter to choose the incumbent. Second, I have not tested the hypothesis that "ugly" districts cause members of Congress to behave differently toward their electoral constituency. Third, the court claims that "ugliness" is of particular importance in

As there was a change in the rules of the game and in the appearance of districts in the 1960's, I focus on the period from 1962-1994. All districts were included in the analyses, except multi-member and at-large districts and a few others noted in the Appendix. Compactness scores, which are the focus of the most attention of all t.d.p.'s, are recorded for all four decadal redistrictings in this period.[247] Since the measurement of malapportionment becomes problematic after equal population standards were applied, as noted earlier, malapportionment scores are omitted from the 1970's decadal redistricting onward. Because of limits in the data sources, intra-decadal redistrictings are omitted, and violations of "traditional" boundaries are omitted from the 1980's decadal redistricting.

T.D.P.'s and Voter Attitudes

---

majority-minority districts. I have included, but not singled, out these districts in my analysis.

[247] I report scores for the normalized perimeter and the area measure, the most popular measures of compactness. All tests were duplicated with the length-width measure and with logged compactness scores on all measures, but the results were not substantially different.

Justice O'Connor, writing for the court,[248] contends that violations of t.d.p.'s, especially in majority-minority districts, sends "pernicious messages" to representatives and voters in such districts. Is there evidence for this contention? Do voters in ugly districts act or feel differently than voters in other districts? Do they trust government less or feel that Congress is less representative? Do they turn out to vote in fewer numbers in congressional elections?

The Court claims that members of Congress from ugly districts will represent only (or to a relatively greater extent) those who elected them. Do members of Congress from ugly districts represent only the winning part of the constituency? Do members from other districts represent everyone in the district, as the Court implies?

Given a particular group of voters with strong and easily identifiable interests, it is possible to directly measure the extent to which each member of Congress votes for the interests of these voters (Cameron, et al. 1996). In general, however, I cannot assume that I know the interest of a district's various constituents *a priori*, and survey data that could be used to estimate the interests of district constituents is incomplete or unavailable. We might expect, however, that the partisan competitiveness of the district should be a rough indicator of the "moderateness" of the district's constituents.

---

[248] In addition to Justice O'Connor, a number of legal scholars argue that redistricting can cause symbolic and expressive harms. (See Chapter 1.)

Accordingly, I constructed a measure of the partisan competitiveness of each district by taking the absolute value of the difference between partisan vote percentages for president in 1992. I used Poole-Rosenthal W-nominate scores to calculate the extremeness of each representative's voting pattern in the 103rd congress.[249] I regressed this measure, a dummy variable for the party of the representative, and compactness and contiguity measures, on the extremeness of the representative.

---

[249] Poole-Rosenthal scores measure the relative location of politicians in a multi-dimensional issue space. They tend to measure ideology better than interest group scores, and are the best currently known predictor of roll-call voting.

I defined the extremity of the representative as the absolute difference between that representatives W-nominate$_1$ score and the median score of the congress (Poole and Rosenthal 1991).

Poole-Rosenthal score data can be freely obtained from the authors through their web-site.

Intercept	0.52*	(0.03)
Democrat	-0.43*	(0.02)
Competitiveness	0.01*	(0.00)
Hispanic %	0.01	(0.06)
Black %	-0.14	(0.07)
Questionable Contiguity	0.05	(0.03)
Normalized Perimeter Compactness	0.00	(0.08)
Area Compactness	0.13	(0.09)
	Adj. R-squared	.66
	N. Obs.	426

Table 6-6. Effect of compactness on extremeness of representative.

None of the compactness parameters is significant at the $p<0.05$ level (two-tailed), although area compactness is significant at the $p<0.10$ level. The sign on the area compactness variable is the reverse of what advocates of compactness expect. In other words, more compact districts, if they have an effect, tend to increase the extremeness of voting behavior.

A regression of the W-nominate scores using demographic variables, and geographic terms *interacted* with party shows that the effect of compactness is more significant for Republicans than for Democrats, and that perimeter and area compactness may have opposite effects. This table shows that for Republicans, perimeter compactness is correlated with more centrist opinions, but area compactness is correlated with more extremist opinions (Table 6-7).

Intercept	0.47*	0.047
Democrat	-0.81*	0.058
Competitiveness	0.00*	0.001
Hispanic %	-0.08	0.068
Black%	-0.09	0.082
Area Compactness (Republican)	0.32*	0.153
Normalized Perimeter Compactness(Republican)	-0.33*	0.132
Questionable Contiguity (Republican)	-0.00	0.096
c_ac (Democrat)	-0.21	0.206
Normalized Perimeter Compactness(Democrat)	0.24	0.177
Questionable Contiguity (Democrat)	-0.07	0.109
	Adj. R-squared	0.85

Table 6-7 Effect of compactness interacted with party on issue position of representative.

The Court contends that "ugly" districts send a pernicious message to voters that political identity should be racial, a message that they claim threatens to deeply divide American society. It is natural to assume that such virulent messages should have some affect on voters' attitudes toward government and Congress. As a test of whether voters in "ugly" districts felt differently about government or Congress, I used individual survey-response data on trust in government and feelings toward Congress from the American National Election Survey. I estimated the effect of t.d.p.'s on the responses of potential voters to three questions[250] about congressional responsiveness, and general trust in government:

---

[250] Surveys of the size necessary to detect district-level effects are rare, and the range of available questions is limited: These questions do not ask about each voter's particular

- Question 610, which reads: "Generally speaking, those we elect to Congress in Washington lose touch with the people pretty quickly. (Agree/Disagree)" Higher values indicate disagreement.

- Question 625, which reads: "How much attention do you think most congressmen pay to the people who elect them when they decide what to do in Congress, a good deal, some, or not much?" Higher values represent more attention.

- Question 626, which is not an individual question, but a five-point index derived from five other questions. Higher values represent more trust in government.

To evaluate the influence of t.d.p.'s on voter's attitudes towards government, I used an ordered probit models with the survey responses as dependent variables.[251] (See, for example, Maddala 1983.) Where possible, I analyzed the first two surveys after each decadal redistricting. Some questions were, however, asked only during certain years. The results of this analysis are presented below (Table 6-8, Table 6-9).

---

representative, nor were all of them asked over the entire period of the N.E.S. Furthermore, even though the National Election Survey is the largest and most detailed survey to address congressional representation, the number of survey participants from each district is small.

[251] O.L.S. is often used for this purpose as well. O.L.S. analysis of ordered categorical variables can be biased if the interval between each pair of categories is not equal. The ordered probit model allows for unequal intervals among categories.

	Trust					
	1964	*1972*	*1974*	*1984*	*1992*	*1994*
Intercept	2.47 (0.10)	0.70 (0.08)	0.42 (0.098)	0.65 (0.082)	0.44 (0.071)	0.25 (0.08)
Percent Deviation from Mean District Population	0.29 (0.11)	--------	--------	--------	--------	--------
Split Sub-Unit Boundaries	0.02 (0.11)	-0.066 (0.055)	0.035 (0.066)	--------	--------	--------
Normalized Perimeter Compactness	-0.46 (0.27)	0.28 (0.25)	0.07 (0.31)	-0.34 (0.24)	-0.39 (0.27)	0.048 (0.25)
Area Compactness	-0.40 (0.30)	-0.27 (0.28)	-0.37 (0.35)	0.52 (0.30)	-0.037 (0.23)	-0.28 (0.30)
Questionable Contiguity or Discontiguous	-0.14 (0.14)	-0.067 (0.71)	-0.64 (0.81)	none in Sample	-0.03 (0.10)	0.13 (0.10)
Number of Observation	1356	2204	1532	1814	2204	1741
Correctly Predicted /% pred. with intercept only	43% /43%	24% /25%	38% /38%	24% /24%	37% /37%	44% /42%
Likelihood	1942	3500	2300	2900	3300	2500

Table 6-8. Relationship between T.D.P.'s and perceived congressional responsiveness and trust in government. The dependent variable in each model is an individual categorical survey response for a question on the ANES survey. The independent variables are measured district-level characteristics. Dashed lines indicate that the independent variable was missing for that year. Standard errors are reported in parentheses. Starred variables are significant at 0.05% (two-tailed test).

Attention				**Lose Touch**	
	1964	*1972*	*1974*	*1972*	*1974*
Intercept	1.1	0.91	0.60	-0.39	-0.44
	(0.10)	(0.088)	(0.10)	(0.09)	(0.12)
Percent Deviation from	0.41	--------	--------	--------	--------
Mean District Population	(0.13)				
Split Sub-Unit Boundaries	0.072	-0.021	0.013	-0.12	-0.03
	(0.11)	(0.61)	(0.08)	(0.063)	(0.082)
Normalized Perimeter	0.31	0.067	0.06	0.64	0.32
Compactness	(0.32)	(0.27)	(0.33)	(0.28)	(0.39)
Area Compactness	0.58	-0.21	0.02	-0.78	-0.45
	(0.29)	(0.30)	(0.36)	(0.32)	(0.43)
Questionable Contiguity or	-0.015	0.74	0.98	-2.8	-3.0
Discontiguous	(0.16)	(0.83)	(0.84)	(12.0)	(15.0)
Number of Observation	1356	2701	1476	1452	1452
Correctly Predicted	43%	52%	53%	69%	69%
/% pred. with intercept only	/43%	/52%	/53%	/69%	/69%
Likelihood	1942	2100	1500	900	900

Table 6-9. Relationship between T.D.P.'s and perceived congressional responsiveness and trust in government. The dependent variable in each model is an individual categorical survey response for a question on the ANES survey. The independent variables are measured district-level characteristics. Dashed lines indicate that the independent variable was missing for that year. Standard errors are reported in parentheses. Starred variables are significant at 0.05% (two-tailed test).

This table lends no support to the hypothesis that t.d.p's have a significant affect on voters' attitudes toward Congress or on voters' trust of government. Although I have omitted the additional tables for brevity, I ran additional analyses adding race and other demographic variables from the N.E.S. survey as additional independent variables, and

found the same results. Although the addition of demographic variables improves the predictive ability of the model, t.d.p.'s remain insignificant (at any conventional level).

Although this data does not support the conclusion, it is still of, course, statistically possible that t.d.p.'s have "expressive" effects. The burden of proof hould, however, be considered. Both legal and scientific tradition holds that we should require evidence before accepting a new claim. Quite simply, no evidence has previously been offered to show that "expressive harms" exist, and this is the first study to attempt to empirically detect them. The lack of evidence found here for expressive harms should signal the court that its attention should be focussed on the well known and easily detectable political effects of redistricting.

T.D.P.'s and Voter Turnout

Although people in "ugly" districts do not admit to feeling different about their member of Congress, they might still behave differently. One obvious way of showing disaffection with the electoral process would be to avoid the voting booth, and previous authors have hypothesized that ugly districts would decrease turnout.[252] Is voter turnout lower in ugly districts?

---

[252] More precisely, previous authors have argued that ugly districts proxy "cognizability," which affects turnout. (See Section 6.1.)

Voter turnout is known to be influenced by demographic factors, regional factors, and electoral factors. Previous research indicates that turnout is influenced by such diverse and interconnected variables as race, region, residential mobility, education, age, income, registration laws, campaign spending and campaign advertising (Ansolabehere et al. 1994; Cox and Munger 1989; Wolfinger and Rosenstone 1980).

Because of the interconnections between turnout variables, choosing "control" variables is difficult. In order to show how well t.d.p.'s alone can proxy political behavior, I estimate a model of turnout with those variables alone, as well as with demographic variables drawn from the 1990 census.

Data on turnout is available for most districts and election periods only in aggregate form.[253] Since a linear probability function is most commonly used for estimating changes in percent turnout, I show the results of this model (using weighted least squares to implement the minimum chi square method (Maddala 1983, p. 28)) for comparison. Both methods lead to the same substantive conclusions (Table 6-10).

---

[253] Rather than assume a model of turnout based entirely on aggregate variables, I also reanalyzed turnout with an individual-level logit model of the decision to vote, and estimate its values from the aggregate data using a log-odds model of the percent of voting-age-population that turned out for Republican and Democratic candidates in the congressional election (King 1989, pgs. 119 & 139). The substantive implications of both models are the same.

	Turnout as a % of VAP (1992)			*(1994)*		
	Shape and Demo-graphic	*Shape*	*Demo-graphic*	*Shape and Demo-graphic*	*Shape*	*Demo-graphic*
Intercept	0.71* (0.21)	0.45* (0.01)	0.79* (0.21)	0.45* (0.186)	0.34* (0.012)	0.46* (0.190)
Normalized Perimeter Compactness	0.087* (0.03)	0.11* (0.047)	------	0.12* (0.029)	0.16* (0.038)	------
Area Compactness	0.038 (0.038)	0.09 (0.052)	------	-0.02 (0.034)	0.03 (0.044)	------
Bad or Questionable Contiguity	-0.042* (0.013)	-0.066* (0.018)	------	0.00 (0.015)	-0.04 (0.020)	------
Percent Black	-0.21* (0.024)	------	-0.25* (0.024)	-0.21* (0.021)	------	-0.24* (0.021)
Percent Hispanic	-0.41* (0.026)	------	-0.43* (0.026)	-0.27* (0.022)	------	-0.27* (0.023)
Percent College+	0.19* (0.06)	------	0.17* (0.065)	0.18* (0.058)	------	0.16* (0.059)
log(Median Income)	-0.020 (0.021)	------	-0.024 (0.022)	-0.01 (0.019)	------	-0.01 (0.019)
Open Seat	0.009* (0.0084)	------	-0.01 (0.0089)	0.02* (0.012)	------	0.03* (0.012)
Adjusted R^2	0.53	0.11	0.49	0.47	0.09	0.44

Table 6-10. Relationship between t.d.p.'s and percent turnout in 1992 and 1994 elections. The independent variables are measured district-level characteristics. Dashed lines indicate that the independent variable was missing for that year. Standard errors (approximate) are reported in parentheses. At-large districts and uncontested districts in which elections were not held are omitted (417 observations).Starred variables are significant at 0.05% (two-tailed test).

These results tend to support O'Rourke's and Grofman's hypotheses, at least strictly: Ill-compact and non-contiguous districts tend to have lower turnout, even when other demographic factors are included in the model. The effects of compactness and contiguity are, however, dwarfed by demographic factors.

Partisan Bias and Electoral Responsiveness

Regardless of the non-effect of t.d.p.'s on individual voter behavior, do t.d.p.'s affect elections in the aggregate? Would elections be different if districts were "perfect"? In the final part of this section, I use Gelman's and King's method of evaluating electoral plans (Gelman and King 1994) and their *Judgeit* program to estimate the change in bias and electoral responsiveness[254] that would have resulted had districts been contiguous, evenly-apportioned (in 1962), compact, and aligned along traditional boundaries (this in 1962 and 1972 elections only).

I use Judgeit to estimate bias and responsiveness in each election using election returns from that election and the following election. I include, as "predictive" variables,

---

[254] Bias and responsiveness have a long history as measures of electoral results (Niemi and Deegan 1978). Responsiveness is, in essence, the elasticity of the seats-votes curve around its midpoint, and bias is a measure of the asymmetry of that curve.

all of the t.d.p. variables that were available for that decade. I then use *Judgeit* to evaluate a hypothetical election in which all districts are perfectly 'pretty'.[255]

Obviously, it is not possible, in real life, for all districts to be perfect on all scores (notably compactness scores). Nor do the resulting hypothetical election results capture what would happen if districts were really perfect. What the hypothetical election results show is the statistical predictive relationship between t.d.p.'s and electoral results.

---

[255] More precise results could be obtained with additional data. In particular, because of limits in the present data-set, it is not possible to do analyses of compactness criteria between 1913-1961, or of intra-decadal redistrictings after 1913. Since the electoral effects of redistricting are strongest immediately after redistricting (Gelman and King 1994), however, the omission of intra-decadal data should tend to increase our estimates of the effect of compactness criteria, if anything. (Without the intra-decadal data, we cannot know if changes in compactness were greater than decadal changes. Nevertheless, the range of the independent variables in the decadal offers a rich range of compactness and malapportionment scores.)

	1962 *(Observed)*	*(Ideal)*	*(I-O)*	**1972** *(Observed)*	*(Ideal)*	*(I-O)*
Responsiveness at Vbar=0.45 to0.55	1.8 (0.063)	1.8 (0.07)	-0.03	1.4 (0.058)	1.4 (0.07)	0.017
Responsiveness at Vbar=Observed	1.8 (0.12)	1.8 (0.12)	-0.01	1.4 (0.11)	1.4 (0.11)	0.0065
Bias at Vbar=0.45 to0.55	-0.21 (0.0052)	-0.013 (0.0057)	0.20*	0.0036 (0.0040)	0.0032 (0.0056)	-0.0004
Bias at Vbar=050	-0.22 (0.0058)	-0.013 (0.0063)	0.20*	0.0039 (0.0044)	0.0036 (0.0060)	-0.0003

	1982 *(Observed)*	*(Ideal)*	*(I-O)*	**1992** *(Observed)*	*(Ideal)*	*(I-O)*
Responsiveness at Vbar=0.45 to0.55	1.7 (0.076)	1.6 (0.071)	-0.021	2.0 (0.067)	2.0 (0.085)	0.038
Responsiveness at Vbar=Observed	1.6 (0.13)	1.6 (0.12)	-0.01	1.8 (0.11)	1.8 (0.14)	0.019
Bias at Vbar=0.45 to0.55	-0.0059 (0.0052)	-0.0031 (0.0066)	0.0028	0.014 (0.0053)	0.014 (0.0054)	0.0006
Bias at Vbar=050	-0.0049 (0.0059)	-0.0023 (0.0072)	0.0026	0.015 (0.0058)	0.016 (0.006)	0.0003

Table 6-11. Comparison of observed congressional bias (positive values indicate a Democratic bias) and responsiveness in first elections following decadal redistrict, compared to "ideal" districts that always follow t.d.p.'s "At-large" districts and uncontested districts that did not hold elections were excluded. Significance is calculated for the columns of differences between ideal and observed, where starred variables are significant at 0.05% (two-tailed test).

The results in Table 6-11 indicate that in the elections immediately following the 1960, 1970, and 1980 decadal redistrictings, t.d.p.'s did not have a large effect. If districts had been ideally shaped, followed all traditional boundaries, and been evenly apportioned, then the 1960's elections would have been somewhat less biased against the Democrats, but little else would have changed. Districting plans, on the whole, would be no more responsive to partisan shifts, and bias in 1972 and in 1982 would not have been significantly altered.

6.5. Discussion

Despite theoretical misgivings about some popular measures of population inequality, all but one of the malapportionment measures that I tested produced very similar evaluations over a large set of real districts. In effect, the courts can use almost any convenient measure of population equality and obtain consistent results. Because of this, the courts were free to pick easily calculable and manageable population measures. This choice of measures was relatively uncontroversial, and the courts' resulting measurements of malapportionment have been predictable.

Unfortunately, this cannot be said of compactness measures. The measures of compactness in this study, which are typical examples of the three most popular types of geographical compactness measures, disagree more often than not, and compactness measures can disagree with assessments of practical contiguity.

Population equality is often held to be intrinsically valuable. The principle of "one person one vote" is enforced by the courts on its own merit, as well as in its instrumental

role in creating fair outcomes. This intrinsic value lends strength to the concept of population equality and seems to make it easier to measure. On the other hand, if there is some intrinsic value to geographic compactness, the majority of proponents for compactness leave it unarticulated, and this study shows that numeric measures fail to agree on it.

This disagreement among measures raises concerns about the judicial manageability of compactness standards. The disagreement that I find among compactness evaluations contradicts Polsby's and Popper's (1991) assertion that compactness measures generally lead to the same results, and it supports Young's (1988), conclusion that compactness is still a "hazy and ill-defined concept" (pg. 114), and Lowenstein's and Steinbergs's (1985) contention that measuring compactness is neither simple nor straightforward.

Districts have changed in recent decades. Although in practically every decadal apportionment from the first through fifty-eighth Congress, some districts violated practical criteria of contiguity;violations of contiguity were few, and were concentrated in a small number of states (New York, Massachusetts, and the Carolinas). The violations of contiguity in the latest round of redistricting far exceeded traditional baselines.

Modern districting plans exceed their predecessors to an even greater extent in the frequency with which their districts violate town, counties and other sub-unit boundaries. Except in dense urban areas, early districts followed county and town boundaries exclusively. Districts split larger boundaries with increasing frequency after the 43rd

Congress, but the frequency of violations of "traditional boundaries" skyrockets following the population requirements imposed by the Court in *Reynolds* and *Wesberry*.

Of the four formal criteria examined in this paper, compactness has been the most controversial. How to define compactness and whether it can be usefully defined are questions that have been debated since the first compactness requirement passed by Congress. Unlike patterns of population equality and contiguity, patterns of compactness depend on which definition of "compactness" that we use. Whether or not there were traditional norms of compactness, and whether the current round of "ugly" districts violate these norms, depends crucially on the precise method that we choose to measure this property. Still, under a common definition of compactness, the normalized-perimeter score, district compactness has dropped steadily since *Reynolds*.

While districts have changed, these changes seem not to have had a large effect on elections. No one would deny that odd shapes and other violations of "traditional districting principles" can sometimes signal odd intents, but this research indicates that they do so only in the larger context of our political knowledge. Odd lines can indicate gerrymanders to the experienced researcher, politician, or judge, but only because the researcher also knows the political and demographic composition of the area on which the lines were imposed. Formal measurements and formal principles, are not, in general, sufficient to accomplish the task of judging politics.

The Court, and some legal scholars, believe that ugly districts change voter behavior. If ugly districts have an effect, it is small. I find no support for the hypothesis that "ugly"

districts send pernicious messages to voters that affect their attitudes toward government or Congress. And although there is some evidence that, as O'Rourke and Grofman argue, violations of "traditional districting principles" reduce turnout, the effect is dwarfed by other well-known influences on turnout.

It has also been claimed that violations of "traditional districting principles" are proxies for gerrymandering. This research shows that, overall, "traditional districting principles" are not good proxies. A proxy is most useful when it provides us with a cheap way of learning about something that is otherwise expensive or impossible to study directly. While "traditional districting principles" are sometimes slightly predictive of bias and turnout, there are much more effective methods to determine the political effects of redistricting (Gelman and King 1994; Kousser 1995).