

Gerrymandering and Legislator Efficiency

John Mackenzie, College of Agricultural Sciences

University of Delaware, Newark, DE 19717

johnmack@udel.edu

The purposes of gerrymandering are to maximize the number of legislative seats that can be won by the political party in charge of redrawing the district boundaries, and to create “safe” seats for the party’s incumbent legislators. This paper develops a geometric index of gerrymandering, uses it to analyze US Representative district boundaries, and explores a few simple hypotheses. Do safer, more gerrymandered districts yield more seniority in Congress? Do Representatives from gerrymandered districts bring as much federal money back to their districts, *ceteris paribus*, as Representatives from non-gerrymandered districts?

Introduction

Gerrymandering involves the redrawing of election district boundaries to give an electoral advantage to a particular candidate or party. It has been recognized as an American political art form since 1812. The term derives from a redrawing of US Representative districts in Massachusetts before the 1812 elections, when Elbridge Gerry was governor. The story goes that two reporters were discussing the new map, and when one commented that the district north of Boston looked like a salamander, the other exclaimed, “No, a Gerry-mander!” (Figure 1).



Figure 1: Satire of the original 1812 “gerrymander”

The US Constitution specifies that Representative seats should be apportioned among the states based on the results of the decennial Census, but says nothing about how states should draw district boundaries for their Representatives. Beyond the Constitutional principle that districts should have near-equal populations so that citizens are equally represented in the House of Representatives, there are several other generally-accepted principles in drawing “fair” boundaries. District boundaries should make sense to laypeople, preferably following naturally-occurring geographic features that define boundaries between communities. And as geography permits, they should each be comprised of

single polygons with reasonable geometric compactness. But politicians are often willing to ignore these principles, and majority parties in the various state legislatures have been refining the art of gerrymandering ever since the term was coined. US courts have generally been reluctant to intervene in these processes unless they involve overt civil rights issues, and in some cases have actually encouraged creative mappings that promote representation of racial minorities.

While the Constitution guarantees that how you vote is secret, the fact that you did or did not vote in an election is public record. Voter registration data typically include the name, address, phone number, sex, age, party registration and voting history of each registered voter. All of this is made available to parties and political candidates, and most of it is available to the public. These voter databases are now routinely geo-coded, and combined with Census block-level SF1 (short-form) data and block-group-level SF3 (long-form) data from the decennial Census, they can readily support high-precision (house-by-house) gerrymandering in a geographic information system.

Theory of gerrymandering

Gerrymandering depends on a heterogeneous geographic distribution of political interests, and only occurs in elections by district rather than at-large. In the schematic diagram below (Figure 2), representing a geographic area with evenly-dispersed population, suppose the people in the west (40% of the total population) all vote for Party A and the people in the east (60% of the total population) all vote for Party B. If the region is split into compact east and west sections of equal proportions, each party wins one seat. But if regional majority Party B controls the redistricting process, and splits the region into less compact north and south sections of equal proportions, it can win both seats and Party A is effectively unrepresented.

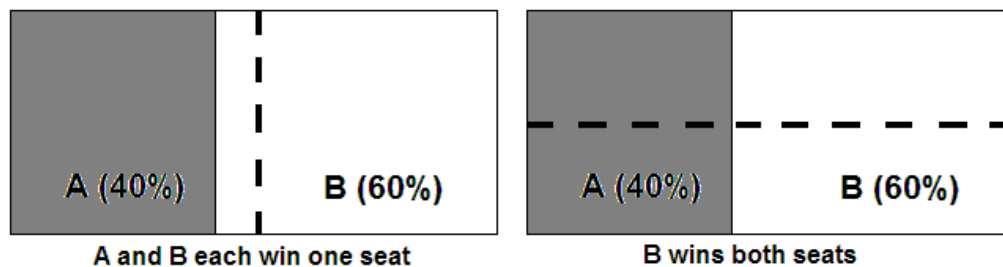


Figure 2: Gerrymandering Schematic

The redistricting problem can be generalized for multiple seats, parties and voting systems. In the two-party case, as the number of districts gets large, the minority party with A% of total voters could theoretically gerrymander districts to win up to 2A% of the seats (e.g., a 26% minority party could eke out wins in up to 52 of 100 districts and gain political control); alternately, a 51% majority party could gerrymander districts to win all the seats. With J equal-sized parties competing for K districts where district winners are determined by simple plurality, the party controlling redistricting could theoretically concede just one seat and sweep all the other seats.

The party in control of redistricting can weaken the opposition party by “packing” as many opposition voters as possible into a minimum number of conceded districts, and/or “cracking” opposition voters among numerous safe districts where they are in the minority. If representatives are required to be residents of their districts, redistricting may “hijack” a district from an incumbent by redrawing the boundary to exclude his house, and/or “kidnap” him into a district where he will lose the next election.

Hypothesized effects of gerrymandering

By pre-determining election outcomes, gerrymandering makes actual voting less consequential, and therefore ought to discourage voter turnout, but I have not found any empirical analysis that demonstrates this. Variations in voter turnout mostly depend on voter age, income, educational attainment, race and ethnicity. Since ballots include multiple races for offices with various geographies, some of which may be closely-contested, one or two “foregone conclusions” on the ballot will not diminish voter interest in other races.

The political efficiency of gerrymandered districts is also unclear. Political economists typically assume that each legislator’s primary incentive is to insure her own re-election, and the typical behaviors of legislators support this assumption. An “effective” legislator delivers large appropriations and other special advantages for the politically-involved constituents in her district, and she extracts substantial contributions to her re-election campaigns from these same constituents. She builds a formidable campaign apparatus to discourage challengers and insure re-election, and gains seniority and legislative clout to sustain her incumbency. Challengers hardly ever beat incumbents except in special circumstances: the incumbent is discredited by some scandal, defeated by hostile gerrymandering, retires or dies.

Political competition, like market competition, is supposed to yield efficiency. In hotly-contested districts, politicians competing to win the support of various constituent groups have to be highly responsive to those groups. In safe districts, well-entrenched legislators, like monopolists, can afford to be lazy. One of the questions I address at the end of this paper is whether legislators in more gerrymandered districts are in fact less “efficient” than other legislators in capturing federal appropriations for their home districts.

Geometric indices of gerrymandering

Any formal analysis of gerrymandering requires some quantitative measurement of it. Researchers have used various geometric measures of polygon complexity—or the inverse, “compactness”—to index the degree to which a political district has been gerrymandered. Niemi et al. (1990) discuss a large number of these measures, e.g., circumscribing the district polygon with a rectangle and calculating the ratio of width to length; calculating the variance of the distances of voters from the interior center of the district; calculating the ratio of the district population to the convex hull of the district or its circumscribing circle or oval; etc.

The most commonly-used measures of polygon complexity are based on perimeter length, or perimeter compared to area. The simplest unit-free, scale-invariant measure is the ratio of a polygon’s area to the square of its perimeter. A circle has maximal compactness of $1/4\pi$ or minimal complexity of $4\pi \approx 12.57$; a square has complexity of 16; an equilateral triangle has complexity of about 20.79, etc.

Numerous researchers have attempted to measure gerrymandering via raw polygon complexity P^2/A , although these measurements do not distinguish natural boundary complexities such as shorelines from the artificial complexities caused by gerrymandering. One rough correction involves dividing each district’s complexity by the boundary complexity of the state in which it is located, but this expedient sacrifices comparability of compactness measures across states. This paper improves upon the raw complexity index by distinguishing artificial from natural district boundaries and scaling the raw compactness measure by the proportion of the perimeter that is politically-drawn.

I obtained digital shapefiles of US Representative district polygons for the 106th, 107th, 108th, 109th and 110th Congresses (1999 through 2008) from the US Census Bureau website (http://www.census.gov/geo/www/cob/bdy_files.html) and used a GIS to calculate polygon areas A (in

square meters) and perimeters P (in meters) for each district. There are minor inconsistencies in the perimeters and areas of equivalent polygons in the different shapefiles because the Census applied slightly different degrees of line simplification to them. The implications of line simplification are discussed later in this paper.

To refine the raw polygon complexity measure as an index of gerrymandering, I used the following geo-processing method to distinguish natural district boundaries (shorelines and state boundaries) from the artificial district boundaries drawn by politicians. First, I dissolved the district polygons to create state polygons that have exact local congruence with district boundaries along state borders. Next, I cracked the district polygons into their component boundary line segments, with the “left” and “right” side district ID’s extracted as attributes for each segment. (The arbitrary “left” and “right” sides of a boundary segment simply reflect the direction in which the segment was digitized.) I then selected the segments that are inside the state polygons and not congruent with a state border (these boundary segments are shown in red in Figure 3), and calculated the lengths of these politically-generated interior segments. The boundary segments that coincide with state borders and are not subject to political manipulation were assigned default interior lengths of zero.

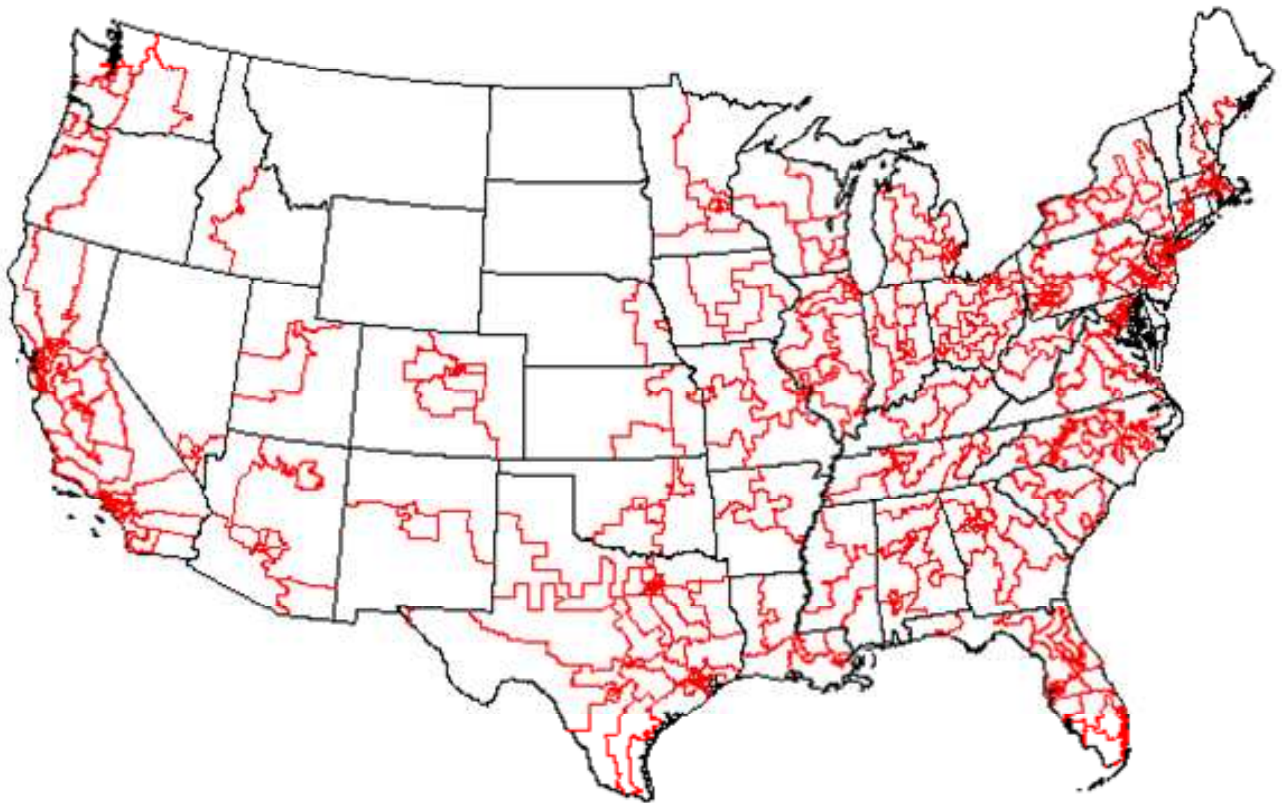


Figure 3: Interior Congressional district boundary segments versus state boundary segments

I summed the interior segment lengths for each “left” side district ID, and for each “right” side district ID, and then summed the left- and right-side aggregate lengths to obtain total interior boundary lengths for each Congressional district polygon. The technical details of this procedure are explained in the Appendix.

The index of gerrymandering can be determined from the proportion of a district’s perimeter that is politically drawn. If $P = g + n$ denotes total district perimeter as the sum of mutable political boundary length g plus immutable natural boundary length n , then this proportion equals g/P , and the polygon complexity attributable to political manipulation is $G = (g/P)(P^2/A) = gP/A$.

Caveats

This index represents a significant logical and functional improvement over raw polygon complexity as a measure of gerrymandering, but it is still affected by natural boundary complexity. While politically-drawn boundary segments, like surveyed boundaries, have finite, well-defined lengths, some natural boundary segments such as shorelines do not (Mandelbrot, 1967). These features cannot be digitized without simplification, and their digitized lengths will depend upon the degree of simplification.

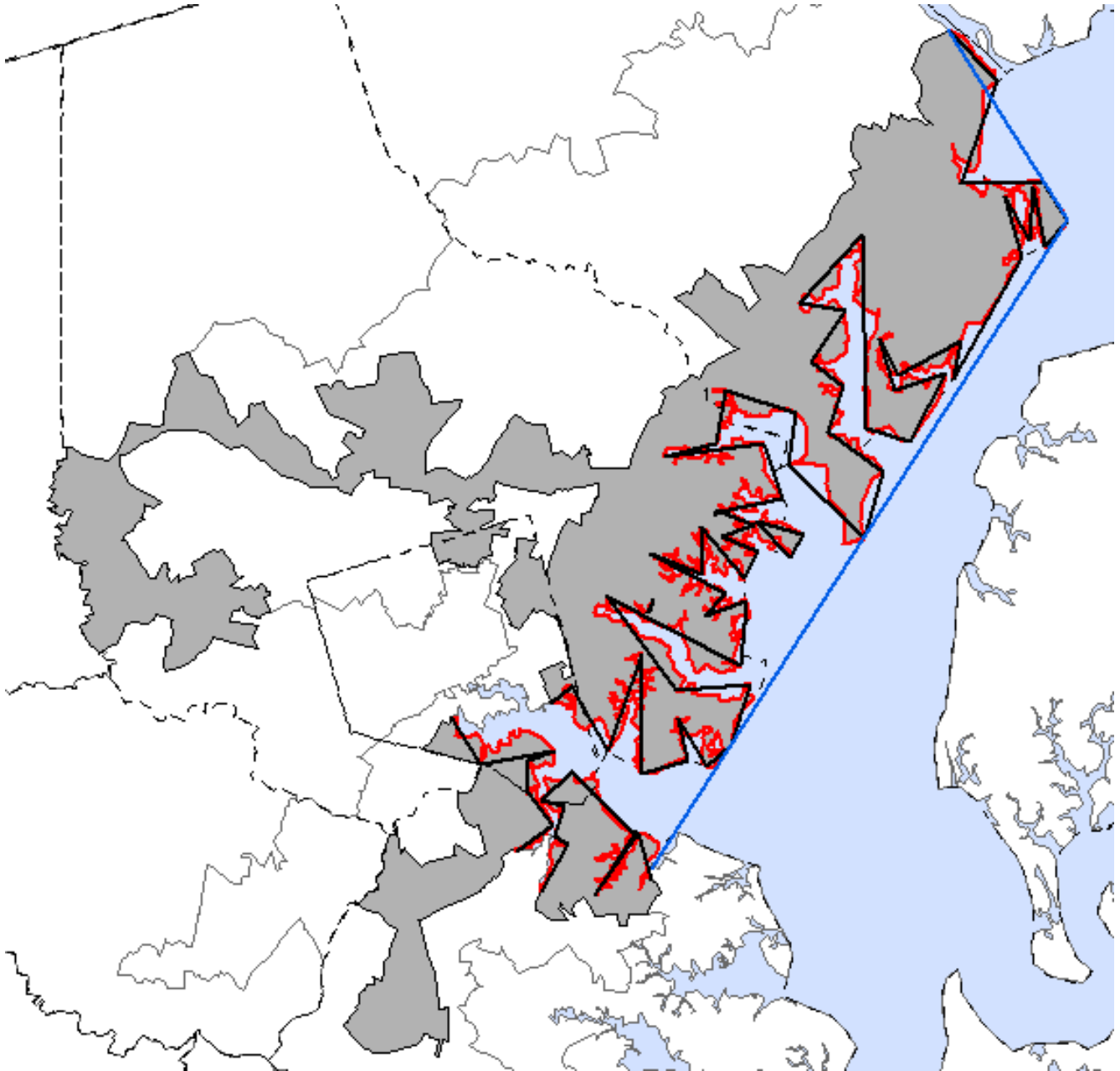


Figure 4: Shoreline simplifications of Maryland's 2nd Representative District

Figure 4 illustrates Maryland's 2nd district as something approaching a worst case. This district earned the top G score (410) as calculated from the 110th Congress shapefile obtained from the Census Bureau, and it includes a lot of complex Chesapeake Bay shoreline. The shoreline included in the perimeter for the original calculation is shown in red. I applied a fairly radical simplification, shown in

black, to the shoreline only, using a point shift tolerance of 2,000 meters. This reduced the shoreline length from 517,151 to 309,348 meters, the total perimeter from 927,962 to 720,159 meters, and the calculated G from 410 to 318, dropping this district's national gerrymander rank from first to fifth. The extreme simplification shown in blue, approximating the convex hull of the shoreline area, would reduce the shoreline length to 66,634 meters, the calculated G to 210 and the district's national gerrymander rank to eighth. The point here is that any complexity measure involving natural boundaries must necessarily have some arbitrariness.

Another criticism of this method is that it fails to account for other "natural" boundary segments of districts, such as county boundaries, rivers, etc. For example, two fairly convoluted inland segments of Maryland's 2nd district boundary coincide with rivers that are also county boundaries (shown as dashed lines) for Anne Arundel and Baltimore counties. Excluding these two segments from the computation of "artificial" perimeter would reduce the G index slightly. This further refinement is technically difficult if the digital county boundary polygons or river lines do not exactly match the corresponding district boundary segments.

Gerrymandering scores

The following maps show calculated gerrymandering indices for the last five Congresses (1999 through 2008) using the same data classifications and color schemes. Many of the changes between the 107th and 108th Congressional district boundaries were driven by reapportionment of seats following the 2000 Census. Some states, most notably Texas and Georgia, revised their district boundaries again after the 108th Congress. The mid-decade redistricting of Texas displaced at least five Democrat Congressmen with Republicans, although in some cases the new districts were actually more compact than the old districts.

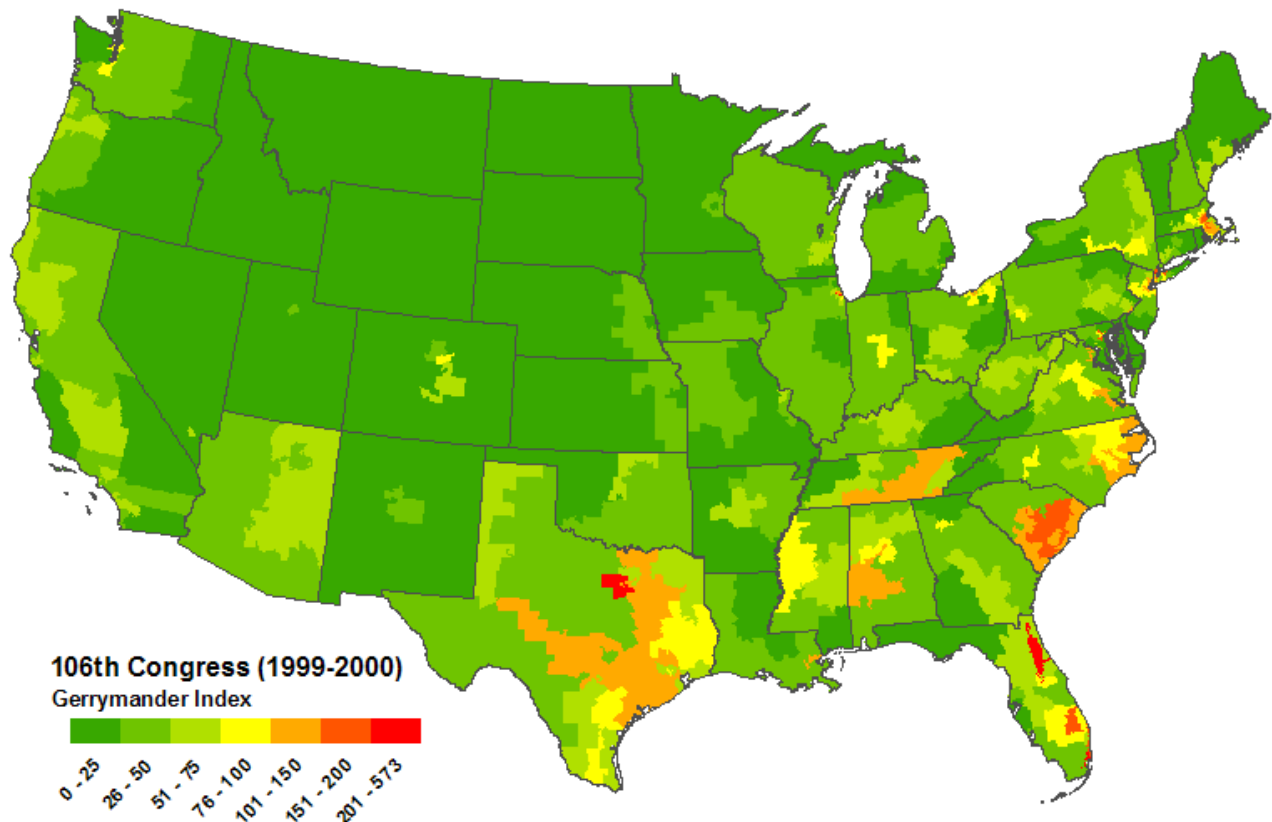


Figure 4a: Gerrymandering indices for 106th Congressional districts

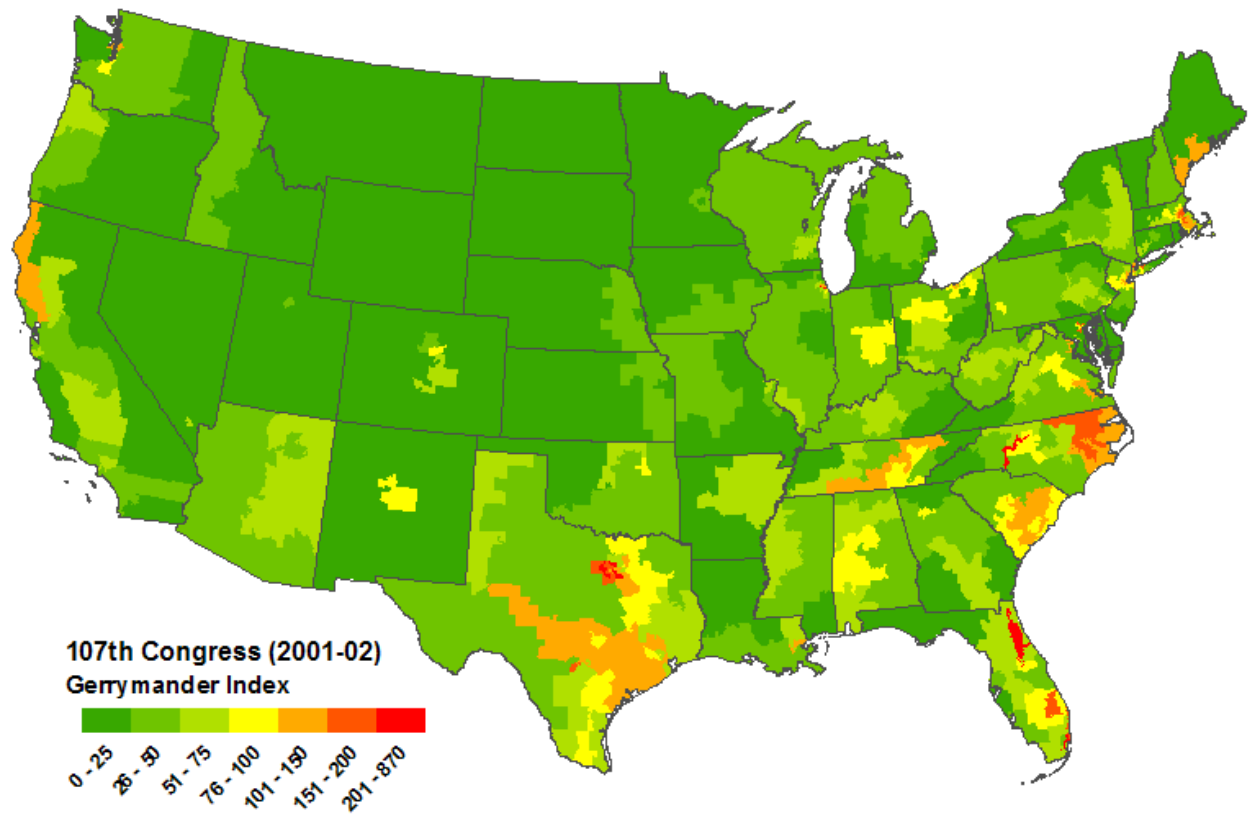


Figure 4b: Gerrymandering indices for 107th Congressional districts

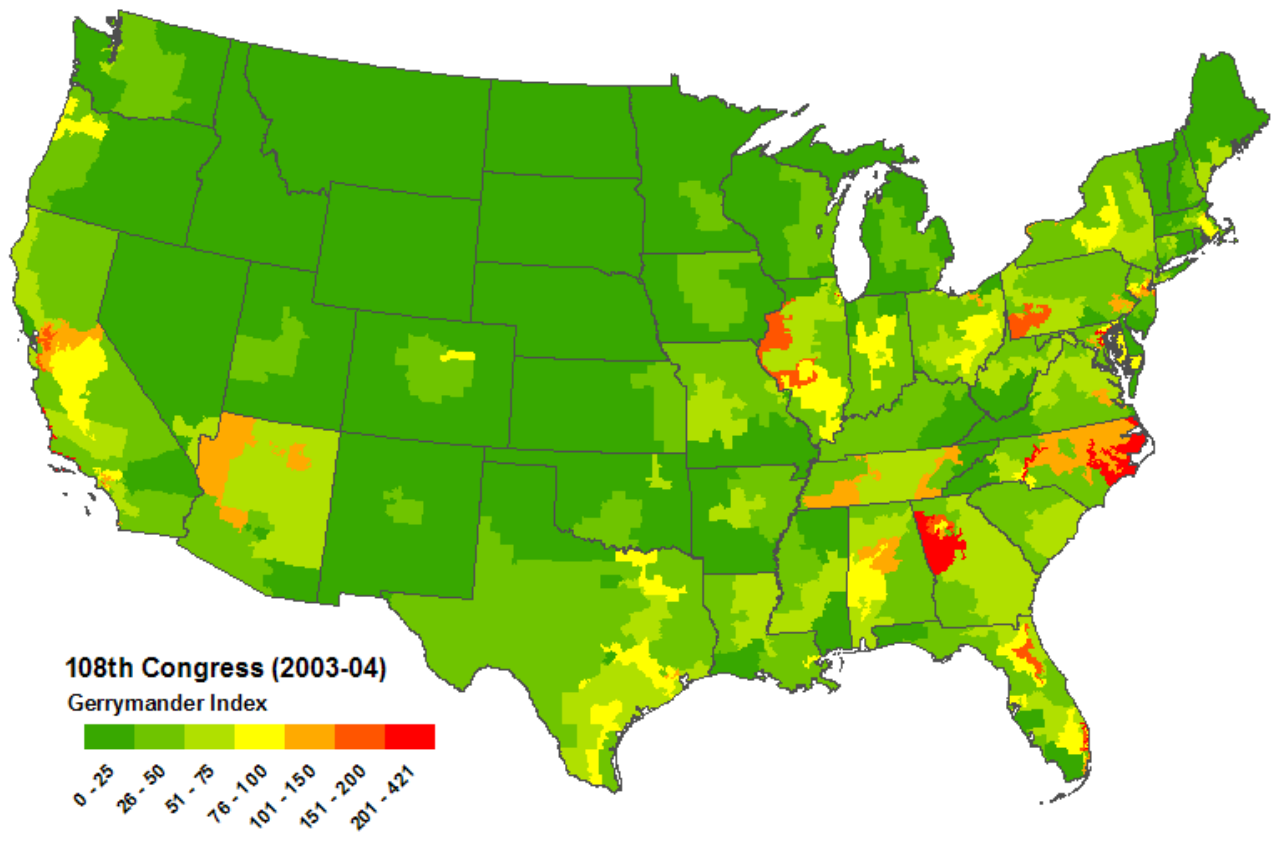


Figure 4c: Gerrymandering indices for 108th Congressional districts

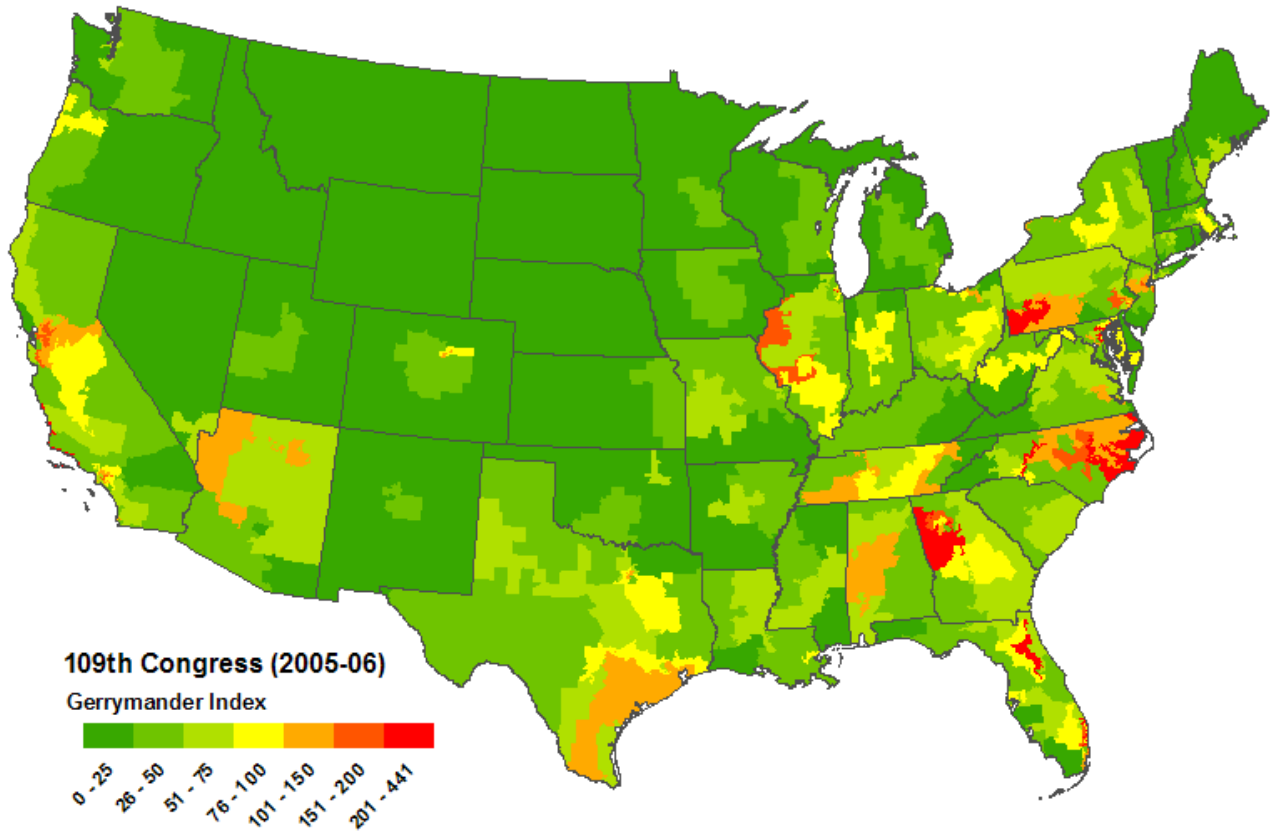


Figure 4d: Gerrymandering indices for 109th Congressional districts

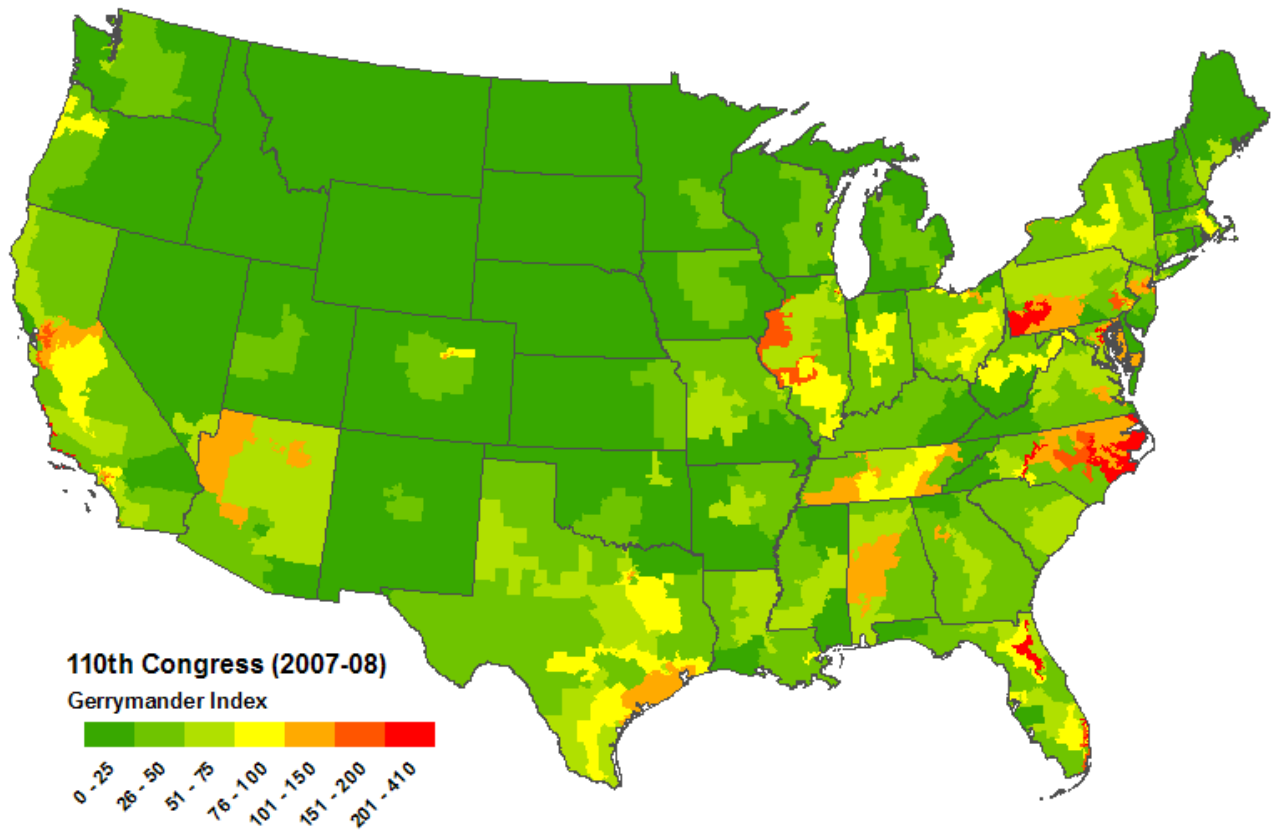
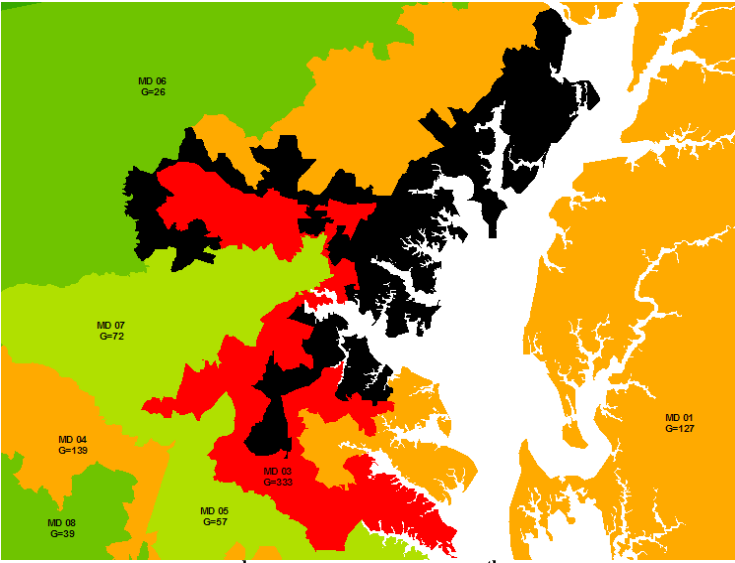


Figure 4e: Gerrymandering indices for 110th Congressional districts

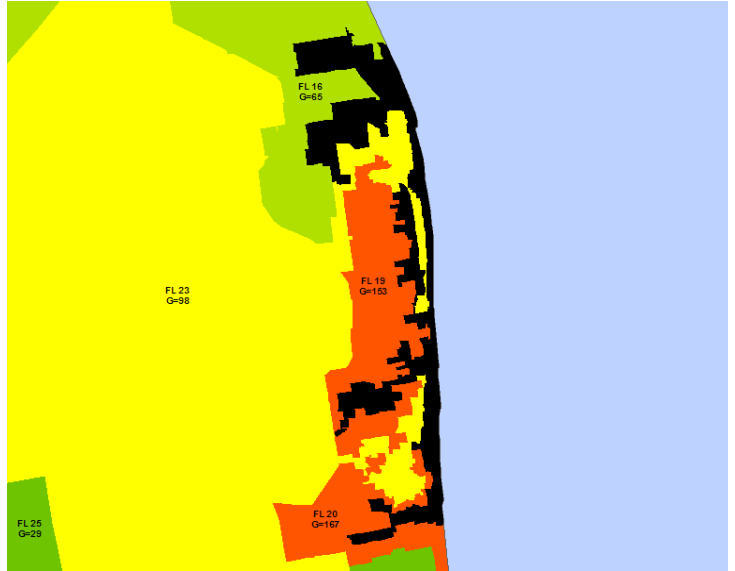
I ranked all 435 Representative districts for the 110th Congress according to their calculated G scores. The 40 most gerrymandered Representative districts of the 110th Congress, sorted by state and degree of gerrymandering ($G > 123$), are listed in Table 1 below. The eight most gerrymandered districts ($G > 200$) are shown in the six maps on the following page. Four of these districts include shoreline segments and three include state border segments that did not factor into their G scores. The 20 states with the highest average G scores are ranked in Table 2.

Table 1: Forty most gerrymandered Representative districts of the 110th Congress, by State

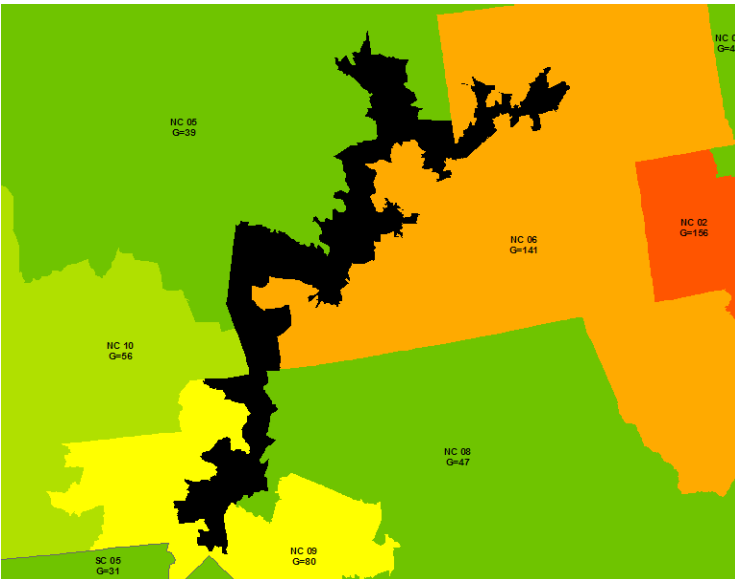
State	District	Area	Perimeter	InteriorLen	Gerrymandr	Rank
CA	23	2,710,712,479	1,117,039	567,987	234.06	8
CA	15	747,995,495	384,115	384,115	197.25	11
CA	39	169,806,474	175,878	175,878	182.17	12
CA	38	272,093,608	212,564	212,564	166.06	17
CA	10	2,809,534,141	664,561	664,561	157.19	18
CA	7	986,400,384	405,769	350,196	144.06	24
CA	11	5,998,547,487	901,539	901,539	135.50	28
CA	29	262,922,748	186,792	186,792	132.71	30
CA	16	601,100,884	273,302	273,302	124.26	38
CA	32	240,132,929	172,222	172,222	123.52	40
FL	22	741,857,852	587,570	489,804	387.94	2
FL	3	5,432,692,567	1,055,151	1,055,151	204.93	9
FL	18	1,217,803,298	875,636	245,657	176.63	14
FL	20	477,717,295	289,217	275,629	166.87	16
FL	19	606,344,475	304,239	304,239	152.66	20
IL	4	102,137,314	185,259	185,259	336.03	4
IL	17	21,468,882,487	2,157,351	1,678,058	168.62	15
MA	8	125,902,426	132,821	122,199	128.91	31
MD	2	929,983,832	927,962	410,812	409.92	1
MD	3	761,306,489	576,950	439,278	332.90	5
MD	4	822,694,728	352,968	325,052	139.46	26
MD	1	9,602,496,943	4,208,660	288,730	126.55	34
NC	12	2,142,809,517	875,449	875,449	357.67	3
NC	3	18,222,836,794	2,836,807	1,640,672	255.41	6
NC	2	10,306,494,675	1,267,739	1,267,739	155.94	19
NC	6	7,739,873,689	1,044,548	1,044,548	140.97	25
NC	1	19,347,573,709	1,786,873	1,499,524	138.49	27
NC	13	5,939,165,507	938,032	798,738	126.15	35
NJ	6	534,664,360	349,879	273,620	179.05	13
NJ	13	173,562,346	178,660	131,164	135.02	29
NJ	10	179,822,486	149,556	149,556	124.38	37
NJ	12	1,664,121,446	481,796	427,679	123.82	39
NY	8	49,676,726	98,497	72,994	144.73	23
NY	12	51,333,141	80,317	80,317	125.67	36
PA	12	7,206,950,016	1,347,178	1,263,574	236.20	7
PA	18	3,720,284,196	892,574	834,877	200.30	10
PA	6	2,121,308,811	568,459	568,459	152.33	21
PA	1	175,387,064	180,737	144,993	149.42	22
PA	14	439,464,746	237,095	237,095	127.91	32
TN	7	16,443,475,459	1,563,860	1,331,073	126.59	33



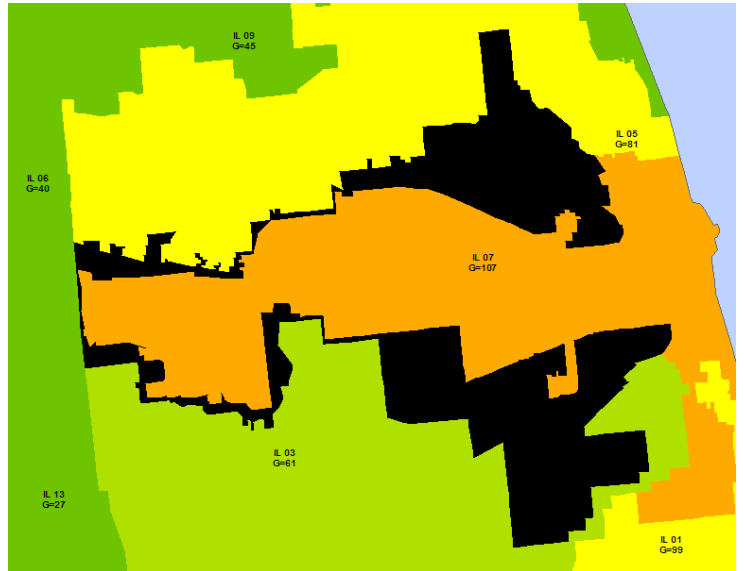
Maryland 2nd (black) G=410 and 4th (red) G=333



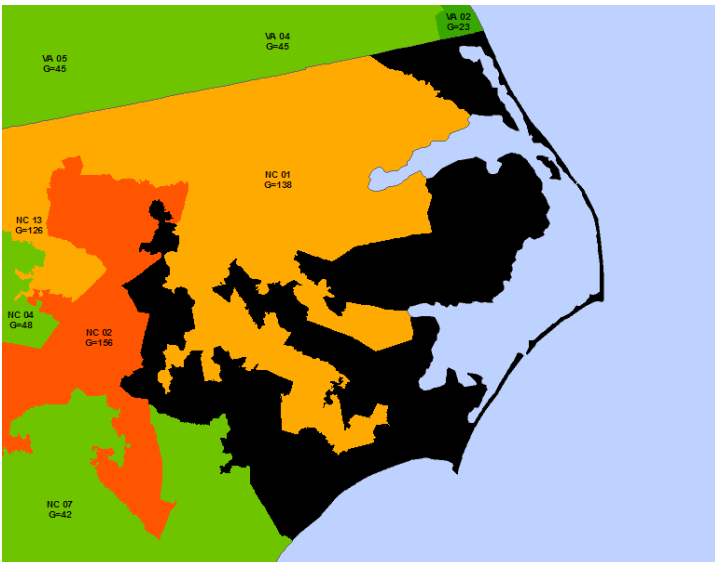
Florida 22nd, G=388



North Carolina 12th, G=358



Illinois 4th, G=336



North Carolina 3rd, G=255



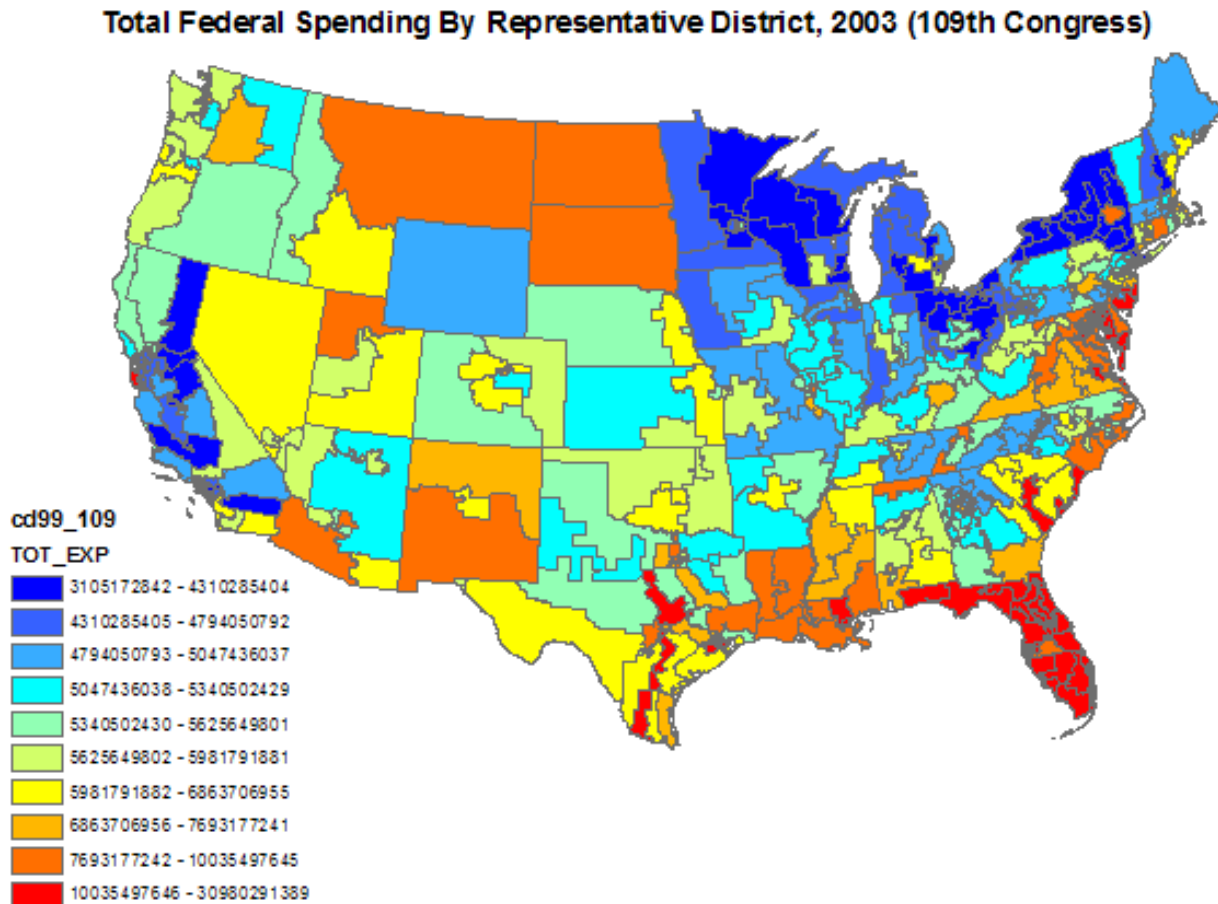
Pennsylvania 12th (black) G=236 and 18th (red) G=200

Table 2: States with Highest Average Gerrymandering Scores

State	Seats	Average G
MD	8	150.3
NC	13	115.5
FL	25	90.4
PA	19	89.1
CA	53	80.6
NJ	13	77.6
IL	19	76.6
TX	32	68.6
AL	7	64.8
TN	9	62.9
MA	10	62.0
VA	11	55.7
NY	29	54.9
OH	18	51.0
LA	7	47.8
CO	7	47.2
WV	3	45.5
GA	13	44.5
AZ	8	44.4
SC	6	43.0

Consequences of Gerrymandering

To obtain a crude measure of legislator “effectiveness,” I aggregated 2003 Congressional appropriations data from the US Census Bureau’s *Consolidated Federal Funds Report* by Representative district. Where appropriations were for multiple districts within a state, I apportioned the dollars equally among the districts. Aggregate federal spending by district is mapped below.



These spending figures include Social Security payments, which is why Florida stands out in this map.

The final two maps show Representative seniority (years in Congress) and the ratio of federal expenditures received versus total federal taxes paid in 2004, by Congressional district. Fortunately, Congressman are mortal, so even the safest seats see turnover (*pace* John Murtha).

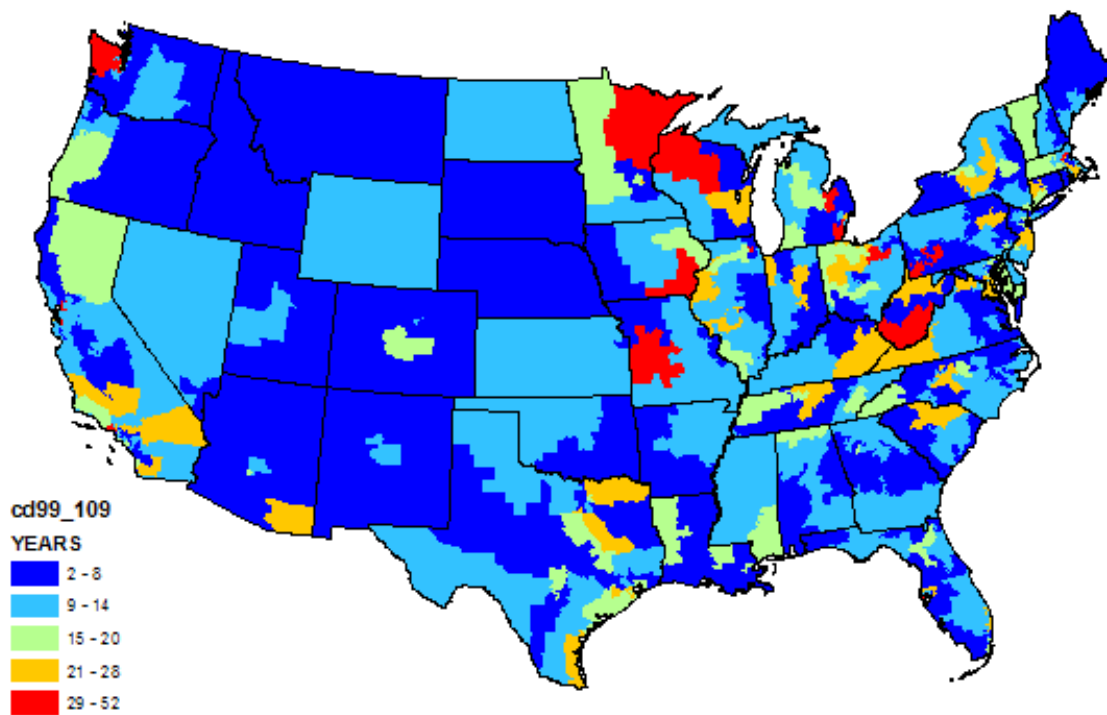
The tax data obtained from the Tax Foundation (www.taxfoundation.org) were calculated from IRS reports. In this map, the “winner” and “loser” districts in the federal money redistribution game are clearly identifiable.

The final image below is a scatterplot of the logarithms of districts’ compactness indices versus the logarithms of federal dollars received per dollar of federal taxes paid. As this scatterplot indicates, the data do not show any significant correlation between my informal gerrymandering index and the federal expenditure/tax ratios of Congressional districts. A trial regression of the expenditure/tax ratio against both district compactness and Representative seniority actually suggested that more *junior* members of Congress from more *compact* districts may fare slightly better in delivering federal dollars to their home districts, but this model explained less than four percent of the total variation in federal appropriations.

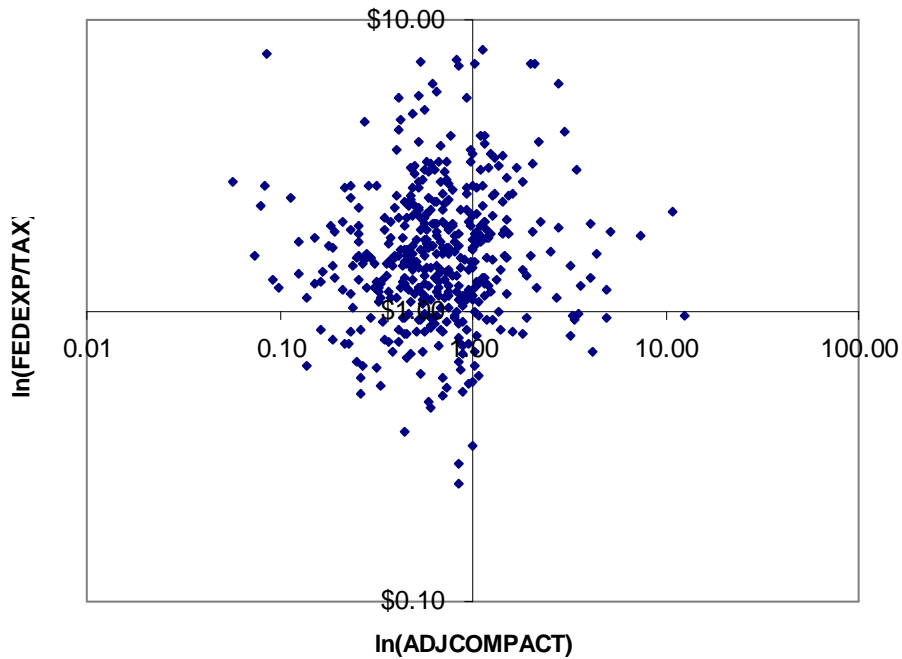
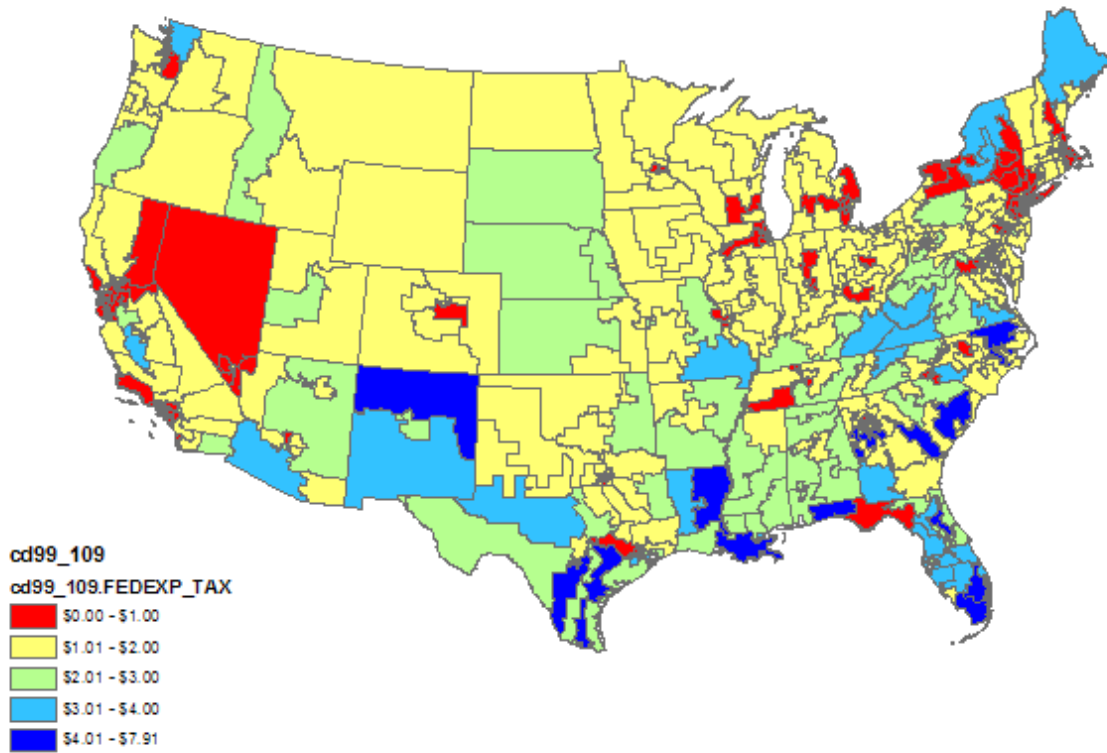
These results are generally consistent with prior studies that show little or no relationship between legislative seniority, campaign fundraising, or victory margins in elections and federal appropriations. While gerrymandering may be decried as dirty politics, my analysis does not suggest that it reduces the effectiveness of US Representatives in capturing (or recapturing) federal dollars for their home districts. Indeed, given the complex political interdependencies between US Representatives and local and state-level officials, other members of their state's Congressional delegation from both parties, their various committee colleagues, and their caucuses, any individualistic modeling of the appropriations process may be naïve.

In a democracy, the voters get what they deserve. Legislators write their own rules, and incumbency gives them many advantages over prospective challengers. As political sport, redrawing district boundaries may distract a lot of politicians from more serious malfeasance.

Congressional Seniority, 2003 (109th Congress)



Federal Spending per Federal Tax Dollar Paid,
by Representative District, 2003 (109th Congress)



Federal Dollars Received per Federal Tax Dollar Paid vs. Geometric Compactness Index of
US Representative Districts, 109th Congress

Conclusions

The Supreme Court has generally been reluctant to intervene in state legislatures' redistricting processes, and states have followed various principles and procedures in re-drawing their political boundaries. While many public interest groups have called for state legislatures to delegate redistricting to independent commissions, only Arizona has actually done so. Other states have advisory commissions, but legislators are typically loath to surrender real redistricting authority. FairVote's websites (<http://www.fairvote.org/gerrymandering/> and <http://archive.fairvote.org/index.php?page=1389>) provide detailed information on redistricting methods and reform initiatives.

While district compactness is widely viewed as desirable, this paper indicates that any geometric compactness measure will likely have a significant degree of arbitrariness. Laypeople may be amused or appalled by highly contorted districts, but some contortions may be justified by equity concerns. The Supreme Court has articulated the principle that districts should represent "communities of interest," and such communities may not be geographically compact. The bizarre shape of Illinois' 4th District connects Chicago's Hispanic communities to give them a voice in Congress. The stringy shapes of coastal districts in California and Florida are partly due to natural concentrations of population along coasts.

One generally agreeable redistricting criterion that is receiving increasing attention is to have district boundaries follow county or Census tract boundaries wherever possible. Census geography is generally reflective of "communities of interest." Congressional district boundaries deserve increasing scrutiny when they cut through smaller-order Census geographies—tracts, block groups and blocks.

References:

Gillman, R. 2002. "Geometry and Gerrymandering", *Math Horizons* 10(1):10-13.

Humphreys, M. 2009. "Can compactness constrain the Gerrymander?"
<http://www.columbia.edu/~mh2245/papers1/gerry.pdf>

Mandelbrot, B. 1967. How long is the coastline of Britain? Statistical self-similarity and fractal dimension. *Science* 156:636-638.

Niemi, R.G., B. Groffman, C Calucci and T. Hofeller 1990. "Measuring compactness and the role of compactness standard in a test for partisan and racial gerrymandering" *Journal of Politics* 52:1155-1181.

Wildgen, J.K. 1995. "Fractal geometry and the boundaries of voting districts" *Spatial and Contextual Models in Political Research*, 1995.

Appendix--GIS procedures:

The GIS analysis was executed with Arc 9.3 (ESRI). I applied a US National Atlas Equal-Area projection to the lat-lon Congressional district boundary shapefile obtained from the Census Bureau. After adding AREA and PERIMETER fields in the attribute table, I used the Calculate Geometry tool to calculate the area (in square meters) and perimeter (in meters) of each district.

I used the Dissolve utility to combine Congressional districts by state ID to obtain a shapefile of exactly congruent state boundaries.

I used the Polygon-to-Line utility to crack the 435 district polygons into 2,232 component boundary segments. These line features include attribute table fields LEFT_FID and RIGHT_FID that identify the polygon feature ID's on the "left" and "right" sides of each line.

I used the Select-by-Location "are within (Clementini)" criterion to select the interior district boundaries, identifying 1,318 district boundary line features as not coincident with any state boundary. I created a LENGTH field in this line shapefile's attribute table, and used the Calculate Geometry tool to calculate the lengths of these interior boundary segments. The unselected segments on state boundaries were assigned LENGTH values of zero.

The next step was create Summarize tables to sum the LENGTH values for each LEFT_FID and RIGHT_FID. I then Joined these two Summarize tables to the original district polygon shapefile, created a new INTLENGTH field in the polygon attribute table, and summed the LENGTH values from the two Summarize tables into that field.

The calculation of G values was straightforward: $G = \text{INTLENGTH} * \text{PERIMETER} / \text{AREA}$.