

A Measure of Bizarreness

Christopher P. Chambers* and Alan D. Miller^{†,‡}

ABSTRACT

We introduce a path-based measure of convexity to be used in assessing the compactness of legislative districts. Our measure is the probability that a district contains the shortest path between a randomly selected pair of its points. The measure is defined relative to exogenous political boundaries and population distributions.

The upcoming decennial census will result in a new legislative redistricting process to be completed in 2012. That year will also mark the two-hundredth anniversary of the *Gerrymander* — that monster of American politics — the bizarrely shaped legislative district drawn as a means to certain electoral ends.

An early diagnosis of this malady did not lead to an early cure. Already in the nineteenth century, reformers introduced anti-gerrymandering laws requiring districts to be

* Division of the Humanities and Social Sciences, Mail Code 228-77, California Institute of Technology, Pasadena, CA 91125, USA. Email: chambers@caltech.edu. Web: <http://www.hss.caltech.edu/~chambers/>

† Faculty of Law and Department of Economics, University of Haifa, Mount Carmel, Haifa, 31905, Israel. Email: admiller@econ.haifa.ac.il. Web: <http://econ.haifa.ac.il/~admiller/>

‡ We would like to thank the editors and an anonymous referee for their helpful comments. In addition, we would like to thank Yaser Abu-Mostafa, Micah Altman, Bruce Bueno de Mesquita, Federico Echenique, Paul Edelman, Timothy Feddersen, Itzhak Gilboa, Daniel Goroff, Tim Groseclose, Catherine Hafer, Paul Healy, Matt Jackson, Jonathan Katz, Ehud Lehrer, R. Preston McAfee, Stephen Miller, Richard Pildes, Daniel Polsby, Robert Popper, Dinakar Ramakrishnan, Eran Shmaya, Alastair Smith, Matthew Spitzer, Daniel Ullman, Peyton Young, and seminar participants at Caltech, the 66th Annual Meeting of the Midwest Political Science Association, the First Annual Graduate Student Conference at the Alexander Hamilton Center for Political Economy at New York University, and the 2009 Joint Mathematics Meetings. Christopher Chambers would like to thank the support of the NSF through grant SES-0751980.

“compact” and “contiguous”,¹ but the disease spread unabated. District shapes grew more odd over time as politicians used modern technology to increase their control over elections. In 1812 a district was said to resemble a salamander; one hundred eighty years later, another was likened to a “Rorschach ink blot test.”²

Redistricting reform has been hampered by a lack of agreement among experts as to what a good districting plan should look like. Some believe that legislatures should mirror the racial, ethnic, or political balance of the population. Others believe that it is more important that districts be competitive or, alternatively, stable. This lack of an ideal has made it difficult to design an algorithm which yields districting plans acceptable to all.

Rather than make districts better by moving them closer to an ideal, we try to make districts less bad by moving them further from an identifiable problem. That problem is bizarre shape. We introduce a new method to measure the bizarreness of a legislative district. The method provides courts with an objective means to identify the more egregious gerrymanders which weaken the citizens’ confidence in the electoral system.

As with so many other aspects of redistricting, there is little agreement as to reason for restricting bizarre shapes. Some argue that while the shape of legislative districts is not important in and of itself, compactness restrictions constrain the set of choices available to gerrymanderers and thereby limit their ability to control electoral outcomes. Others believe that bizarrely shaped districts cause direct harm in the “pernicious” messages that they send to voters and their elected representatives.³

Laws restricting the shapes of legislative districts have been unsuccessful, in part because courts lack established criteria to determine whether a particular shape is allowable. Lawyers, political scientists, geographers, and economists have introduced multiple methods to measure district compactness.⁴ However, none of these methods is widely accepted, in part because of problems identified by Young (1988), Niemi *et al.* (1990), and Altman (1998).

Part of the difficulty in defining a measure of compactness is that there are many conflicting understandings of the concept. According to one view the compactness standard exists to eliminate elongated districts. In this sense a square is more compact than a rectangle, and a circle may be more compact than a square. According to another

¹ Thirty-five states require congressional or legislative districting plans to be compact, forty-five require contiguity, and only Arkansas requires neither. See NCSL (2000). There may also be federal constitutional implications. See *Shaw v. Reno*, 509 U.S. 630 (1993); *Bush v. Vera*, 517 U.S. 959 (1996).

² *Shaw v. Reno*, 509 U.S. at 633.

³ “Put differently, we believe that reapportionment is one area in which appearances do matter.” *Shaw v. Reno*, 509 U.S. at 647. The direct harm that arises from the ugly shape of the legislative districts is generally referred to as an “expressive harm.” See Pildes and Niemi (1993).

⁴ “Contiguity” is generally understood to require that it be possible to move between any two places within the district without leaving the district. See for example Black’s Law Dictionary which defines a “contiguous” as touching along a surface or a point (Garner, 2004).

view compactness exists to eliminate oddly shaped districts.⁵ According to this view a rectangle-shaped district is better than a district shaped like a Rorschach blot.

We follow the latter approach. While it may be preferable to avoid elongated districts, the sign of a heavily gerrymandered district is bizarre shape. To the extent that elongation is a concern, it should be studied with a separate measure.⁶ These are two separate issues, and there is no natural way to weigh tradeoffs between bizarreness and elongation.

We note that, in some cases, bizarrely shaped districts may be justified by compliance with the Voting Rights Act of 1965.⁷ It is not clear whether any of these bizarre shapes could have been avoided by districting plans which satisfy the constraints of the act.⁸ Whether a bizarrely shaped district is necessary to satisfy civil rights law is a matter for the courts.⁹ Our role is only to provide a meaningful standard by which the court can determine whether districts are bizarrely shaped.

The basic principle of *convexity* requires a district to contain the shortest path between every pair of its points. Circles, squares, and triangles are examples of convex shapes, while hooks, stars, and hourglasses are not. (See Figure 1.) The most striking feature of bizarrely shaped districts is that they are extremely non-convex. (See Figure 2.) We introduce a measure of convexity with which to assess the bizarreness of the district.

The *path-based* measure we introduce is the probability that a district contains the shortest path between a randomly selected pair of its points.¹⁰ This measure always returns a number between zero and one, with one being perfectly convex. To understand how our measure works, consider a district containing two equally sized towns connected by a very narrow path, such as a road. (See Figure 3(a).) Our method assigns this district a measure of approximately one-half. A district containing n equally sized towns connected by narrow paths is assigned a measure of approximately $1/n$.¹¹ (See Figure 3(b).) If the

⁵ Writing for the majority in *Bush v. Vera*, Justice O'Connor referred to "bizarre shape and noncompactness" in a manner which suggests that the two are synonymous, or at least very closely related. If so then a compact district is one without a bizarre shape, and a measure of compactness is a measure of bizarreness.

⁶ Elongated districts are not always undesirable. See Figure 5.

⁷ See 42 U.S.C. 1973c.

⁸ Individuals involved in the redistricting process often attempt to satisfy multiple objectives when creating redistricting plans. It may be the case that the bizarreness of these districts could be reduced by sacrificing other objectives (such as creating safe seats for particular legislators) without hurting the electoral power of minority groups. As a matter of law, it is not clear that the Voting Rights Act necessarily requires bizarre shapes in any case.

⁹ The Supreme Court has held that, irrespective of the Voting Rights Act, "redistricting legislation that is so bizarre on its face that it is 'unexplainable on grounds other than race'" is subject to a high level of judicial scrutiny. *Shaw v. Reno*, 509 U.S. at 643. See also Pildes and Niemi (1993).

¹⁰ Versions of this measure were independently discovered by Lehrer (2007) and Žunić and Rosin (2002). These works do not discuss population weighting or exogenous boundaries.

¹¹ Alternatively one might use the reciprocal, where the measure represents the equivalent number of disparate communities strung together to form the district. The reciprocal is always a number greater or equal to one, where one is perfectly convex. A district containing n towns connected by narrow paths is assigned a measure of approximately n .

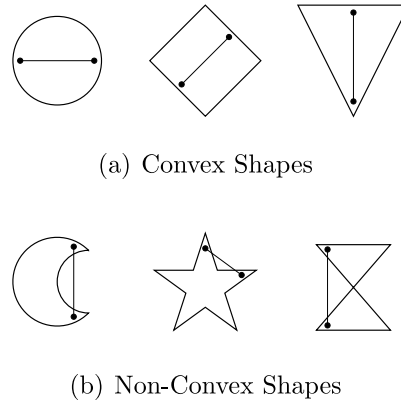


Figure 1. Convexity.

n towns are not equally sized, the measure is equivalent to the Herfindahl–Hirschman Index (Hirschman, 1964).¹²

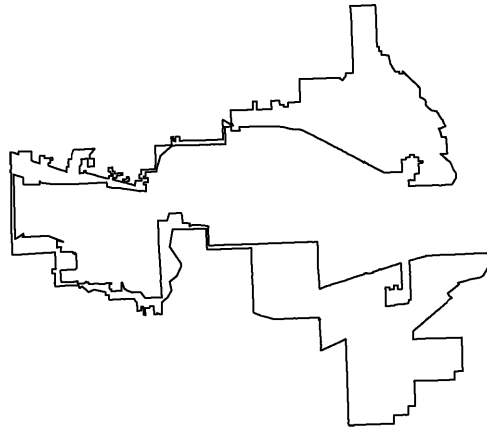
Ideally, a measure of compactness should consider the distribution of the population in the district. For example, consider the two arch-shaped districts depicted in Figure 4. The districts are of identical shape, thus the probability that each district contains the shortest path between a randomly selected pair of its points is the same. However, the populations of these districts are distributed rather differently. The population of district A is concentrated near the bottom of the arch, while that of district B is concentrated near the top. The former district might represent two communities connected by a large forest, while the latter district might represent one community with two forests attached.

Population can be incorporated by using the probability that a district contains the shortest path between a randomly selected pair of its residents. In practice our information is more limited — we do not know the exact location of every resident, but only the populations of individual census blocks. We can solve this problem by weighting points by population density. The population-weighted measure of district A is approximately one-half, while that of district B is nearly one.¹³

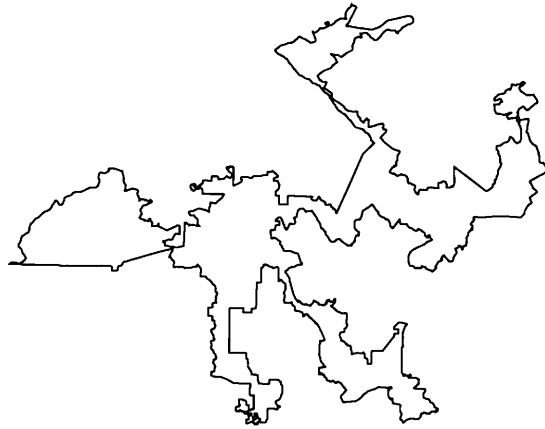
One potential problem is that some districts may be oddly shaped simply because the states in which they are contained are non-convex. Consider, for example, Maryland’s Sixth Congressional District (shown in Figure 5 in gray). Viewed in isolation, this district is very non-convex — the western portion of the district is almost entirely disconnected from the eastern part. However, the odd shape of the district is a result of the state’s boundaries, which are fixed. We solve this problem by measuring the probability that

¹² If x_i is the size of town i , then the measure of the district is $\sum_{i=1}^n x_i^2 \left[\sum_{j=1}^n x_j \right]^{-2}$.

¹³ Note that, under the population-weighted approach, a district may have a perfect score even though it has oddly shaped boundaries in unpopulated regions. The ability to draw bizarre boundaries in unpopulated regions is of no help to potential gerrymanderers.



(a) 4th District, Illinois

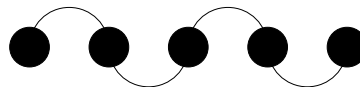


(b) 13th District, Georgia

Figure 2. Congressional Districts, 109th Congress.



(a) Two Circular Towns



(b) Five Circular Towns

Figure 3. Towns connected with narrow paths.



Figure 4. Same shapes, different populations.



Figure 5. Sixth District, Maryland, 109th Congress.

a district contains the shortest path *in the state* between a randomly selected pair of its points. The adjusted measure of Maryland's Sixth Congressional District is close to one.

Our measure considers *whether* the shortest path in a district exceeds the shortest path in the state. Alternatively, one might wish to consider the *extent* to which the former exceeds the latter. We introduce a parametric family of measures which vary according to the degree that they penalize deviations from convexity. At one extreme is the measure we have described; at the other is the degenerate measure, which gives all districts a measure of one regardless of their shape.

Related Literature

Individual District Compactness Measures

A variety of compactness measures have been introduced by lawyers, social scientists, and geographers. Here we highlight some of basic types of measures and discuss some of their weaknesses. A more complete guide may be found in surveys by Young (1988), Niemi *et al.* (1990), and Altman (1998).

Most measures of compactness fall into two broad categories: (1) dispersion measures and (2) perimeter-based measures. Dispersion measures gauge the extent to which the district is scattered over a large area. The simplest dispersion measure is the length-to-width test, which compares the ratio of a district's length to its width. Ratios closer to

one are considered more compact. This test has some support in the literature, most notably from Harris (1964).¹⁴

Another type of dispersion measure compares the area of the district to that of an ideal figure. This measure was introduced into the redistricting literature by Reock (1961), who proposed using the ratio of the area of the district to that of the smallest circumscribing circle. A third type of dispersion measure involves the relationship between the district and its center of gravity. Measures in this class were introduced by Boyce and Clark (1964) and Kaiser (1966). The area-comparison and center of gravity measures have been adjusted to take account of district population by Hofeller and Grofman (1990) and Weaver and Hess (1963), respectively.

Dispersion measures are widely criticized, in part because they consider districts reasonably compact as long as they are concentrated in a well-shaped area. (See Young, 1988.) We point out a different (although related) problem. Consider two disjoint communities strung together with a narrow path. *Disconnection-sensitivity* requires the measure to consider the combined region less compact than at least one of the original communities. None of the dispersion measures are disconnection-sensitive. An example is shown in Figure 6.¹⁵

Perimeter measures use the length of the district boundaries to assess compactness. The most common perimeter measure, associated with Schwartzberg (1966), involves comparing the perimeter of a district to its area.¹⁶ Young (1988) objected to the Schwartzberg measure on the grounds that it is overly sensitive to small changes in the boundary of a district. Jagged edges caused by the arrangement of census blocks may lead to significant distortions. While a perfectly square district receives a score of 0.785, a square shape superimposed upon a diagonal grid of city blocks has a much longer perimeter and a lower score, as shown in Figure 7(a).¹⁷ Figure 7 shows four shapes, arranged according to the Schwartzberg ordering from least to most compact.

Taylor (1973) introduced a measure of indentation which compared the number of reflexive (inward-bending) to non-reflexive (outward-bending) angles in the boundary

¹⁴ The length-to-width test seems to have originated in early court decisions construing compactness statutes. See *In re Timmerman*, 100 N.Y.S. 57 (N.Y. Sup. 1906).

¹⁵ The length-width measure is the ratio of width to length of the circumscribing rectangle with minimum perimeter. See Niemi *et al.* (1990). All measures are transformed so that they range between zero and one, with one being most compact. The Boyce-Clark measure is $\sqrt{1/(1+bc)}$, where bc is the original Boyce-Clark measure (Boyce and Clark, 1964). The Schwartzberg measure used is the variant proposed by Polsby and Popper (1991) (originally introduced in a different context by Cox, 1927), or $(1/sc)^2$, where sc is the measure used by Schwartzberg (1966).

¹⁶ This idea was first introduced by Cox (1927) in the context of measuring roundness of sand grains. The idea first seems to have been mentioned in the context of district plans by Weaver and Hess (1963) who used it to justify their view that a circle is the most compact shape. Polsby and Popper (1991) also supported the use of this measure.

¹⁷ The score of the resulting district decreases as the city blocks become smaller, reaching 0.393 in the limit.



COMPACTNESS MEASURES		
	<i>District:</i>	
	I	II
<i>Dispersion Measures</i>		
Length-Width	0.63	1.00
Reock	0.32	0.44
Area to Convex Hull	0.57	0.70
Boyce-Clark	0.15	0.29
<i>Other Measures</i>		
Path-Based Measure	0.84	0.42
Schwartzberg	0.29	0.14
Taylor	0.40	0.20

Figure 6. District II is formed by connecting district I to a copy of itself. Disconnection-sensitivity implies that I is more compact.

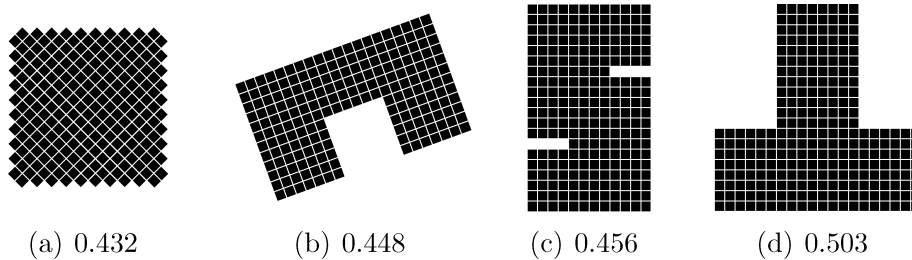


Figure 7. Schwartzberg measure.

of the district. Taylor's measure is similar to ours in that it is a measure of convexity. Figure 8 shows six districts and their Taylor measures, arranged from best to worst.

Lastly, Schneider (1975) introduced a measure of convexity using Minkowski addition.¹⁸ For more on the relationship between convex bodies and Minkowski addition, see Schneider (1993).

¹⁸ Schneider's measure is closely related to an earlier measure of convexity introduced by Arrow and Hahn (1971).

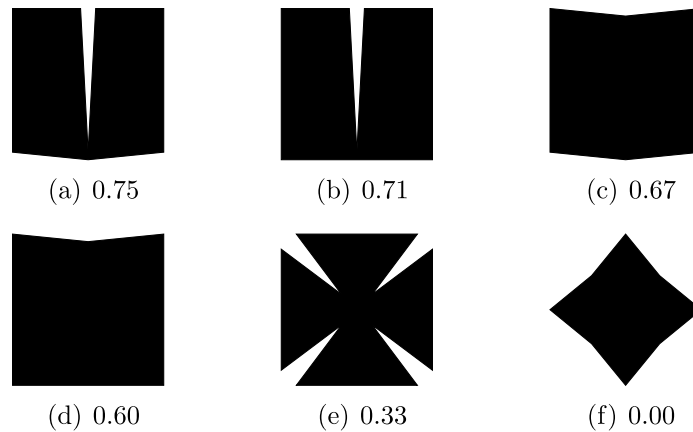


Figure 8. Taylor's measure.

Districting-Plan Compactness Measures

In addition to these measures of individual legislative districts, several proposals have been introduced to measure entire districting plans. The “sum-of-the-perimeters” measure, found in the Colorado Constitution, is the “aggregate linear distance of all district boundaries.”¹⁹ Smaller numbers indicate greater compactness. An alternative method was introduced by Papayanopoulos (1973). His proposal can be described through a two-stage process. First, in each district, the sum total of the distances between each pair of residents is calculated. The measure for the plan is then the sum of these scores across the districts. Smaller numbers again indicate greater compactness. More recently, Fryer and Holden (2007) proposed a related measure which uses quadratic distance and which is normalized so that an optimally compact districting plan has a score of one.

A potential problem, raised by Young (1988), is that these measures penalize deviations in sparsely populated rural areas much more severely than deviations in heavily populated urban areas. For example, Figure 9 shows five potential districting plans for a four-district state with sixteen equally sized population centers (represented by dots). The upper portion of the state represents an urban area with half of the population concentrated into one-seventeenth of the land. Papayanopoulos scores are given, although we note that the sum-of-the-perimeters and Fryer–Holden measures give identical ordinal rankings of these districting plans.

According to these measures, the ideal districting plan divides the state into four squares (Figure 9(a)). The plan with triangular districts is less compact (Figure 9(b)), and the plan with wave-shaped districts fares the worst (Figure 9(c)). However, the measure is more sensitive to deviations in areas with lower population density. The

¹⁹ Colo. Const. Art. V, Section 47.

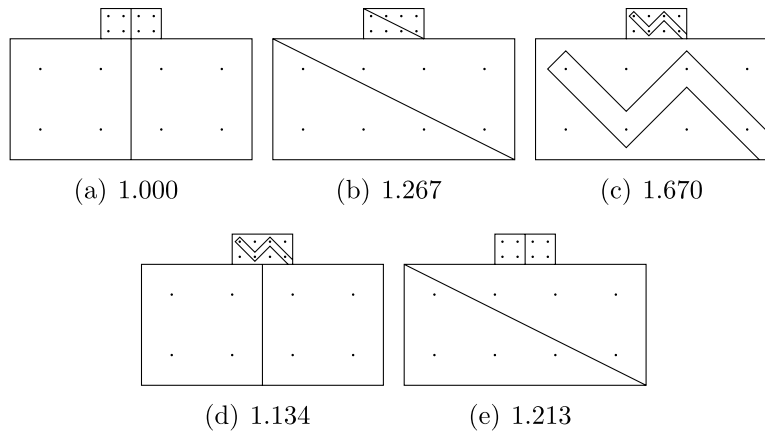


Figure 9. Urban gerrymandering.

plan in Figure 9(d), which divides the rural area into perfect squares and the urban area into low-scoring wave-shape districts, is considered more compact than the plan in Figure 9(e), which divides the rural area into triangles and the urban area into perfect squares.

An alternative approach is to rank state-wide districting plans using the scores assigned to individual districts. Examples include the *utilitarian* criterion, which is the average of the districts' scores (see Papayanopoulos, 1973), and the *maxmin* criterion, which is simply the lowest of the scores awarded the districts under the plan. This approach allows for the ranking of both individual districts and entire districting plans as required by Young (1988).

The ideal criterion depends in large part on the individual district measure with which it is used. We advocate the use of the maxmin criterion with our path-based measure on the grounds that it restricts gerrymandering the most. The maxmin criterion is also consistent with the U.S. Supreme Court's focus on analyzing individual districts as opposed to entire districting plans.²⁰ However, if some districts must necessarily be noncompact (a common problem with the Schwartzberg measure) then the utilitarian criterion may be more appropriate.

Other literature

Vickrey (1961) shows that restrictions on the shape of legislative districts are not necessarily sufficient to prevent gerrymandering. In Vickrey's example there is a rectangular

²⁰ This focus might stem from the Court's understanding of the right to vote as an individual right, and not a group or systemic right. This understanding may have influenced other measures used in the redistricting context, such as the "total deviation" test. See Edelman (2006).

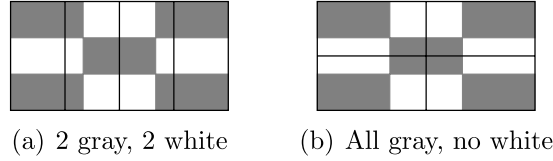


Figure 10. Vickrey's example.

state in which support for the two parties (white and gray) are distributed as shown in Figure 10. With one district plan, the four legislative seats are divided equally; with the other district plan, the gray party takes all four seats. In both plans, the districts have the same size and shape.

Compactness measures have been touted both as a tool for courts to use in determining whether districting plans are legal and as a metric for researchers to use in studying the extent to which districts have been gerrymandered. Other methods exist to study the effect of gerrymandering — the most prominent of these is the seats–votes curve, which is used to estimate the extent to which the district plan favors a particular party as well as the responsiveness of the electoral system to changes in popular opinion. For more see Tufte (1973).

THE MODEL AND PROPOSED FAMILY OF MEASURES

The Model and Notation

Let \mathcal{K} be the collection of compact sets in \mathbb{R}^n whose interiors are path-connected (with the usual Euclidean topology) and which are the closure of their interiors. Elements of \mathcal{K} are called **parcels**. For any set $Z \subseteq \mathbb{R}^n$ let $\mathcal{K}_Z \equiv \{K \in \mathcal{K} : K \subseteq Z\}$ denote the restriction of \mathcal{K} to Z .

Consider a path-connected set $Z \subseteq \mathbb{R}^n$ and let $x, y \in Z$. Let $\mathcal{P}_Z(x, y)$ be the set of continuous paths $g : [0, 1] \rightarrow Z$ for which $g(0) = x$, $g(1) = y$, and $g([0, 1]) \subset Z$. For any path g in $\mathcal{P}_Z(x, y)$, we define the length $l(g)$ in the usual way.²¹ We define the distance from x to y within Z as:

$$d(x, y; Z) \equiv \inf_{g \in \mathcal{P}_Z(x, y)} l(g).$$

We define $d(x, y; \mathbb{R}^n) \equiv d(x, y)$. This is the Euclidean metric.

²¹ That is, suppose $g : [0, 1] \rightarrow Z$ is continuous. Let $k \in \mathbb{N}$. Let $(t_0, \dots, t_k) \in \mathbb{R}^{k+1}$ satisfy for all $i \in \{0, \dots, k-1\}$, $t_i < t_{i+1}$. Define $l_t(g) = \sum_{i=1}^k \|g(t_k) - g(t_{k-1})\|$. The length (formally, the **arc length**) of g is then defined as $l(g) = \sup_{k \in \mathbb{N}} \sup_{\{t \in [0, 1]^k : t_i < t_{i+1}\}} l_t(g)$. Technically this quantity may be infinite, but we abstain from a discussion of this issue as it appears to be of no practical relevance.

Let \mathcal{F} be the set of density functions $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ such that $\int_K f(x)dx$ is finite for all parcels $K \in \mathcal{K}$. Let $f_u \in \mathcal{F}$ refer to the uniform density.²² For any density function $f \in \mathcal{F}$, let F be the associated probability measure so that $F(K) \equiv \int_K f(x)dx$ represents the population of parcel K .²³

We measure compactness of districts relative to the borders of the state in which they are located. Given a particular state Z ,²⁴ we allow the measure to consider two factors: (1) the boundaries of the legislative district and (2) the population density.²⁵ Thus, a **measure of compactness** is a function $s_Z : \mathcal{K}_Z \times \mathcal{F} \rightarrow \mathbb{R}_+$.

The Basic Family of Compactness Measures

As a measure of compactness we propose to use the expected relative difficulty in traveling between two points within the district. Consider a legislative district K contained within a given state Z . The value $d(x, y; K)$ is the shortest distance between x and y which can be traveled while remaining in the parcel K . To this end, the shape of the parcel K makes it relatively more difficult to get from points x to y the lower the value of

$$\frac{d(x, y; Z)}{d(x, y; K)}. \quad (1)$$

Note that the maximal value that Expression (1) may take is one, and its smallest (limiting) value is zero. Alternatively, any function $g(d(x, y; Z), d(x, y; K))$ which is scale-invariant, monotone decreasing in $d(x, y; K)$, and monotone increasing in $d(x, y; Z)$ is interesting; Expression (1) can be considered a canonical example. The numerator $d(x, y; Z)$ is a normalization which ensures that the measure is affected by neither the scale of the district nor the jagged borders of the state. We obtain a parameterized family of measures of compactness by considering any $q \geq 0$; so that $[d(x, y; Z)/d(x, y; K)]^q$ is our function under consideration, defining

$$\left[\frac{d(x, y; Z)}{d(x, y; K)} \right]^\infty = \begin{cases} 1, & \text{if } \frac{d(x, y; Z)}{d(x, y; K)} = 1 \\ 0, & \text{otherwise} \end{cases}.$$

Note that for $q = 0$, the measure is degenerate. This expression is a measure of the relative difficulty in traveling from points x to y . Our measure is the expected relative difficulty over all pairs of points, or:

$$s_Z^q(K, f) \equiv \int_K \int_K \left[\frac{d(x, y; Z)}{d(x, y; K)} \right]^q \frac{f(y)f(x)}{(F(K))^2} dy dx. \quad (2)$$

²² We define $f_u(x) = 1$.

²³ Similarly, the uniform probability measure $F_u(K)$ represents the area of parcel K .

²⁴ The state Z is typically chosen from set \mathcal{K} but is allowed to be chosen arbitrary; this allows the case where $Z = \mathbb{R}^n$ and the borders of the state do not matter.

²⁵ The latter factor can be ignored by assuming that the population has density f_u .

We note a few important cases. First, the special case of $q = +\infty$ corresponds to the measure described in the introduction, which considers whether the district contains the shortest path between pairs of its points.²⁶ Second, we can choose to measure either the compactness of the districts' shapes (by letting $f = f_u$) or the compactness of the districts' populations (by letting f describe the true population density). Third, if $Z = \mathbb{R}^n$, our measure describes the compactness of the legislative district without taking the state's boundaries into consideration.

Discrete Version

Our measure may be approximated by treating each census block as a discrete point. This may be useful if researchers lack sufficient computing power to integrate the expression described in (2).

Let $Z \in \mathbb{R}^n$ be a state as described above, and let $K \in \mathcal{K}_Z$ be a district. Let $\mathcal{B} \equiv \mathbb{R}^n \times \mathbb{Z}_+$ be the set of possible census blocks, where each block $b_i = (x_i, p_i)$ is described by a point x_i and a non-negative integer p_i representing its center and population, respectively. Let $Z^* \in \mathcal{B}^m$ describe the census blocks in state Z and let $K^* \subset Z^*$ describe the census blocks in district K . The approximate measure is given by:

$$s_{Z^*}^q(K^*) \equiv \left[\sum_{b_i \in K^*} \sum_{b_j \in K^*} \left[\frac{d(x_i, x_j; Z)}{d(x_i, x_j; K)} \right]^q p_i p_j \right] \left[\sum_{b_i \in K^*} \sum_{b_j \in K^*} p_i p_j \right]^{-1}.$$

DATA

To illustrate our measure we have calculated scores for all districts in Connecticut, Maryland, and New Hampshire during the 109th Congress. (See Figures 11–13.) Because of limitations in computing power we use the discrete approximation.

Dark lines represent congressional district boundaries, while shading roughly follows population distributions. Table 1 contains scores for our path-based measure as well as three others: the Schwartzberg measure, the Reock measure, and Convex Hull measure, which compares the area of a district to that of its' Convex Hull.²⁷ The small numerals in

²⁶ Mathematically, there may be two shortest paths in a parcel connecting a pair of residents. The issue arises when one state is not simply connected. For example, two residents may live on opposite sides of a lake which is not included in the parcel. In this general case, our measure is the probability that at least one of the shortest paths is contained in the district for any randomly selected pair of residents.

²⁷ In calculating the path-based measure, we have assumed a Mercator projection, so that the state of Colorado would be considered convex. For other measures, we have projected districts onto a spherical globe, to get accurate measurements of district area and boundary length. To calculate perimeters for the Schwartzberg measure we summed the lengths of the line segments that form the district boundary. In some cases, natural state boundaries (such as the Chesapeake Bay) added significantly to the total length. The Census data we used did not allow us to calculate district

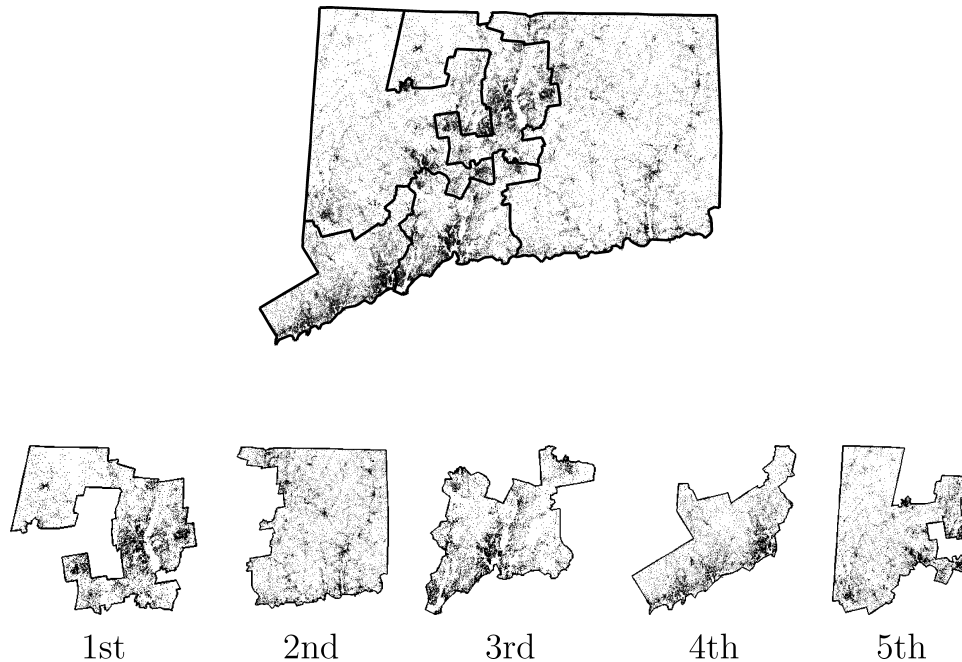


Figure 11. Connecticut.

parentheses give the ordinal ranking of the district according to the respective measure. Thus, according to our measure, Connecticut's Fourth District is the most compact, with a nearly perfect score of 0.977, followed by Maryland's Sixth District (0.926). Maryland's Third District is the least compact with a score of 0.140, which makes it slightly less compact than seven equally sized communities connected with a narrow path. (See Figure 3). The Schwartzberg measure ranks Connecticut's Second District as the most compact and Maryland's First District as the least compact. Like the Schwartzberg measure, the Reock and Convex Hull measures rank Connecticut's Second District as the most compact district. The Reock measure, however, ranks Maryland's Sixth District as the least compact district, while the Convex Hull measure places Maryland's Second District in last place. For these fifteen districts, the ordinal rankings (between our measure and the any of the other measures) agree on fewer than seventy-five percent of the pairwise comparisons.

Our measure gives strikingly different results than the others with respect to Connecticut's Fifth District and Maryland's Sixth District. All assign a high rank to one

tri-junctions (as recommended by Schwartzberg, 1966), although it seems unlikely that this would have a substantial effect on the calculation in this case. We do not know whether practitioners use a different method to calculate these scores.

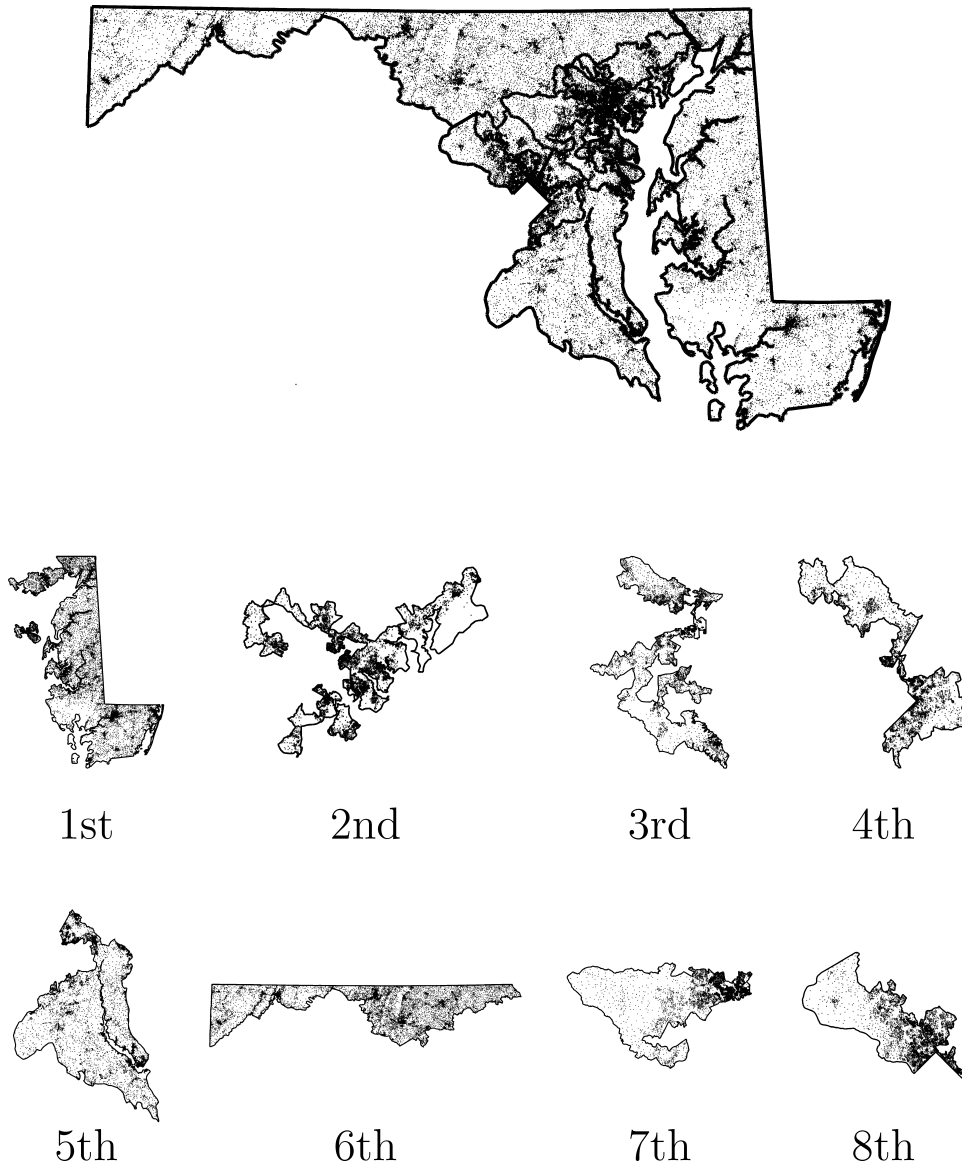


Figure 12. Maryland.

of the districts and a low rank to the other, but the order is reversed. The difference primarily stems from two factors: state boundaries and population.

Maryland's Sixth District has a very low area-perimeter ratio owing to its location in the sparsely populated panhandle of western Maryland and to the ragged rivers

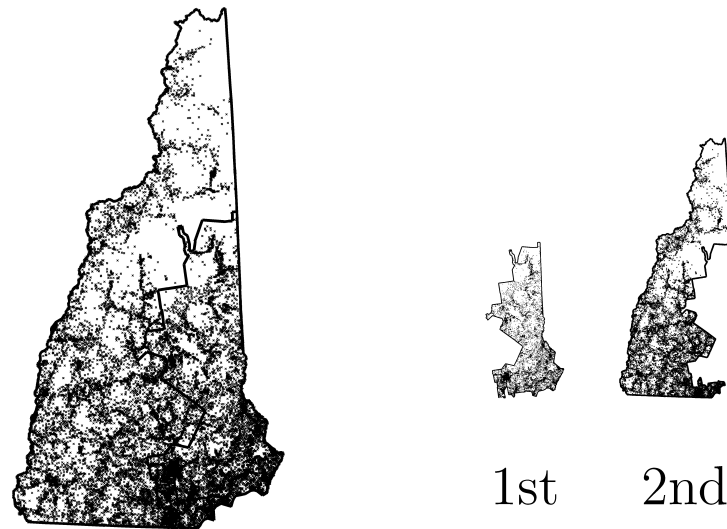


Figure 13. New Hampshire.

which makes up its southern and eastern borders. Its long shape makes the minimum circumscribing circle very large relative to its area, and its' convex hull includes a lot of territory outside of the district, mostly in West Virginia. Our path-based measure, however, takes the state boundaries into account and thus gives this district a high score.

Connecticut's Fifth District, however, has a much higher area-perimeter ratio: the generally square shape of the district compensates for the two appendages protruding from its eastern side. It also has high Reock and Convex Hull measures — the shape of the district, with appendages, fits nicely into a circle, and the appendages are relatively close to each other. However, the appendages reach out to incorporate several urban areas into the district. (See for example, the southeastern portion of the northern appendage and the eastern part of the southern appendage.) Because the major population centers are relatively disconnected from each other, our path-based measure assigns this district a low score of 0.481, which is slightly less compact than two equally sized communities connected with a narrow path. (See Figure 3.)

Although the Reock and Convex Hull measures are similar (both compare the area of the district to that of an ideal figure), they can produce very different results. For example, Connecticut's First District is assigned a high score by the Reock measure (the outside of the district is roughly circular) but a low score by the Convex Hull measure (it has a large hole on the inside). New Hampshire's Second District, on the other hand, is assigned a low score by the Reock measure (it is a long district from North to South) but a high score by the Convex Hull measure. It received a low score according to the path-based measure because the deviations from convexity are in the areas of highest population density — the southeastern corner of the state.

Table 1. Legislative district scores.

District	Measure:	Path-based	Schwartzberg	Reock	Convex Hull
<i>Connecticut:</i>					
1st		0.609 (8)	0.161 (9)	0.430 (3)	0.657 (10)
2nd		0.860 (4)	0.412 (1)	0.560 (1)	0.832 (1)
3rd		0.891 (3)	0.235 (4)	0.347 (4)	0.662 (9)
4th		0.977 (1)	0.305 (3)	0.293 (8)	0.676 (7)
5th		0.481 (12)	0.228 (5)	0.511 (2)	0.749 (2)
<i>Maryland:</i>					
1st		0.549 (10)	0.016 (15)	0.283 (10)	0.520 (12)
2nd		0.294 (14)	0.019 (14)	0.200 (12)	0.377 (15)
3rd		0.140 (15)	0.029 (13)	0.194 (14)	0.402 (14)
4th		0.366 (13)	0.083 (11)	0.198 (13)	0.444 (13)
5th		0.517 (11)	0.066 (12)	0.299 (7)	0.684 (6)
6th		0.926 (2)	0.119 (10)	0.121 (15)	0.562 (11)
7th		0.732 (6)	0.174 (8)	0.339 (5)	0.727 (3)
8th		0.657 (7)	0.204 (7)	0.310 (6)	0.690 (5)
<i>New Hampshire:</i>					
1st		0.801 (5)	0.290 (6)	0.293 (9)	0.670 (8)
2nd		0.561 (9)	0.370 (2)	0.233 (11)	0.705 (4)

CONCLUSION

We have introduced a new measure of district compactness: the probability that the district contains the shortest path connecting a randomly selected pair of its points. The measure can be weighted for population and can take account of the exogenously determined boundaries of the state in which the district is located. It is an extreme point in a parametric family of measures which vary according to the degree that they penalize deviations from convexity.

REFERENCES

- Altman, M. 1998. Districting Principles and Democratic Representation. Ph.D. thesis. California Institute of Technology.
- Arrow, K. J. and F. H. Hahn. 1971. *General Competitive Analysis*. Holden Day.
- Boyce, R. R. and W. A. V. Clark. 1964. "The Concept of Shape in Geography." *Geographical Review* 54: 561–572.
- Cox, E. P. 1927. "A Method of Assigning Numerical and Percentage Values to the Degree of Roundness of Sand Grains." *Journal of Paleontology* 1: 179–183.

- Edelman, P. H. 2006. "Getting the Math Right: Why California has too Many Seats in the House of Representatives." *Vanderbilt Law Review* 59: 297–346.
- Fryer, Jr., Roland G. and Richard Holden. 2007. "Measuring the Compactness of Political Districting Plans." Working Paper.
- Garner, B. 2004. *Black's Law Dictionary*. Eighth ed. Thomson West.
- Harris, C. C. 1964. "A Scientific Method of Districting." *Behavioral Science* 9: 219–225.
- Hirschman, A. O. 1964. "The Paternity of an Index." *The American Economic Review* 54(5): 761–762. <http://www.jstor.org/stable/1818582>
- Hofeller, T. and B. Grofman. 1990. "Comparing the Compactness of California Congressional Districts under Three Different Plans: 1980, 1982, and 1984." In *Political Gerrymandering and the Courts*, ed. Bernard Grofman. New York: Agathon.
- Kaiser, H. F. 1966. "An Objective Method for Establishing Legislative Districts." *Midwest Journal of Political Science* 10: 200–213.
- Lehrer, E. 2007. Private Communication.
- NCSL. 2000. *Redistricting Law 2000*. National Conference of State Legislatures.
- Niemi, R. G., B. Grofman, C. Carlucci and T. Hofeller. 1990. "Measuring Compactness and the Role of a Compactness Standard in a Test for Partisan and Racial Gerrymandering." *The Journal of Politics* 52: 1155–1181.
- Papayanopoulos, L. 1973. "Quantitative Principles Underlying Apportionment Methods." *Annals of the New York Academy of Sciences* 219: 181–191.
- Pildes, R. H. and R. G. Niemi. 1993. "Expressive Harms, "Bizarre Districts," and Voting Rights: Evaluation Election-District Appearances After *Shaw v. Reno*." *Michigan Law Review* 92: 483–587.
- Polsby, D. D. and R. D. Popper. 1991. "The Third Criterion: Compactness as a Procedural Safeguard Against Partisan Gerrymandering." *Yale Law and Policy Review* 9: 301–353.
- Reock, E. C. 1961. "A Note: Measuring Compactness as a Requirement of Legislative Apportionment." *Midwest Journal of Political Science* 5: 70–74.
- Schneider, R. 1975. "A Measure of Convexity for Compact Sets." *Pacific Journal of Mathematics* 58: 617–625.
- Schneider, R. 1993. *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge University Press.
- Schwartzberg, J. E. 1966. "Reapportionment, Gerrymanders, and the Notion of "Compactness"." *Minnesota Law Review* 50: 443–452.
- Taylor, P. J. 1973. "A New Shape Measure for Evaluating Electoral District Patterns." *The American Political Science Review* 67: 947–950.
- Tufte, E. 1973. "The Relationship Between Seats and Votes in Two-Party Systems." *The American Political Science Review* 67: 540–554.
- Vickrey, W. 1961. "On the Prevention of Gerrymandering." *Political Science Quarterly* 76: 105–110.
- Weaver, J. B. and S. W. Hess. 1963. "A Procedure for Nonpartisan Districting: Development of Computer Techniques." *The Yale Law Journal* 73: 288–308.
- Young, H. P. 1988. "Measuring the Compactness of Legislative Districts." *Legislative Studies Quarterly* 13: 105–115.
- Žunić, J. and P. L. Rosin. 2002. A Convexity Measurement for Polygons. In *13th British Machine Vision Conference*. pp. 173–182.