DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

A MEASURE OF BIZARRENESS

Christopher P. Chambers and Alan D. Miller

SOCIAL SCIENCE WORKING PAPER 1272

July 2007

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 will contain the shortest path between a randomly selected pair of its' points. The measure is defined relative to exogenous political boundaries and population distributions.

JEL classification numbers: D72, K00, K19

Key words: Compactness, Contiguity, Convexity, Disconnection Sensitivity, Path-based Measure, Elections, Gerrymandering

A Measure of Bizarreness*

Christopher P. Chambers and Alan D. Miller

1 Introduction

Hundreds of years ago, legislators discovered that the ultimate composition of a legislature is not independent of the means through which district boundaries are drawn. Hoping to stave off unemployment, legislators learned to master the art of *gerrymandering*: carefully drawing district boundaries to increase their electoral chances and political power. Like certain forms of painting and ballet, this art became more and more noticeable by the odd shapes it produced.[1]

Past attempts on the part of political reformers to fight gerrymandering have led to the introduction of vague legal restrictions requiring districts to be "compact and contiguous."[2] The vagueness of these legal terms has led to the introduction of several methods to measure district "compactness."[3] However, none of these methods is widely accepted, in part because of problems identified by Young [19] and Altman [1]. We argue that these laws were introduced with the aim of eliminating bizarrely shaped districts. To this end we introduce a measure of "bizarreness."

The primary problem with gerrymandering is that elections become less competitive when legislators draw district lines to strengthen their reelection chances. The "bizarre" shapes which result are merely a side-effect of this process.[4] Reformers have focused on

*Division of the Humanities and Social Sciences, Mail Code 228-77, California Institute of Technology, Pasadena, CA 91125. Emails: chambers@hss.caltech.edu; alan@hss.caltech.edu. We would like to thank Paul Healy, Ehud Lehrer, R. Preston McAfee, Matthew Spitzer, Peyton Young, and seminar participants at Caltech for their comments.

[1] In 1812 a district was said to resemble a salamander; one hundred eighty years later, another was likened to a "Rorschach ink blot test." *Shaw v. Reno*, 509 U.S. at 633.

[2] Thirty-five states require congressional or legislative districting plans to be "compact", forty-five require "contiguity", and only one requires neither. See [8]. There may also be federal constitutional implications. See Shaw v. Reno, 509 U.S. 630 (1993); Bush v. Vera, 517 U.S. 959 (1996).

[3] "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. [4]

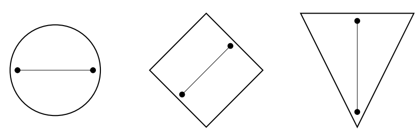
[4] However, the U.S. Supreme Court has held that "bizarre shape and noncompactness" of districts is not only evidence of unconstitutional manipulation of district boundaries but also "part of the constitutional problem." See Shaw v. Reno, 509 U.S. 630 (1993); Bush v. Vera, 517 U.S. 952, 959 (1996).

compactness because, while there is no consensus as to how district boundaries should be drawn, bizarre shapes are clearly identifiable as a symptom of gerrymandering.

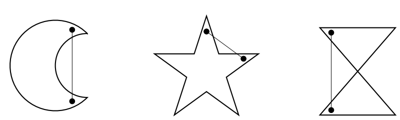
Part of the difficulty of defining a measure of compactness is that there are many conflicting understandings of the concept. According to one view the compactness standard exists to eliminate elongated districts. In this sense a square is more compact than a rectangle, and a circle may be more compact than a square. According to another view compactness exists to eliminate bizzarely shaped districts. According to this view a rectangle-shaped district would be better than a district shaped like a Rorschach blot.[5]

We follow the latter approach. While it may be preferable to avoid elongated districts, the classic sign of a heavily-gerrymandered district is bizzare shape.[6] We introduce a measure of convexity with which to assess the bizzareness of the district.[7] To the extent that elongation is a concern, it should be studied with a separate measure. These are two separate issues, and there is no obvious way to weigh tradeoffs between convexity and elongation.

The basic principle of *convexity* requires a district to contain the shortest path between any two 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 bizzarely shaped districts is that they are extremely non-convex. (See Figure 2.)



(a) Convex Shapes



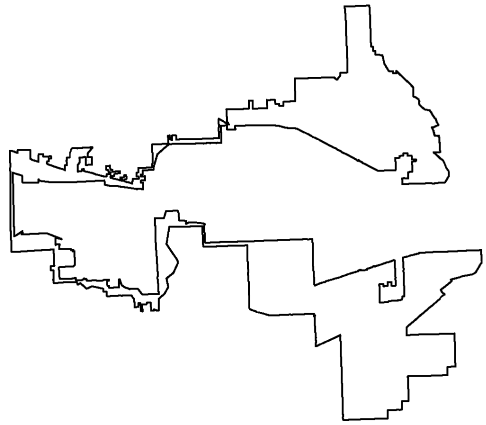
(b) Non-Convex Shapes

Figure 1: Convexity

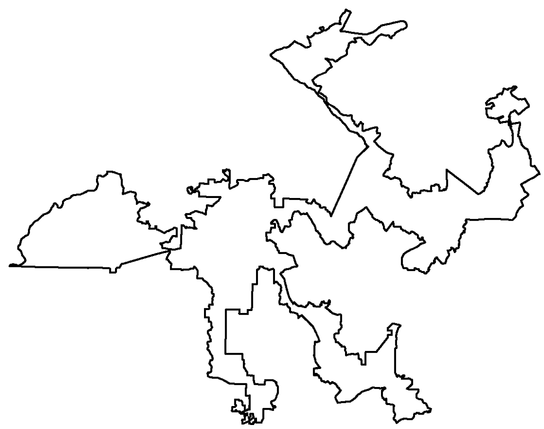
[5]The majority opinion in *Shaw v. Reno* noted that one district had been compared to a "Rorschach ink blot test" by a lower court and a "bug splattered on a windshield" in a major newspaper. 509 U.S. at 633.

[6]Note that the term gerrymander comes from the bizzare shape of a Massachussets legislative district which, in the view of a political cartoonist, resembled a salamander. Had the controversial district merely resembled a rectangle, the process of district manipulation would possibly be referred to as a *gerrytangle*.

[7]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 bizzareness.



(a) 4th District, Illinois



(b) 13th District, Georgia

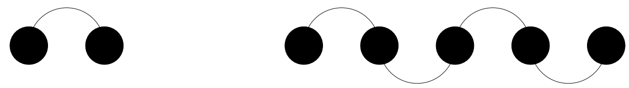
Figure 2: Congressional Districts, 109th Congress

The *path-based* measure we introduce is the probability that a district will contain the shortest path between a randomly selected pair of its' points.[8] This measure will always return 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 would assign this district a measure of approximately one-half. A district containing n towns connected by narrow paths would be assigned a measure of approximately $1/n$.[9] (See Figure 3(b).)

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 will contain the shortest path between a randomly selected pair of its' points is the same. However, the

[8] A version of this measure was independently discovered by Ehud Lehrer.

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



(a) Two Towns (b) Five Towns

Figure 3: Towns connected with narrow paths.

Figure 4: Same shapes, different populations

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 second district might represent one community with two forests attached.

Population can be incorporated by using the probability that a district will contain the shortest path between a randomly selected pair of its' residents. In practice our information will be more limited — we will 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.

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

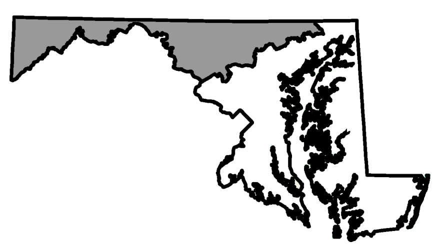


Figure 5: 6th District, Maryland, 109th Congress

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.

1.1 Related Literature

1.1.1 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 [19], Niemi et. al. [9], and Altman [1].

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 it's width. Ratios closer to one are considered more compact. This test has had some support in the literature, most notably Harris [5].[10]

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 [11], 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 it's center of gravity. Measures in this class were introduced by Boyce and Clark [2] and Kaiser [7]. The area-comparision and center of gravity measures have been adjusted to take account of district population by Hofeller and Grossman [6], and Weaver and Hess [18], respectively.

Dispersion measures have been widely criticized, in part because they consider districts reasonably compact as long as they are concentrated in a well-shaped area (Young [19]). 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.[11]

[10] 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).

[11] The length-width measure is the ratio of width to length of the circumscribing rectangle with minimum perimeter. See Niemi et. al. [9]. All measures are transformed so that they range between zero and one, with one being most compact. The Boyce-Clark measure is $\sqrt{\frac{1}{1+bc}}$, where bc is the original Boyce-Clark measure [2]. The Schwartzberg measure used is the variant proposed by Polsby and Popper

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

Compactness Measures

Dispersion Measures	*District:*	**I**	**II**
Length-Width		0.63	1.00
Area to Circumscribing Circle		0.32	0.44
Area to Convex Hull		0.57	0.70
Boyce-Clark		0.15	0.29
Other Measures			
Path-Based Measure		0.84	0.42
Schwartzberg		0.29	0.14
Taylor		0.40	0.20

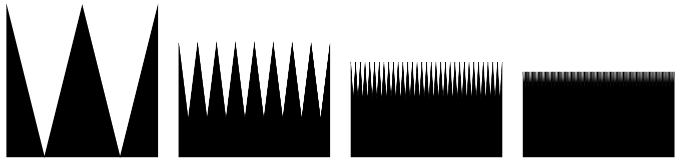
Perimeter measures use the length of the district boundaries to assess compactness. The most common perimeter measure, associated with Schwartzberg [14], involves comparing the perimeter of a district to its area.[12] An objection to the Schwartzberg measure is that it is overly sensitive to small changes in the boundary of a district. Figure 7 shows a sequence of shapes with their associated Schwartzberg measures. While the shapes in the sequence become arbitrarily close to a rectangle, the perimeter is increasing, and thus the Schwartzberg measure of the shapes decreases rapidly. By contrast, the path-based measure considers these shapes more compact the closer they become to rectangles.[13]

Taylor [15] introduced a measure of indentation which compared the number of reflexive (inward-bending) to non-reflexive (outward-bending) angles in the boundary 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.

[10] (originally introduced in a different context by Cox [3], or $(\frac{1}{sc})^2$, where sc is the measure used by Schwartzberg [14].

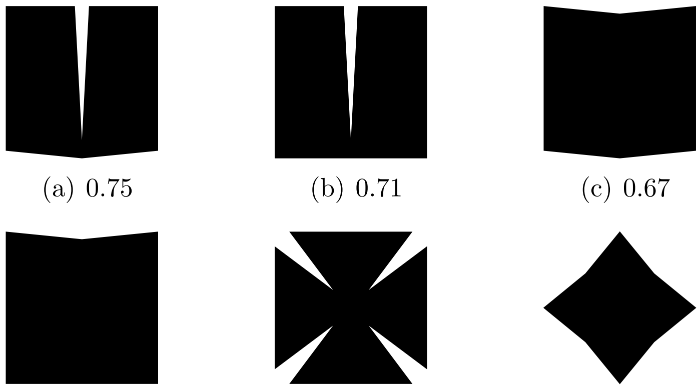
[12] This idea was first introduced by Cox [3] in the context of measuring roundness of sand grains. The idea first seems to have been mentioned in the context of district plans by Weaver and Hess [18] who used it to justify their view that a circle is the most compact shape. Polsby and Popper [10] have also supported the use of this measure.

[13] This example is theoretical; the extent to which this problem occurs in practice is a matter of debate. Young [19] suggests that jagged edges caused by the arrangement of census blocks may lead to the sort



(a) 0.143 (b) 0.059 (c) 0.019 (d) 0.005

Figure 7: Schwartzberg measure



(a) 0.75 (b) 0.71 (c) 0.67

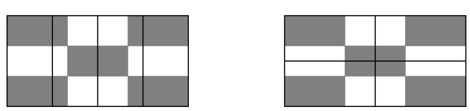
(d) 0.60 (e) 0.33 (f) 0.00

Figure 8: Taylor's measure

Lastly, Schneider [12] introduced a measure of convexity using Minkowski addition. For more on the relationship between convex bodies and Minkowski addition, see Schneider [13].

1.1.2 Other literature

Vickrey [17] showed that restrictions on the shape of legislative districts are not necessarily sufficient to prevent gerrymandering. In Vickrey's example there is a rectangular state in which support for the two parties (white and gray) are distributed as shown in Figure 9. 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.



(a) 2 gray, 2 white (b) All gray, no white

Figure 9: Vickrey's example

Compactness measures have been touted both as a tool for courts use in determining whether districting plans are legal and as a metric for researchers to use in studying the

of problem discussed here.

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 [16].

2 The model and proposed family of measures

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

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

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

Let $\mathcal{F}$ be the set of density functions $f : \mathbb{R}^n \to \mathbb{R}_+$ 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. 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.[14]

We measure compactness of districts relative to the borders of the state in which they are located. Given a particular state Z,[15] we allow the measure to consider two factors: (1) the boundaries of the legislative district, and (2) the population density.[16] Thus, a **measure of compactness** is a function $s_Z : \mathcal{K}_Z \times \mathcal{F} \to \mathbb{R}_+$.

2.2 The basic family of compactness measures

As a measure of compactness we propose to use the expected relative difficulty of traveling between two points within the district. Consider a legislative district K contained within

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

[15] 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.

[16] The latter factor can be ignored by assuming that the population has density f_u.

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

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

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 $p \geq 0$; so that $\left[\frac{d(x,y;Z)}{d(x,y;K)}\right]^p$ 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 $p = 0$, the measure is degenerate. This expression is a measure of the relative difficulty of travelling from points x to y. Our measure is the expected relative difficulty over all pairs of points, or:

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

The special case of $p = +\infty$ corresponds to the measure described in the introduction, which considers whether the district contains the shortest path between pairs of its points.

3 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.

3.1 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 in subsection 2.1 and let $K \in \mathcal{K}_Z$ be a district. Let $\mathscr{B} \equiv \mathbb{R}^n \times \mathbb{Z}_+$ 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 $Z^* \in \mathscr{B}^m$ 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:

$$s^p_{Z^*}(K^*) \equiv \left[\sum_{\substack{b_i, b_j \in K^* \\ i \neq j}} \left[\frac{d(x_i, x_j; Z)}{d(x_i, x_j; K)} \right]^p p_i \, p_j \right] \left[\sum_{\substack{b_i, b_j \in K^* \\ i \neq j}} p_i \, p_j \right]^{-1} .$$

References

[1] M. Altman. *Districting Principles and Democratic Representation.* PhD thesis, California Institute of Technology, March 1998.

[2] R. R. Boyce and W. A. V. Clark. The concept of shape in geography. *Geographical Review*, 54:561–572, 1964.

[3] E. P. Cox. A method of assigning numerical and percentage values to the degree of roundness of sand grains. *Journal of Paleontology*, 1:179–183, 1927.

[4] B. Garner. Black's law dictionary.

[5] C. C. Harris. A scientific method of districting. *Behavioral Science*, 9:219–225, 1964.

[6] T. Hofeller and B. Grofman. "Comparing the compactness of California congressional districts under three different plans: 1980, 1982, and 1984.". In B. Grofman, editor, *Political Gerrymandering and the Courts.* New York: Agathon, 1990.

[7] H. F. Kaiser. An objective method for establishing legislative districts. *Midwest Journal of Political Science*, 10:200–213, 1966.

[8] National Conference of State Legislatures. *Redistricting Law 2000.*

[9] R. G. Niemi, B. Grofman, C. Carlucci, and T. Hofeller. Measuring compactness and the role of a compactness standard in a test for partisan and racial gerrymandering. *The Journal of Politics*, 52:1155–1181, 1990.

[10] D. D. Polsby and R. D. Popper. The third criterion: Compactness as a procedural safeguard against partisan gerrymandering. *Yale Law and Policy Review*, 9:301–353, 1991.

[11] E. C. Reock. A note: Measuring compactness as a requirement of legislative apportionment. *Midwest Journal of Political Science*, 5:70–74, 1961.

[12] R. Schneider. A measure of convexity for compact sets. *Pacific Journal of Mathematics*, 58:617–625, 1975.

[13] R. Schneider. *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge University Press, 1993.

[14] J. E. Schwartzberg. Reapportionment, gerrymanders, and the notion of "compactness". *Minnesota Law Review*, 50:443–452, 1966.

[15] P. J. Taylor. A new shape measure for evaluation electoral district patterns. *The American Political Science Review*, 67:947–950, 1973.

[16] E. Tufte. The relationship between seats and votes in two-party systems. *The American Political Science Review*, 67:540–554, 1973.

[17] W. Vickrey. On the prevention of gerrymandering. *Political Science Quarterly*, 76:105–110, 1961.

[18] J. B. Weaver and S. W. Hess. A procedure for nonpartisan districting: Development of computer techniques. *The Yale Law Journal*, 73:288–308, 1963.

[19] H. P. Young. Determining the compactness of legislative districts. *Legislative Studies Quarterly*, 13:105–115, 1988.