Chapter 4. Modeling the Effect of Mandatory District Compactness on Partisan Gerrymanders

4.1. The Increasing Importance of Compactness in Redistricting

From the time that geographical districting was first used in the United States, irregularly shaped districts have been subjects of popular attention, evoking criticism from the press, but little action from the courts. In recent years, however, the courts have begun to scrutinize district lines. To aid the courts in this scrutiny, the academic community has developed formal methods of measuring the geographic compactness of districts.

I use a three-stage model to examine the electoral effects of formal district criteria. First, I equate the task of drawing compact districts to an optimization problem, which I solve with combinatorial optimization techniques. By treating the problem in this way, I am able to draw thousands of compact district plans that are free from personal bias. Second, I generate many possible population maps, according to different clustering functions. With these maps I abstract away the eccentricities of any one local area, and focus on the electoral effects of general population characteristics on redistricting. Third, I examine the electoral outcomes that would be most likely under each plan, and I relate these outcomes to the geographical compactness of the district.

District planners now may want to go in the direction set by the courts and create compact plans. However, how will they tell whether their plans are compact? If well-known, effective, fair compactness standards existed, the district planner’s job would be simple. Unfortunately, more than thirty compactness measures have been proposed, and none of these measures has been rigorously examined. Scholars disagree about the consistency of these measures, their effectiveness in preventing electoral manipulation, and their neutrality.

Current empirical research does not help district planners to determine which compactness measures to use, which are effective, and which are neutral. It is to this research we now turn.

It is a common assertion that compactness standards will constrain partisan gerrymanderers. This is a straightforward prediction, but misses the point. It is a simple mathematical truth that the maximum of a constrained optimization problem is less than or equal to the maximum of the same unconstrained problem. Hence if one thinks of redistricting as such a mathematical problem, in which one group has control of the redistricting process and acts single-mindedly to maximize a single goal, then any sufficiently restrictive constraint on redistricting plans will reduce the ability of that group to obtain its goal. This argument applies equally well whether the group is a team of Republican redistricting experts or a minority Political Action Committee, and whether the goal is to maximize the probability of partisan control of the legislature, or to maximize the number of minority opportunity districts.

There are many ways of constraining gerrymandering, including eliminating redistricting altogether; the important question is, what are the consequences of constraining redistricting in this way: Can compactness be measured consistently and sensibly? If so, which compactness measures should we use? How restrictive do compactness standards have to be in order to have an effect? To what extent will compactness standards limit redistricting for “good government” goals? Are compactness standards neutral — or will they systematically benefit certain political groups?

Several authors have examined the question of consistency with formal methods (Young 1988), and I will not repeat that work here. I will focus, instead, on the questions of effectiveness and neutrality, which have been debated so recently and hotly. I will pay special attention to the interaction of compactness rules and population characteristics.

4.2. Modeling Partisan Gerrymanders Under Compactness Rules

This chapter uses computer simulation to model the effects of geographical compactness on politics. Formal and empirical analysis in this area is limited. Formal models of redistricting are extremely simplified. Although formal models of redistricting exist, most of these ignore geography altogether, and those that include it do not model compactness or natural population distributions.

Empirical analysis of districting criteria cannot avoid selection bias. Real districts are not random samples but intentional creations, and there are many potential causes for “ugly” districts. Ill-compact districts may be caused by geographical constraints; by an underlying unevenness in the distribution of population across a state; by attempts to follow “natural” political boundaries; or by political attempts to manipulate lines for the benefit of communities of interest, racial minorities, political parties, or incumbents. These causes are difficult to measure and they may interact in complex and confounding ways — district lines drawn to protect incumbents in one district may be compact, in and of themselves, but may cause a neighboring district, drawn in absence of any political motive, to be ill-shaped. For these reasons, an unguided analysis of ex-post non-compact districts says little about the effects of compactness as a restraint on the political process ex-ante.

We can use simulation to draw districts that are driven only by compactness, and we can create a nearly unlimited numbers of these. Simulation allows us to abstract away from the geographical and political eccentricities of any single plan, and to directly analyze the relationships among the shapes of districts, the distribution of political groups, and the outcomes of elections – relationships that can be used as hypotheses to guide empirical analysis. As two of the strongest advocates of compactness write: "Enough knotty statistical issues must be overcome that probably the only way to settle this point (the effect of compactness standards) is through … running thousands of computer models of compact districts and seeing what happens" (Polsby and Popper 1991, 335 fn.).

Historically, computers have been frequently used to create redistricting plans — but usually as a tool to assist human planners (Browdy 1990a).[118] In Chapter 2, I showed

---

[118] There are some notable exceptions to this usage: Shepard and Jenkins (1970) and Taylor (1973) also use automation procedures to examine a range of districting options, but apply their techniques to the analysis of a particular election, rather than to an election rule, while Engstrom and Wilder (1977), Taylor and Johnston (1979), and O'Loughlin

how individual districts could be mathematically characterized. In this chapter, I treat redistricting, as a whole, as a mathematical set-partitioning problem, and use automated districting techniques to generate a series of arbitrarily drawn district plans. I complement these automated districting techniques with general combinatorial optimization algorithms that have been used successfully on similar problems in computer science. These same algorithms can be applied to a variety of formal districting rules. Because compactness standards are central to the current debate over redistricting, I use this procedure to focus upon plans drawn under equal population and compactness rules.[119]

(1982) argue that districts created automatically should be used as a benchmark with which to detect gerrymandering.

[119] You should note that, in these simulations, I model explicitly only compactness and equal population rules. In particular, I do not require plans to be contiguous. I do this for several reasons:

• First, in order to isolate the effects of compactness, the constraint set and value functions were kept as simple as possible.

• Second, contiguity is often, in practice, an ill-defined or vacuous requirement. Practically any set of regions can be made contiguous if lines are drawn finely enough.

4.2.1. Mathematically Characterizing Redistricting

If you were a mathematically-inclined district planner, you might characterize redistricting as a partitioning problem. (See, for example, Gudgin and Taylor (1979)) You would simplify the problem a bit by pretending that the state that you wish to redistrict is composed of indivisible[120] census blocs. Then, you would write out a function to evaluate partitions of blocs.[121] Finally, you would solve for the maxima of the problem — you would find the partition with the highest value. If you could perform this procedure then you would have the best district plan possible. Adding compactness standards does not make this problem more difficult to formulate. You can bring

---

• Third, compactness requirements encompass contiguity: the maximally compact plan will not be measurably and avoidably noncontiguous.

[120] This is not far from the truth since most population data is, at best, limited to the census-bloc level of detail.

[121] A partition divides a set into component groups that are exhaustive and exclusive. More formally:

For any set $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, a *partition* is defined as
a set of sets $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$ $s.t.$
(1) $\forall x_i \in \mathbf{x},\ \exists \mathbf{y}_j \in \mathbf{Y}, s.t.\ x_i \in \mathbf{y}_j$
(2) $\forall i, \forall j \neq i, \mathbf{y}_j \cap y_i = \varnothing$

compactness standards into the problem either as constraints to optimization, or as supplements to your value function.

This characterization of redistricting is useful whether you are trying to find the least biased plan or the most effective gerrymander: If you are an altruistic social planner, you would use a value function that weighed all of the social benefits and costs of redistricting. An altruist might include such factors as preserving county boundaries, maintaining the competitiveness of districts, and minimizing the "bias" of the plan in your value function (Lijphart 1989). A partisan, on the other hand, might attempt to maximize the number of safe party seats, or the probability of their party being in control of the legislature. Alternatively, a self-interested incumbent might attempt simply to maximize the probability of retaining her seat in upcoming elections.

Because altruistic social planners are likely to be outnumbered by partisans, incumbents, and other self-interested individuals, we may wish to impose rules on the redistricting process. Whether we require that all districts have the same population, that they respect political boundaries, maintain geographic contiguity, or that they comply with compactness criteria, we can represent these requirements as constraints on our optimization problem. Approaching redistricting mathematically has two advantages: This approach can help us to draw better districts, and it can help us to predict the effects of particular redistricting rules.

4.2.2. Analyzing The Effects Of Redistricting Rules

How can we use a mathematical characterization of redistricting to predict the effects of redistricting rules? Suppose you are a party leader, intent on producing a partisan gerrymander, and suppose that a citizens' group introduces an initiative requiring all districts to be compact. Should you expend political resources to fight this initiative?

To make this decision, you will have to estimate the effectiveness of the partisan gerrymander you expect to obtain when there are no rules, and then you must weigh that estimate against the effectiveness of the gerrymander which you expect to obtain if you are forced to draw contiguous districts. In mathematical terms, you subtract the value of the optimal partition of the constrained problem from that of the optimal partition of the unconstrained problem. If the difference is big enough, then you should fight the contiguity rule.

In this example, we assumed that the best plan that is found would then be *chosen*. This corresponds to the situation in which an organized group has substantial control over the districting process. In essence, this is a "game" played between a party and the courts. The controlling party submits a redistricting plan, which the court may accept, modify or reject.[122]

4.2.3. Creating Arbitrary Redistricting Plans

While we can easily *formulate* redistricting as a partitioning problem, this problem may be difficult to solve. Political scientists have used a number of different methods to

---

[122] Some caution would be warranted in extending this model of partisan gerrymanders to incumbent gerrymanders, in which each incumbent vies for a district that maximizes her chance of reelection. Unlike a partisan gerrymander, created by party leadership, the incumbent gerrymander may not be the result of an individual choice *per se*, but the result of a strategic game with multiple players. In practice, even if the optimal incumbent gerrymander were known, it might not be the plan that emerges from the smoke-filled rooms in the back of the state house.

Despite the possibility of game-theoretic complications in incumbent gerrymanders, it is valuable to understand how district rules affect the optimal plan. In some cases, a plan will, in effect, be chosen by party leadership or by some other unified group. Even if the redistricting process is best modeled as a game in this case, we still must understand the payoffs to players under different rules (i.e., the expected value of the optimal partition), before we can analyze equilibria of the game.

search for compact districts. At the same time, computer scientists have developed similar techniques to search for optimal partitions.

Exact methods systematically examine *all* legal districts either explicitly or implicitly. Explicit enumeration ("brute force" search) methods literally evaluate every possible plan (examples include (Gudgin and Taylor 1979; Shepherd and Jenkins 1970)),[123] but more sophisticated methods, such as "branch and bound," exclude groups of solutions that are obviously sub-optimal.

Unfortunately, as the number of population blocs increases, the number of potential plans grows so rapidly that no computer can evaluate all of them explicitly. Furthermore, redistricting problems belong to a set of problems for which it is widely believed that no guaranteed, feasible, optimization methods exist.[124]

---

[123] A close examination of these algorithms reveals that in order to make the programs finish in a reasonable amount of time the authors use unproven "short-cuts," making them, in actuality, heuristic. (See Chapter 5.)

[124] Technically, these problems can be proved to be in the set of problems that computer scientists have labeled "NP-Complete." (See Chapter 5.) Most computer scientists believe that to solve an NP-complete problem exactly requires you to spend computation time that grows exponentially with the problem's size ("n" above), although

In the larger simulations, because of these barriers to finding an exact solution to larger problems, I turn to *heuristic* methods for finding compact districts. These methods by definition, do not guaranty an optimal solution, but often perform well in optimization. Most heuristics for locating optimal partitions are based on the principle of iterative improvement. In the simplest of these methods, known as *hill climbing*, you start with a set of randomly generated redistricting plans and repeatedly look for small changes to the plan that improve it — stopping when there are no small alterations that can yield an improvement. Several previous researchers have used variants of hill climbing methods to draw new district plans or to make improvements to existing districts (Liittschwager 1973; Moshman and Kokiko 1973; Nagel 1972; Rose Institute of State and Local Government 1980; Vickrey 1961; Weaver and Hess 1963). I use a variant of hill climbing similar to Nagel's method. In addition, where it is possible to derive the most compact plan from geometric arguments. I use a simple "descent" algorithm: I start with the most compact plan and use hill climbing to create less compact variants of it. (See the appendix to this chapter.) In the illustration below, I show an example of this process (Figure 4-1).

---

there may sometimes be ways to more quickly find approximate solutions. See Papadimitriou 1994 (1994) for a general introduction to NP-completeness.

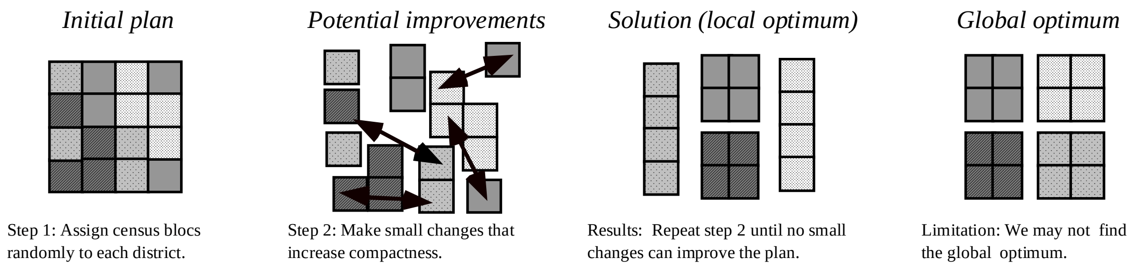

Figure 4-1. An example of creating a compact plan through "hill climbing."

Simple hill climbing methods, however, sometimes produce results that are far from optimum, because these methods are trapped easily in local optima. To minimize this problem in the simulations, I test hill-climbing methods against a number of methods that have been successfully used to solve similar problems in other fields: simulated annealing, genetic algorithms, and Monte Carlo methods. This variety of different methods helps to ensure that results are not being driven by quirks in the optimization process.

Simulated annealing is one of the most successful of combinatorial optimization methods. It is based on a mathematical analogy to the slow cooling of metal. If the value function being optimized is sufficiently well behaved, simulated annealing asymptotically converges to the optimum value (van Laarhoven and Aarts 1989). It has been previously recommended for use in redistricting (Browdy 1990b). *Genetic algorithms* are search algorithms based on an analogy to natural selection and genetic combination. Potential solutions to the optimization problem are defined as genetic strings, which can be mutated or "crossed" with other strings. A group of potential solutions then competes to survive and reproduce in the next generation.

Chandrasekharam (1993) demonstrates the effectiveness of genetic algorithms for graph-partitioning problems, which are somewhat similar to the redistricting problem. For details on the algorithms used in this chapter, see the appendix.

I cover the range of procedures that are available for drawing compact districts. Should compactness standards be legally mandated, district planners will have little choice but to turn to such techniques. Thus the plans that I produce, though simpler than real district plans, are similar in principle to those that will be produced should compactness standards become widespread.

4.2.4. Measuring Compactness

All of the optimization methods that I have discussed are flexible enough to accommodate a variety of value functions. Yet, choosing a particular compactness measure was a special challenge because previous researchers have proposed over thirty distinct measures of compactness (Niemi, et al. 1991; Young 1988). Neither the courts nor political scientists recognize a single standard for measuring compactness. In addition, although many states require compactness, only three states (Iowa, Colorado, and Michigan) define the term (Grofman 1985). What compactness standards represent the set best?

Most measures in the compactness literature fall into one of three categories: "area-based" measures, "perimeter-based" measures, and "population-based" measures. For the simulations, I chose a compactness standard from each of these categories. In addition, each of the measures that I use duplicates, as closely as is practical, legal standards of compactness the United States:

To measure the compactness of a district's area, I compared the area of the district to the area of the smallest box bounding that district.[125] A plan's compactness is defined as the mean compactness of its districts. This measure duplicates, as much as possible, a compactness requirement used in Iowa and in Michigan.[126]

---

[125] Some similar compactness measures use a bounding circle or convex polygon. While these other shapes are theoretically preferable, I use a bounding box here because of the discreteness of the simulation, map, and the desire to have my standard duplicate current legal standards. Furthermore, given the large granularity of population blocs in this simulation this measurement is quite similar to comparing districts to the bounding circle, while being more efficient to compute.

[126] These measures existed in state statutes or constitutions at the time of writing, but are usually not enforced. The 1980 Iowa General assembly Bill generally defines compactness as, "Compact districts are those which are square, rectangular or hexagonal in shape to the extent permitted by natural or political boundaries." The Michigan constitution also specifies, generally, that its state house districts should be "as nearly square in shape as possible." Iowa also offers several operational definitions of compactness, the first of which is "the absolute value of the difference between length and width" (Grofman 1985, 180 fn.). While this absolute value is not equivalent to my

To measure the compactness of a plan's boundaries, I calculated the total perimeter of all its districts; the best plan minimizes this total. Colorado currently uses this measure to evaluate districts (Grofman 1985).

To measure the compactness of a district's population,[127] I calculated the moment-of-inertia for the district's population.[128] A plan's compactness is defined as the mean compactness of its districts. This approach is similar to the measure in force in Iowa, which has the only population measure currently in effect in the U.S.[129]

---

bounding-box measure, it is a similar, if cruder, attempt to capture the squareness of a district.

[127] Since in the simulations, all population blocs have the same weight (although the partisan proportion in each bloc may vary), population-based measures produces results identical to analogous area-based measures.

[128] Formally, this is $\frac{P}{\sqrt{\sum_{x \in X}\left(p_x d(x)\right)^2}}$, where P is the total population of the district, p_x is the population of a particular census bloc, X is the set of all census blocs in the district, and *d()* is the geographical distance from the center of the census bloc to the population-center of the district.

[129] The Iowa measure calls for taking the ratio of the "dispersion of population" around the population center of the district, to the dispersion around the geographic center

Where possible I normalized their scores to fall within the (0,1] interval. Plans that have a score of "1" are as compact as possible. (This normalization was not possible for the perimeter measure, since it is not always possible to know the value of the perimeter of the most compact plan.)

4.2.5. Simulated Politics

In the simulation, *m*x*n* grids represent maps. Two different political groups populate each map. To separate the effect of compactness from the effects of equal population standards, each census bloc is normalized to have one hundred voters, so that only the proportion of each type of voter varies across population blocs. Each group runs a candidate in each district, and members of that group will vote "sincerely" for a candidate identified with that group. These assumptions best fit polarized, bipartisan elections.

Although I assume political groups have symmetric voting behavior, I allow them to be distributed across the state "map" differently. I duplicated the simulations using three different models to determine the political composition of each census bloc.

---

of the district. Since in this simulation I use only maps with uniform population densities, these two centers are always identical, and hence this ratio will always be one. Under these conditions, the population moment-of-inertia measure that I use better captures population dispersion.

I first performed a set of simulations using a very simple population model that I will refer to as the "uniformly-random" distribution. Unless otherwise noted a total of 10,000 plan/population distribution combinations were examined for each simulation run. In this model, the population of each census bloc is drawn from the same normal distribution.[130]

Clustering is a feature of realistic population geography models (Garner 1969). And in the second set of simulation runs, I used a "clustered" distribution to model the distribution of political groups. In this "clustered" distribution one political group is concentrated into *r* compact clusters, each of size *s*, and each randomly located. Similar cluster models have been used previously to explain voting behavior; in particular, Gudgin & Taylor (1979) find that the well known "cube law" of elections can be explained by a variation of the cluster population model.

In the third set of simulation runs, I use a more complicated clustering process that is based on Schelling's (1978) neighborhood formation model. In Schelling's model, persons in two different groups are at first randomly distributed on a map, then if an individual is surrounded by too many individuals of the other type, they can move to any adjacent square, if they prefer. In each round, every person is offered an opportunity to

---

[130] I duplicated each of the simulations in this model, substituting uniform distributions (with the same mean) for normal distributions, but the simulation results were indistinguishable.

move; when no one can improve their location by moving, the neighborhood is 'stable' (Schelling 1978). My model is similar to Schelling's, although it is aggregated at the census bloc level.[131] I show a typical population distribution that this model generated (Figure 4-2):

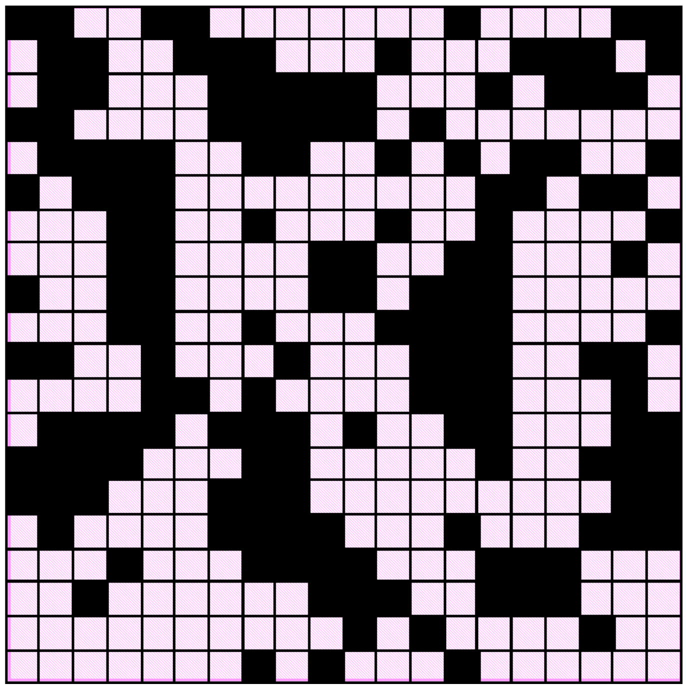

Figure 4-2. A map of 20x20 census blocs, with two political groups distributed across it using a Schelling distribution. The black squares represent census blocs primarily occupied by the minority political group.

131 Schelling requires that individuals move into empty spaces, whereas I allow two willing individuals to trade places.

4.3. Evaluating The Electoral Effects Of Compactness: Simulation Results

The results from these thousands of simulations reveal three interesting properties of compact plans and of compactness standards. First, the distribution of compact plans shows that compactness measures are useful only for comparing similar plans, but not for making absolute measurements of plans. Second, the simulation shows the difficulty of drawing compact plans under some measures — we should avoid these particular measures if we want to minimize the potential for gerrymandering. Finally, the simulation shows that district compactness can systematically influence election results.

4.3.1. Compactness Measures Are Relative Measures

I used an exhaustive analysis of districts to generate all possible plans for a number of small maps. I use a box-plot to compare the distribution of compact plans under the perimeter measure, for different map sizes, shapes, and numbers of districts (Figure 4-3):

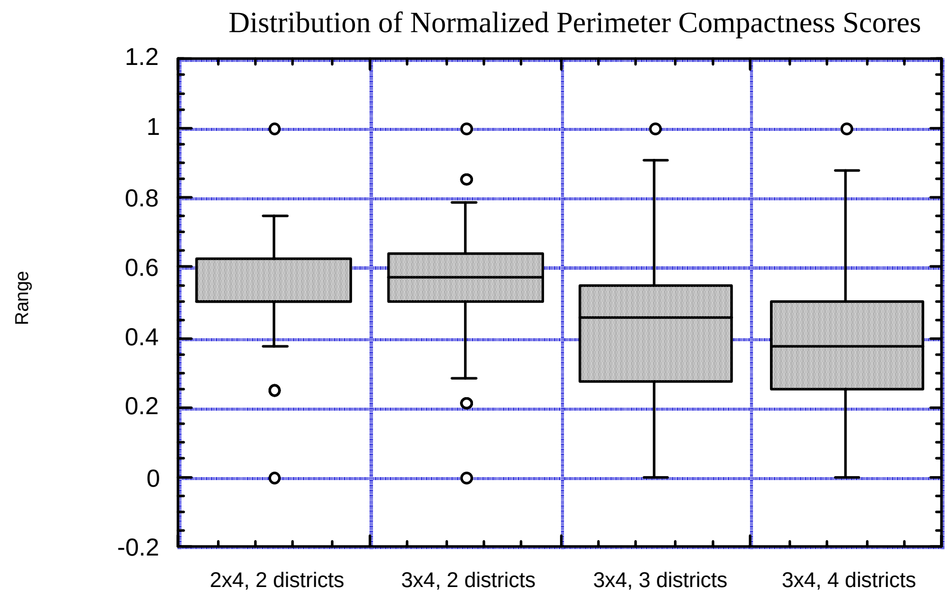

Figure 4-3. Box plots of Plan Compactness. The perimeter measure is normalized to the (0,1) interval.

Looking at any one of these box-plots, you can easily see that compact plans are scarce — most plans fall *far* short of the optimum. Furthermore, the scores of most plans cluster in a very narrow range of compactness values.

If you compare the box-plots for different maps, you can see that while the distribution of scores for each map is similar, the values of the minimum, maximum, and median scores are quite different. Given only the compactness scores of two plans, you can make reasonable comparisons between them only if these plans partition the same map into identical numbers of districts — a compactness score is meaningless outside of its specific context. For example, a score of “.5” is in the bottom decile of plans for the first map in Figure 4-3, and in the *top* decile for the last map.

Some authors have proposed that the courts mandate a minimal level of compactness for district plans. Given the shape of the distribution of compactness scores observed here, the effectiveness of compactness standards for limiting manipulation is likely to be very sensitive to the particular minimum level specified. If the minimum level is set high, the vast majority of plans will fail to meet the standard — it may be difficult to draw any plans at all. If, on the other hand, it is set at the middle of the distribution, the ability to gerrymander may be virtually unaffected.

Empirical studies of compactness scores must also take note of both their nonlinearity and their sensitivity to geographic context. Suppose that your ability to gerrymander is roughly proportional to the number of plans from which you can choose: Then, you will find it immensely more difficult to create an effective gerrymander that scores in the top 99th-percentile than to draw a plan with a slightly lower relative score. Furthermore, since compactness scores will depend on state boundaries, you may find it easy to create a gerrymander that scores “.90” in a state with regular boundaries, like Iowa, and impossible to create any plan at all that scores above “.75” in a state like Maryland. Differences in population distribution can be expected to further cloud such comparisons, as the ability to draw compact districting plans will be affected by equal population constraints. In general, comparing the compactness of plans across different states has little value.

4.3.2. Some Compactness Standards Make Detection Of Gerrymanders Difficult

In Table 4-1, I compare the performance of each optimization method. (This measure is limited to cases where the optimal plan can be deduced from regularities in the shape and population distribution.) Both the hill-climbing method and the genetic algorithm were equally successful in finding optimal plans, although the genetic algorithm was too computation-intensive to use on the larger maps.[132] Unlike these two methods, both the Monte Carlo procedure and simulated annealing performed poorly.[133]

---

[132] The time required to find a solution using the hill climbing method seemed to grow quadraticly in the number of census blocs (in time-complexity notation) while the convergence rates for genetic algorithms and simulated annealing grew at an even faster rate. Consequently, for maps larger than 5x5, I used descent and hill-climbing methods exclusively.

[133] The annealing procedure that I used sometimes made trades that would cause the population of the districts to become unbalanced, lowering the overall score of the plans. Once it did this, it was usually unable to recover because future changes to the plan were unlikely to bring the plan back into balance. Although I felt constrained to use well-documented and generalized algorithms, in practice, it should be possible to modify the algorithm for better performance with given redistricting goals.

Grid Size	# of Districts	Number of Possible Plans	Measure	Best Possible Score	Mean of 1000 Random District (std dev)	Hill-Climb mean (std dev)	Anneal Mean[134] (std dev)	Genetic Mean (std dev)
3x4	4	15400	perimeter	32	42 (3.0)	32.1 (0.41)	---------	32 (0)
3x4	4	15400	area	0.75	0.33 (0.07)	0.73 (0.07)	0.45 (0.096)	0.73 (0.03)
3x4	4	15400	moment[135]	1.5	0.98 (.12)	1.0 (0.02)	1.2 (0.1)	1.49 (0.06)
5x5	5	5.2×10^{12}	perimeter	unknown	86 (4.6)	54.8 (2.8)	---------	55.5 (4.3)
5x5	5	5.2×10^{12}	area	unknown	0.23 (.03)	0.42 (0.065)	0.02 (0.07)	0.42 (0.046)
5x5	5	5.2×10^{12}	moment	unknown	0.61 (0.04)	0.99 (0.001)	---------	1.03 (0.05)[136]
8x8	4	$5.0x10^{53}$	perimeter	64	203 (9.3)	90.5 (8.7)	---------	----------
8x8	4	$5.0x10^{53}$	area	1.0	0.25 (.006)	0.27 (0.015)	---------	-----------
8x8	4	$5.0x10^{53}$	moment	1.0	0.34 (.007)	0.65 (0.02)	---------	-----------
9x9	9	$1.5x10^{65}$	perimeter	108	295 (7.3)	150 (9.8)	---------	-----------
9x9	9	$1.5x10^{65}$	area	1.0	0.12 (.006)	0.18 (0.026)	---------	-----------
9x9	9	$1.5x10^{65}$	moment	1.0	0.32 (.009)	0.87 (0.03)	---------	-----------
20x20	8	$3.0x10^{728}$	perimeter	160	1508 (13)	-------------	---------	-----------

134 Annealing returned plans that violated equal population constraints, these plans were assigned a compactness value of 0.

135 If measured over a continuous area, the moment of inertia measure for a shape can be no larger than one, but this condition is violated in very small discrete approximations.

136 Sample size in this case was 148, because computation exceeded time limit.

20x20	8	$3.0x10^{728}$	area	1.0	0.06 (.001)	-------------	---------	-----------
20x20	8	$3.0x10^{728}$	moment	1.0	0.13 (0.0)	0.36 (0.003)[137]	---------	-----------

Table 4-1. Performance of algorithms using different measures of compactness.[138]

In fact, the most striking differences in this chart are not among methods, but among compactness measures. I show the ratio ("approximation ration") between the mean value of the plans created using the best optimization method to the value of the most compact plans possible. Notice that, in general, these methods were much more successful at finding compact plans under the perimeter standard and moment-of-inertia standard than under the area-based standard. What does this tell us about the properties of these different standards? (Figure 4-4)

---

[137] Sample size in this case was 287, because computation exceeded time limit.

[138] For each table entry I performed 1000 simulation samples, unless otherwise noted. When a cell filled is filled with dashes it means that the specified algorithm was not able to complete a significant number of iterations before the time limit (several days) expired.

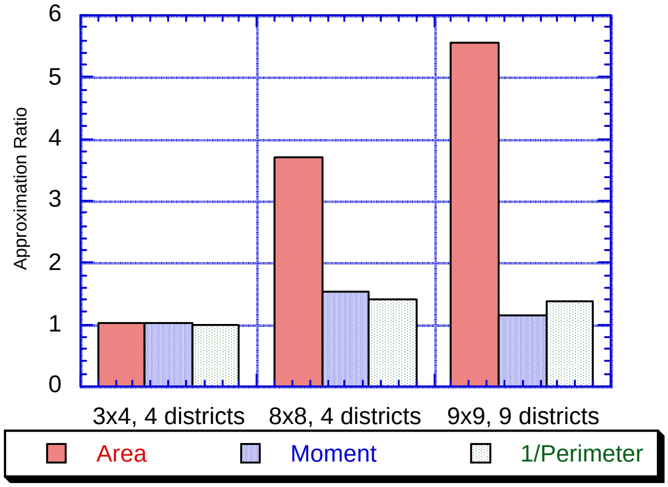

Figure 4-4. To obtain an approximation ratio, I divide the best possible score by the mean reached by the best algorithm.[139] The best possible ratio is one, which means that the algorithm always reaches the best possible solution.

Remember that all of these optimization methods rely upon iterative improvement. In other words, they operate through change that is gradual and limited. Since, as we have seen, these methods work well for the perimeter-based standard, we can conclude that the perimeter and moment standards are sensitive to small changes in a district plan; on the other hand, the area-based measure is much less sensitive to small changes— to improve the plan's area compactness we need to change districts radically.

---

[139] I use the inverse of the perimeter here because the perimeter measure grows when a shape becomes less compact, unlike the others.

Changing districts radically can be politically difficult and can interfere with other redistricting goals, such as preserving natural boundaries and communities of interest. While the simulations ignore these concerns, the courts should not. Because of these difficulties, the courts will find the perimeter-based standard easier to manage than the area-based standard. (In coming to this conclusion, I retain the assumption from Section 4.2.2 that a single party substantially controls the redistricting process, and is able to create plans that the court must then either approve, modify, or reject.)

Furthermore, for most real district maps we will not know the value of the most compact plan beforehand, a situation that is exacerbated by the area-based measure. Since it is likely to mislead us with plans that are locally optimal, but which fall far short of the most compact plans, the area-based measure allows gerrymanderers much greater leeway in designing their districts. Altruistic district planners will suffer as well, as they may expend unnecessary effort trying to improve a plan that is already very close to optimal.

4.3.3. Compactness Standards May Create Opportunities For Political Manipulation.

In addition to failing to prevent gerrymanders, there is a further consideration that has not been suggested in the literature as yet: The process of evaluating plans under a compactness standard might well induce strategic behavior that would harm the reapportionment process.

Finding the maximally compact plan is, as I have indicated, a very difficult mathematical problem. In practice, it will often not be possible to determine whether a plan is “optimally compact,” especially for plans composed of a large numbers of census blocs. Moreover, as I argued in the Section 4.3.2, the simulation results indicate that if we do not know the value of the optimal plan, we cannot set reasonable an “absolute” compactness limit.

Instead of using some absolute measuring value, we will have to compare plans against each other, or simply search for “improvements” to any proposed plans. It has even been suggested that when two plans are proposed, the most compact should automatically be implemented (Polsby and Popper 1991).

Unfortunately, in a strategic political environment, in which plans are compared only with each other, the very shape of districts becomes valuable information to your opponents: If you hide your plan, there is a chance that opponents will mistakenly believe their plan to be the most compact, which is to your advantage. Whereas if you reveal your plan, you give up this strategic advantage without gaining anything. In sum, because compactness standards give district planners an incentive to hide information, these standards may increase political manipulation.

4.3.4. Compactness Standards Are Not Politically Neutral

In this next section, I will show that there is a systematic relationship between compactness standards, population distribution and electoral advantage. The specific effect that these standards will have on redistricting, however, will depend on the political institutions used to create districts. Here I examine the effects of compactness on partisan gerrymanders and on automated redistricting plans.

Arbitrarily Selected Compact Districts

Polsby and Popper 1991 claim that if the court adopts a policy of automatically accepting the most compact districting plan proposed to them, then through competition among political groups, gerrymandering will disappear. This view of arbitrary compact district plans is relatively recent, but scholars have long argued that we should simply use a computer to generate arbitrary district plans, following only the principles of compactness, contiguity and population equality (Harris 1964; Kaiser 1966; Weaver and Hess 1963). Suppose that we did manage to create districts arbitrarily, following only the principles of compactness and population equality, as these scholars desire. Would this be a neutral solution to the gerrymandering problem? Can we reach color-blindness if we choose the first horn of Justice Souter's dilemma by awarding primacy to compactness standards?

In Table 4-2 I show the correlation between compactness and electoral results in such a case. Since neither compactness scores nor seats were distributed normally, I also report Somers's d along with the correlation measures. Somers's d is similar to Kendall's Tau, except that it treats ties asymmetrically, ignoring ties on the dependent variable (number of minority seats, in this case). I use it in this case because of the number of minority districts takes on only a few values, leading to many ties that would distort Kendall's measure. See (Liebetrau 1983) for a discussion of these measures.

Grid Size	Number of Districts	Cluster Size	Number of Clusters	Minority Percentage of Population	Mean Minority Controlled Districts (std dev)	Correlation Between Minority Controlled Districts and Compactness (Somers's d)
5x5	5	1	3	12%	0.02 (0.13)	0.0 (-0.03)
5x5	5	1	5	25%	0.34 (0.49)	0.0 (-0.01)
5x5	5	1	12	48%	2.37 (0.6)	0.06 (0.00)
5x5	5	4	1	16%	0.35 (0.47)	0.39 (0.44)
5x5	5	4	2	32%	1.17 (0.63)	0.36 (0.35)
5x5	5	4	3	48%	2.37 (0.6)	0.07 (0.08)
5x5	5	9	1	36%	1.46 (0.6)	0.32 (0.26)
8x8	4	1	26	40%	0.47 (0.43)	0.01 (-0.02)
8x8	4	9	3	42%	0.93(0.65)	0.46 (0.47)
8x8	4	16	1	25%	0.24(0.42)	0.46 (0.61)
8x10	8	9	3	34%	1.38(0.89)	0.58 (0.52)
20x20	8	1	100	25%	0.04(0.20)	0.0 (0.0)
20x20	8	4	40	40%	3.1 (1.2)	0.29 (0.23)
20x20	8	9	18	41%	3.84(1.85)	0.43 (0.33)
20x20	8	36	2	18%	1.38(1.03)	0.66 (0.65)
20x20	8	36	4	36%	4.1 (1.56)	0.58 (0.52)

Table 4-2. The effects of perimeter compactness on the representation of clustered minorities.[140]

(10,000 Samples were performed for each grid/district combination)

---

[140] I generated ten thousand district plan-population combinations for each set of district and population parameters. Since I controlled for the other parameters by keeping them constant across runs, a standard correlation measure adequately represents the linear association between the number of districts captured by a minority and the compactness of districting plans. Genetic algorithms, hill climbing and descent methods were used for the 5 by 5 case, while hill-climbing and descent methods were used for all other cases. Since the results from each method were similar, only hill-climbing results are reported.

As I had suggested earlier, the electoral effects of compactness depend upon the geographic distribution of political groups. When all political groups are thoroughly geographically mixed, no district, compact or not, can contain a majority of a minority group. Even when political groups form small geographic clusters, if these clusters are dispersed geographically, as in the uniform and normal population distribution models, then compact districts are no more likely to elect candidates from one political group than from another. This is not because the district drawing process is politically neutral in these circumstances. On the contrary, in these cases compactness has no effect on electoral outcomes because geographical districting itself embodies such a powerful majoritarian bias[141] that minority political groups are unlikely to win seats under *any* circumstances.[142]

---

[141] I use the term "bias" here in the descriptive, rather than normative sense — arbitrarily drawn districts tend to award district-share in excess of the majority's share of the population. See Grofman (1982) for a similar analysis of the majoritarian bias inherent in redistricting.

[142] Also important will be the relative geographic size of the district, minority clusters, and census blocs. For example, census blocs that were large relative to the size of minority clusters could make it difficult or impossible to create a minority-majority district, even intentionally.

Consider the 12th entry in the table, which describes a simulation run on a 20 by 20 grid where the minority political group populated 100 of four hundred census blocs in the state. Although the minority group makes up 25 percent of the voting population of the state, it loses nearly every election simply because it has the misfortune not to be geographically concentrated. The only way for minority political groups to win any seats in these circumstances would be if we tailored districts expressly to their boundaries, linking small concentrated clusters or minorities. While districts created in this way will almost certainly be noncompact, for such purposes *non*compactness will be a necessary condition, but almost never a sufficient one. Compactness makes it impossible for dispersed minority political groups to gain representation. I will return to this issue in Section 4.3.2.

These results may understate the electoral effects of compactness on minority representation when we consider the assumptions we made in the simulation about turnout. In the simulations, we assumed that each political group turned out to vote at the same rate and voted strictly for their own party. This is a simplification that helped to reveal the general dynamics of voting in compact districts, but which may bias our predictions. In particular, if the minorities that are geographically dispersed also turn out at a lower rate than the majority political group, or if they have a higher rate of cross-over voting, it will be even more difficult to draw compact districts that would allow minorities an opportunity to elect a candidate of choice.

In contrast, compactness helps combat majority bias when minority political groups are geographically concentrated. As Figure 4-5 shows, there is a strong positive correlation between a plan's compactness and the number of seats captured by such a political group. Why do we see such a correlation? The explanation for this is straightforward: Fortuitously, when both districts and minorities are very compact, a concentrated minority will sometimes fall completely within the district lines of the "optimal" plan. By contrast, under any other circumstances short of a purposeful minority gerrymander, the majoritarian bias inherent in geographical redistricting makes minority controlled districts extremely unlikely (Figure 4-5).

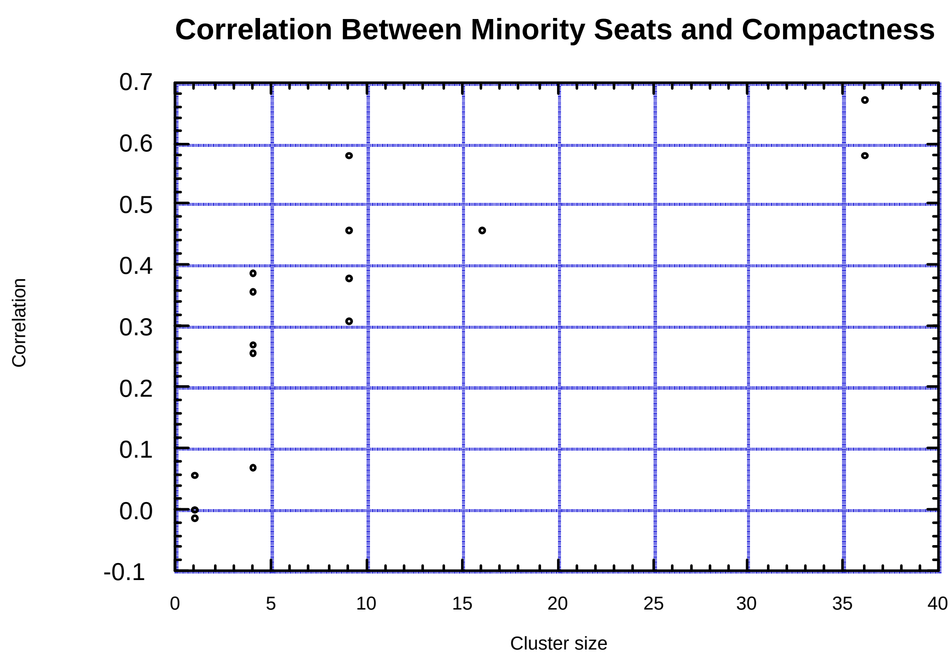

Figure 4-5. Correlation between minority seats and perimeter compactness (based on data in Table 4-2).

This phenomenon is not isolated to perimeter compactness, to small numbers of districts, or to compact clusters. Table 4-3 shows us the same patterns when we use the moment-of-inertia measure for compactness. Table 4-4 shows us somewhat weaker patterns when minorities are grouped in less compact clusters and into more districts.

Grid Size	Number of Districts	Cluster Size	Number of Clusters	Minority Percentage of Population	Mean Minority Controlled Districts (std. dev.)	Correlation Between Minority Controlled Districts and Compactness (Somers's d)
8x8	4	1	26	40%	0.47 (0.53)	-0.01 (-0.01)
8x8	4	4	6	38%	0.48 (0.55)	0.14 (0.15)
8x8	4	9	3	42%	1.2 (0.73)	0.52 (0.52)
8x8	4	16	1	25%	0.41 (0.49)	0.53 (0.57)
8x10	8	9	3	34%	1.38 (0.93)	0.62 (0.55)
20x20	8	4	40	40%	0.76 (0.72)	0.15 (0.12)
20x20	8	9	18	41%	1.46 (0.86)	0.25 (0.19)

Table 4-3. The effects of moment-of-inertia compactness on the representation of clustered minorities.

(10,000 Samples were performed for each grid/district combination.)[143]

[143] Hill-climbing and descent methods were used for these cases. Since the results from each method were similar, only hill-climbing results are reported.

	Grid Size	*Number of Districts*	*Minority Percentage of Population*	*Mean Minority Controlled Districts (std dev)*	*Correlation Between Minority Controlled Districts and Compactness (Somers's d)*
Perimeter	8x10	4	34	0.33 (0.48)	0.10 (0.12)
	20x20	16	10	0 (0)	0 (0)
	20x20	16	25	0.08 (0.28)	0.08 (0.16)
	20x20	16	40	2.88 (1.14)	0.25 (0.19)
Moment	8x10	4	34	0.37 (0.49)	0.01 (0)
	20x20	16	10	0 (0)	0 (0)
	20x20	16	25	0.08 (0.28)	0.04 (0.06)
	20x20	16	40	2.87 (1.11)	0.13 (0.09)

Table 4-4. The effects of compactness on the representation of Schelling-distributed minorities.

(10,000 Samples were performed for each grid/district combination.)[144]

Compact Partisan Gerrymandering

The previous section examines the effects of compactness when the creation of district plans is arbitrary, in one common sense of the word. Arbitrary district plans, although a distinct possibility in the future, have yet to come into wide practice. On the other hand, many authors past and present claim that partisan gerrymandering is widespread (Congressional Quarterly Staff 1993; Griffith 1974). Furthermore, although there are other types of gerrymandering (incumbent gerrymandering, for example), partisan gerrymandering is likely to become more pervasive as term limits and spending

[144] Hill-climbing and descent methods were used for these cases. Since the results from each method were similar, only hill-climbing results are reported.

limits weaken the individual incumbent relative the political party. What are the effects of compactness when it is applied to partisan gerrymanders?

Obviously, if one party is in complete control of drawing districts all of the time and in all places, its ability to create districts will be limited by practically any restriction, compactness included. Suppose, however, that different parties substantially control the redistricting process at different times and different places. If the courts continue to use compactness as a red flag to mark plans for judicial review, the party in control may try to produce the most compact partisan gerrymander that they can. How will compactness affect electoral results in these circumstances?

Table 4-5 and Table 4-6 compare the relationship between compactness and seats for the minority party when compact districts are arbitrarily selected to the case where partisans try to gerrymander compactly. In this second case, I altered the simulations to find, for each party, the most compact plan subject to the constraint of partisan seat maximization.[145]

---

[145] Gerrymandering to capture the maximum number of seats is only one possible partisan objective. In the real world, where there is uncertainty over voters' behavior, partisans might try to maximize the probability of controlling the legislature instead. I use seat maximization in this simulation because I have assumed certainty, and because for some minority population distributions, it may be impossible a-priori to capture the legislature.

The effects in these tables are somewhat more complicated than in Table 4-3 and Table 4-4. Here we see two major effects. First, we see the effect that we would expect to see, given these tables — if the minority party is populous and compact enough, the minority party benefits from a compactness rule: They will be able to produce maximal gerrymanders that are more compact, on average, than the maximal gerrymanders for the majority party, given the same population distribution. (See rows 2–4, 9, and 12–14 in Table 4-5.)

	Grid Size	*Number of Districts*	*Cluster Size*	*Number of Clusters*	*Minority Percentage of Population*	*ARBITRARY PLANS Correlation Between Compactness and Minority Controlled Districts (Somers's d)*	*COMPACT GERRY - MANDER Correlation Between Compactness and Minority Controlled Districts (Somers's d)*
Perimeter	8x8	4	1	26	40%	0.01 (-0.02)	-0.20 (-0.19)
	8x8	4	9	3	42%	0.46 (0.47)	0.23 (0.22)
	8x8	4	16	1	25%	0.46 (0.61)	0.18 (0.22)
	8x10	8	9	3	34%	0.58 (0.52)	0.19 (0.19)
	20x20	8	1	100	25%	0.0 (0.0)	0.00 (0.00)
	20x20	8	4	40	40%	0.29 (0.23)	-0.23 (-0.22)
	20x20	8	9	18	41%	0.43 (0.33)	-0.02 (-0.02)
	20x20	8	36	2	18%	0.66 (0.65)	-0.15 (-0.27)
	20x20	8	36	4	36%	0.58 (0.52)	0.06 (0.06)
Moment of Inertia	8x8	4	1	26	40%	-0.01 (-0.01)	-0.58 (-0.65)
	8x8	4	4	6	38%	0.14 (0.15)	-0.45 (-0.44)
	8x8	4	9	3	42%	0.52 (0.52)	0.14 (0.16)
	8x8	4	16	1	25%	0.53 (0.57)	0.09 (0.12)
	8x10	8	9	3	34%	0.62 (0.55)	0.27 (0.24)
	20x20	8	4	40	40%	0.15 (0.12)	-0.63 (-0.64)
	20x20	8	9	18	41%	0.25 (0.19)	-0.56 (-0.55)

Table 4-5. The effects of compact gerrymanders on clustered minorities.[146]

(500 samples were for each grid/district combination.)

---

[146] Hill-climbing was used for these cases.

	Grid Size	*Number of Districts*	*Minority Percentage of Population*	*ARBITRARY PLANS Correlation Between Compactness and Minority Controlled Districts (Somers's d)*	*COMPACT GERRYMANDER Correlation Between Compactness and Minority Controlled Districts (Somers's d)*
Perimeter	8x10	4	34	0.10 (0.12)	-0.16 (-0.13)
	20x20	16	10	0 (0)	0.00 (0.00)
	20x20	16	25	0.08 (0.16)	0.09 (0.08)
	20x20	16	40	0.25 (0.19)	0.18 (0.16)
Moment of Inertia	8x10	4	34	0.01 (0)	-0.47 (-0.46)
	20x20	16	10	0 (0)	0.00 (0.00)
	20x20	16	25	0.04 (0.06)	-0.51 (-0.50)
	20x20	16	40	0.13 (0.09)	-0.73 (-0.67)

Table 4-6. The effects of compact gerrymanders on Schelling-distributed minorities.

(500 samples were for each grid/district combination.)[147]

On the other hand, remember that if the minority party is weak or dispersed, compactness did not help them very much when districts were created automatically. As I noted in Section 0, the minority would need noncompact districts to capture seats. Hence we see in these cases that compactness harms the minority party: The majority party will be able to produce maximal gerrymanders that look much better than the gerrymanders produced by the minority party.[148]

---

[147] Hill-climbing was used for these cases.

[148] Note that compactness can obviously have no effect where the minority party population is small enough and scattered enough that they cannot capture any districts,

4.4. Discussion

The simulation shows us features of compactness standards. First, compactness effects are nonlinear. Electoral manipulation is much more severely constrained by high compactness than by moderate compactness. Any empirical study of the relationship between gerrymandering and compactness must use models that can accommodate these nonlinearities.

Second, compactness effects are context-dependent. The difficulty of drawing compact plans is significantly affected by the shape of the state being divided, as well as by the compactness measure used. Similarly, differences in population geography may affect the difficulty of drawing compact, equal-population plans. Therefore, comparisons of compactness between states are misleading: statistical studies of the electoral effects of compactness should take this into account – for example, using time-series rather than comparing compactness across states.

Third, compactness standards can have asymmetric effects on different political groups if those groups are distributed in geographically different ways. The geography of district lines alone is not sufficient to diagnose a gerrymander: A majority which is purposefully attempting to exclude a large but geographically diffuse political minority from the political arena will want to draw districts that are as compact as possible —

---

even if given substantial opportunity to gerrymander. (See row 5 in Table 4-5 and rows 2 and 6 in Table 4-6.)

whereas the same majority, with the same purpose, facing a geographically concentrated minority, will want to draw noncompact districts. To correctly predict the electoral effects of a set of district lines, you must know the geographical distribution of all the relevant political groups. Geography does matter — but you must interpret it within a political context.

Fourth, this study shows that compactness can also significantly disadvantage geographically *concentrated* minorities, in the context of a partisan gerrymander. We might expect compactness to have the greatest political effect on racial and ethnic minorities, and on the parts of the Democratic Party that they disproportionately support, because, these minorities are disproportionately concentrated in large cities.[149]

The electoral effects of compactness, however, will not be limited to these groups; the simulations results apply in general to any large minority group that is politically cohesive and geographically concentrated – including any "community of interest" that has these characteristics. The argument that "communities of interest" that are *not* compact are likely to be hurt by compactness standards is fairly straightforward. Minorities that are geographically dispersed are already at a disadvantage in the redistricting process, especially where the building blocks for districts are large, and

---

[149] Only 19% of the white population resides in cities of 100,000 or more, while 57% of African-Americans, 56% of Hispanics and 53% of Asian-American's reside in these dense urban areas (G.P.O. 1993).

compactness is likely to make this disadvantage even greater. What has not previously been noted, however, and what these simulations show, is that even “communities of interest” that are compact can be systematically damaged by compactness standards, in some institutional environments. This contradicts the optimistic assumption that some proponents of compactness have made (Polsby and Popper 1993), that compact communities are likely to benefit from compactness standards.

(Allegations that compactness requirements harm dispersed minorities are often based on the implicit assumption that in the absence of a compactness requirement, districts will be drawn so as to maximize the total number of minority districts. If this is the case, any requirements on districting at all will tend to reduce minority seats, as I discuss in Section 4.3.)

The Court has declared that optimal compactness is not required -- redistricting is not a “beauty contest.” If compactness is to be used at all, however, such beauty contests are going to be difficult to avoid, for several reasons: The extent to which compactness is likely to vary dramatically with the particular level of compactness that is considered “good enough.” Moreover, to set this threshold appropriately, the Court would have to consider separately the particulars of each measure of compactness, and the geography of each state. The only practical method of implementing effective compactness measures are likely to require the beauty contests that the court wishes to avoid.

Such beauty contests would have little to recommend them. Compactness standards, rather than being the neutral standard that the court envisions, are likely to have distinctly partisan effects. These simulation results contradicts the view of compactness advocates and bears out Lowenstein's (1985) assertion that compactness is not a partisan-neutral standard because of the way that Democrats are concentrated geographically.

These simulations also raise questions for the use of ill-compactness as evidence of political gerrymandering. Parties with distributed support are likely to able to draw districts to their advantage without violating triggering any alarms.

The effects of compactness rules will also depend upon an important factor that the Court has largely ignored in recent cases -- the larger institutional mechanism for drawing districts. Compactness rules will have different effects — rules that benefit geographically concentrated minority parties under a system of automated redistricting may have different effects under a system dominated by partisan gerrymandering. The apparent mathematical universality of compactness rules is, in fact, illusory — to understand the political effects of compactness requires an intensely local appraisal.

Appendix: Computer Techniques

The optimization routines used in this chapter were written in the *C* language, and run on a variety of Unix workstations. Although, the *C* implementations were, for the most part, specific to this chapter, they are based on a number of publicly documented algorithms. This appendix details the Monte Carlo, genetic, simulated annealing, and genetic algorithm used in the chapter.

The Monte Carlo procedure I used was quite simple — I create a district by adding population units at random, until the districts are approximately the correct size. Hill climbing can build on this random assignment.

My hill climbing procedures used the “greedy” method described in (Nagel 1965): each round the program examines every possible combination of single shifts (moving a census block from one district to another) and single trades (i.e., a pair of compensating shifts between two districts). After examining all possibilities, the trade that most improved the plan was executed, and the process repeated. The program stopped when it could find no further improvement to make. Reverse hill climbing operated similarly, except that first changes were made to decrease fitness of the maximally compact plan, until the program reached the target level of non-compactness. And second, to avoid deterministically reusing the same set of non-compact plans, and the fitness decreasing changes were chosen randomly rather than greedily.

The genetic algorithm optimizer was based upon the algorithms described in Goldberg (1989). Following this terminology, each district was encoded as a haploid string with separate positional information (i.e., a string of {district assignment, census block} pairs). Each population (population=750) of strings was subjected to mutation (p=.001), inversion (p=.01), and pmx-crossover (p=.05), as described in Goldberg. Strings were then chosen for the next generation by repeated random drawing (with replacement) of pairs of strings, the fitter string succeeding to the next generation. This process was repeated for one hundred generations.

Simulated annealing code came directly from Lester Ingber's Adaptive Simulated Annealing (ASA) package (v. 2.2), as described in (Ingber 1989). I started by representing each census block as an integer, in the ASA framework, which was assigned a district number by the annealing process. In view of the poor performance of the annealing algorithm, I then modified the code slightly, to create a "combinatoric" type to represent each census block. The only difference between the "integer" and "combinatoric" types was that for combinatoric types only the probability of the variable being changed lowered with the annealing temperature; whereas for integer types, the probablilty and the magnitude of the change lowered. The logic behind the change was that since district numbers are arbitrary, changing a census bloc from district 1 to district 2 is not necessarily a "smaller" change, in any relevant measure, than shifting a block from district 1 to district 10. These results are reported in the chapter, and were only slightly better than the original results using integers.

The simulation programs in this chapter used uniform and normal random numbers. Simulations are often susceptible to quirks in random number generators, and the random number generators in many standard *C* libraries are poorly written and non-portable across platforms. Because of this, I used the portable *Ranlib* package written by Barry W. Brown, and James Lovato at the Department of Biomathematics, The University of Texas, M.D. Anderson Cancer Center. The algorithms that this package uses to generate uniform and normal random number routines are described in (L'Ecuyer and Cote 1991) and (Ahrens and Dieter 1973).