

Optimal Gerrymandering in a Competitive Environment

John N. Friedman and Richard T. Holden*

December 9, 2008

Abstract

We analyze a model of optimal gerrymandering where two parties receive a noisy signal about voter preferences from a continuous distribution and simultaneously design districts in different states and in which the median voter in a district determines the winner. The form of the optimal gerrymander involves “slices” of extreme right-wing voters that are paired with “slices” of left-wing voters, as in Friedman and Holden (2008). We also show that, as one party controls the redistricting process in more states, that party designs districts so as to spread out the distribution of district median voters from a given state. (*JEL D72, H10, K00*).

*Friedman: University of California at Berkeley, Evans Hall, 5th floor, CA 94720. email: jfriedman@post.harvard.edu. Holden: Massachusetts Institute of Technology Sloan School of Management, E52-410, 50 Memorial Drive, Cambridge MA 02142. email: rholden@mit.edu. Friedman acknowledges the support of the RWJ Scholars in Health Policy Program.

1 Introduction

A growing literature analyzes *gerrymandering*, the process by which politicians draw the boundaries of their own electoral districts. To simplify the analysis, however, most papers have focused on the simplest case—that in which one party controls the redistricting of one state (Owen and Groffman 1988, Gilligan and Matsusaka 1999, Friedman and Holden 2008). In practice, of course, Republicans and Democrats each control the districting process in a number of states. Thus the environment is best represented as a two-player game rather than a control problem. A key feature of this game is the number of states which are controlled by a given party. For example, in 2002 the G.O.P. gained control of redistricting in: Florida, Idaho, Kansas, Michigan, Pennsylvania and Texas. This gave them a net gain of 95 districts in which they controlled the redistricting process¹. The Democratic party had a net gain of just 1 district². How do shifts like this in control of redistricting affect equilibrium strategies? This is the question we address in this paper.

We build on our work in Friedman and Holden (2008) to provide a treatment of the districting game in an environment where the median voter in a district is decisive. The analysis that follows has two parts. First, we extend the analysis of Friedman and Holden (2008) to a multi-state, multi-party environment. The key result from this analysis is that the form of the optimal gerrymander in Friedman and Holden (2008) is the same in the richer environment. Specifically, when signals are sufficiently precise, the party in control forms districts by matching a group of right-wing voters with a group of left-wing voters, with these “slices” of voters becoming progressively less extreme as the district becomes less favorable to the redistricting party.

Having established the basic form of the optimal gerrymander in the Friedman-Holden framework, we then analyze a more abstract model where players simply control district

¹They lost control of New Hampshire and its 2 districts.

²See Friedman and Holden (2009) for a detailed breakdown. We treat CA as being previously controlled by the Democrats since they had partisan control in 1972 and court imposed plans modified this in 1982 and 1992 before partisan control by Democrats in 2002.

medians subject to constraints. This is a model of the districting game at what is arguably its most general level. We compute comparative statics on optimal district formation with respect to key parameters of the redistricting game. Most importantly, we show that as one party controls the redistricting in more states, that party creates districts that are less homogenous within a state. Viewed from the specific model of redistricting in Section 2, this implies that a greater number of right-wing voters are matched with more left-wing voters when the party controls more states. This increases the effective representation of extreme supporters of both parties.

The work most closely related to ours is an elegant paper by Gul and Pesendorfer (2008). They characterize the set of equilibria using an ingenious argument which restates the game as a control problem involving maximization of the number of seats won at cutoff values of an aggregate shock to voter preferences. This also allows them to provide the important comparative static on the consequences for the optimal gerrymander as the number of states districted by a particular party changes. One simplifying assumption which Gul and Pesendorfer (2008) make is that there are only two types of voters. In a single state model it is known that the familiar pack-and-crack strategy obtains with only two voter types, but not with more types (Friedman and Holden 2008, Proposition 3). By contrast, our first result on the matching slices strategy contrasts with this. Despite the additional complexity which the assumption of a continuum of voter types brings with it, we are also able to analyze the impact of more general objective functions that the simple majoritarian function Gul and Pesendorfer analyze. We make use of certain useful mathematical results on *Pólya frequency functions* to perform this analysis.

The remainder of the paper is organized as follows. Section 2 shows that the matching slices strategy of Friedman and Holden (2008) obtains in the redistricting game. Section 3 considers a general model of competitive redistricting and shows how the optimal strategy changes as the proportion of districts controlled by one party changes. Section 4 contains some concluding remarks.

2 The Optimal Gerrymander

In this section we extend the model in Friedman and Holden (2008) to include two parties and many states.

2.1 The Model

There are two parties D and R . In a given state s , there a unit mass of voters with heterogenous preference points β , but the parties apportioning voters only observe a noisy signal of this parameter, denoted σ .³ We denote the joint distribution as $f_s(\beta, \sigma)$ and the posterior conditional distribution, or the “conditional preference distribution,” $g_s(\sigma|\beta)$. The marginal distribution of signals in the state s , or the “signal distribution,” is denoted $h_s(\sigma) = \int f_s(\sigma, \beta) d\beta$. Let μ_{ns} denote the median voter in district n in state s , and $\psi_{ns}(\sigma)$ denote the mass of voters placed in such a district by the gerrymanderer. Suppose in state s there are N_s districts. There are two constraints on the formation of districts. First, all districts in state s must contain the same mass of voters $\frac{1}{N_s}$. Second, each voter in state s must appear in exactly one district in state s . Aggregate uncertainty in state s occurs with distribution function B , so that the probability the Republicans win district n in state s is $B(\mu_{ns})$.

We assume that each party redistricts some states. To do so, we assume that there are S states comprising a total of N districts. Suppose that party D creates the districts in states $1 \leq s \leq S_D$, party R does the same in states $S_D < s \leq S_R$, and states $s > S_R$ are redistricted exogenously to the model; this could represent bipartisan gerrymandering (in which no single party controls the organs of redistricting in a state) or court-mandated apportionment. Each party p has value function $W_p : [0, 1] \rightarrow \mathbb{R}$, whose domain is the fraction of seats (districts) won in the election. We assume that each W_p is strictly increasing, and that parties maximize expected payoffs.

³Note that one can model this reduced form “bliss point” approach as the implication of an assumption that voters have preferences over policy outcomes that satisfy “single-crossing” and that all candidates from a given party in a given state share a policy position. See Friedman and Holden (2008) for this treatment.

We assume that parties move simultaneously. This assumption matches the reality that 49 states must (by state law) redistrict within a window of about six months, after the release of the preliminary census but in time to organize the next Congressional elections. Furthermore, redistricting is typically a long and involved process, so that states cannot afford to wait for other states to complete their redistricting process. We focus on Nash equilibria of this game. The choice variables of each party are the signal distributions in their districts; thus, the party R may choose $\{\psi_{ns}(\sigma)\}_{s=S_D+1, n=1}^{s=S_D+S_R, n=N_s}$ while the democrats may choose $\{\psi_{ns}(\sigma)\}_{s=1, n=1}^{s=S_D, n=N_s}$.

Formally, party D faces the problem

$$\begin{aligned} & \max_{\{\psi_{ns}(\sigma)\}_{s=1, n=1}^{s=S_D, n=N_s}} EW_D \left(\frac{1}{N} \left(\sum_{s=1}^{S_D} \sum_{n=1}^{N_s} d_{ns} + \sum_{s=S_D+1}^{S_R} \sum_{n=1}^{N_s} d_{ns} + \sum_{s=S_R+1}^S \sum_{n=1}^{N_s} d_{ns} \right) \right) \\ \text{s.t.} \quad & \int_{-\infty}^{\infty} \psi_{ns}(\sigma) d\sigma = \frac{P_s}{N_s} \quad \forall n, s \\ & \sum_{n=1}^{N_s} \psi_{ns}(\sigma) = h_s(\sigma) \quad \forall \sigma, s \\ & 0 \leq \psi_{ns}(\sigma) \leq h_s(\sigma) \quad \forall n, \sigma, s. \end{aligned}$$

and party D solves a parallel problem where d_{ns} is a dummy variable equal to one if party D wins the election in district n in state s . Party R faces a similar problem but with different state districting schemes as the choice variables.

We now make two assumptions about the nature of the relationship between the signal σ and the true voter preference β . First, we require that the signal σ is informative about the underlying preference β , in a specific sense.

Condition 1 (Informative Signal Property) *Let $\frac{\partial G_s(\beta|\sigma)}{\partial \sigma} = z_s(\beta|\sigma)$. Then*

$$\frac{z_s(\beta|\sigma')}{z_s(\beta|\sigma)} < \frac{z_s(\beta'|\sigma')}{z_s(\beta'|\sigma)}, \quad \forall \sigma' > \sigma, \beta' > \beta, s$$

This property is similar to the Monotone Likelihood Ratio Property, and if a the signal shifts only the mean of the conditional preference distribution, then this property is equivalent to MRLP.⁴

Second, we require a form of unimodality.

Condition 2 (Central Unimodality) *For all s , $g_s(\beta | \sigma)$ is a unimodal distribution where the mode lies at the median.*

Note that, without loss of generality (given Condition 1), we can “rescale” σ such that $\sigma = \max_\sigma g_s(\beta | \sigma)$. The two parts of Condition 2 essentially require that β is distributed “near” σ , and not elsewhere.⁵

2.2 The Form of the Optimal Gerrymander

We can now state the first of two main results of this section.

Proposition 1 *Suppose that Conditions 1 and 2 hold. Then for a sufficiently precise signal the optimal districting plan in any equilibrium, for each party p , in each state s , can be characterized by breakpoints $\{u_{ns}\}_{n=1}^{N_s}$ and $\{l_{ns}\}_{n=1}^{N_s}$ (ordered such that $u_{1s} > u_{2s} > \dots > u_{N_s-1,s} > l_{N_s-1,s} > \dots > l_1 \geq -\infty$) such that*

$$\begin{aligned} \psi_{1s} &= \begin{cases} h(\sigma) & \text{if } \sigma < l_1 \text{ or } \sigma > u_1 \\ 0 & \text{otherwise} \end{cases}, \\ \psi_{ns} &= \begin{cases} h(\sigma) & \text{if } l_{n-1} < \sigma < l_n \text{ or } u_{n-1} > \sigma > u_n \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 1 < n < N, \\ \text{and } \psi_{N_s,s} &= \begin{cases} h(\sigma) & \text{if } \sigma > l_{N-1} \text{ or } \sigma < u_{N-1} \\ 0 & \text{otherwise} \end{cases}. \end{aligned}$$

⁴See footnote 11 of Friedman and Holden (2008) for a simple proof of this.

⁵For a more detailed discussion on this property, see Friedman and Holden (2008).

This result establishes that “cracking” is not optimal, so that parties find it optimal to group the most partisan voters into one district within a given state. Parties still may wish to “pack” those least favorable voters into segregated districts, though. We now provide conditions under which packing too is not optimal.

Proposition 2 *Suppose that Conditions 1 and 2 hold, and the signal is of sufficiently high quality. Then in any set of equilibrium redistricting strategies, there exists n and $\sigma < \sigma'$ such that $\mu_{ns} > \mu_{N_s s}$ and $\sigma \in \psi_{ns}$, $\sigma' \in \psi_{N_s s}$ for all n, s .*

Thus, in Proposition 2, we rule out the possibility of “packing” as well. We refer to this strategy, in its purest form (as in Proposition 2), as a “matching slices” strategy, since the parties find it optimal to match slices together from extreme ends of the signal distribution, working in to the middle of the distribution. Figure 1 is an example of a strategy (in a single state with five districts) that satisfies the conditions in Propositions 1 and 2.

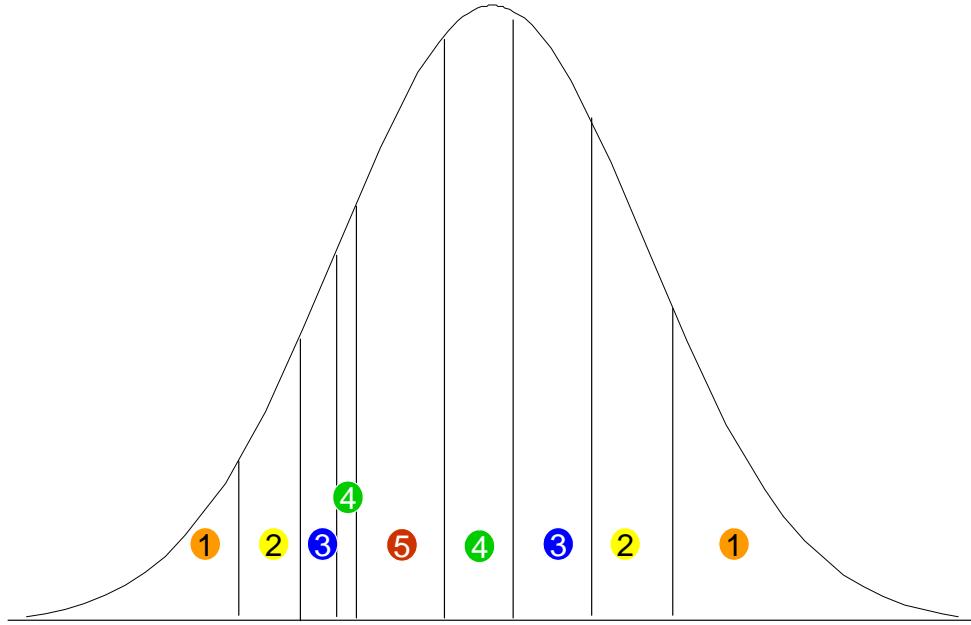


Figure 1: The form of the optimal gerrymander

These results extend those in Friedman and Holden (2008) to the richer setting in which parties do not control all districts, but instead control only a fraction of the relevant districts.

Furthermore, parties must apportion voters within preexisting states, which further limits their flexibility. To understand intuitively why the original results extend to this broader case, consider the gain to party D from winning a given district, as opposed to losing it. If the value function is non-linear, this value may change greatly depending on the party R 's districting plan for their states, the nature of the exogenous redistricting, or the set of states controlled by the party R . But holding all else fixed—which is precisely what happens at Nash equilibrium—an increase in the probability of winning the given district increases the value function *linearly*. Thus, the trade-offs between districts in this more complicated model differ only from those in a simpler model (in which a party maximizes the sum of the probabilities of winning districts, or the expected number of districts won) by constant terms. A party may adjust by altering the number of right-wingers in the upper “slice” of each district in a given state, but the fundamental characterization of the optimal strategy, as described in Propositions 1 and 2, remains the same.

3 A Generalized Model

The above model is useful for characterizing equilibrium strategies, but it can be abstracted from to some degree when considering comparative statics. At the most basic level, each party constructs districts so as to choose median voters in those districts, subject to constraints given by the primitives of the problem.

3.1 The National Model and Party Preferences

Denote the median voter in a district n in state s by μ_{ns} . Let the feasible set of medians for player R be Ω_R . There are two types of shocks: national and district-specific, denoted ϕ and ψ respectively, with cdfs B and C . These shocks are mean zero and symmetrically

distributed. Party R 's vote-share in district n in state s is

$$V_{ns} = V(\mu_{ns} + \psi_{ns} + \phi)$$

where V is a strictly increasing function and $V(0) = \frac{1}{2}$. Party R wins such a district if and only if $V_{ns} > \frac{1}{2}$. Thus, the probability of winning such a district, conditional on the aggregate shock, is

$$\Pr(\text{win}) = C(\mu + \phi).$$

There are two parties, R and D . Party R may design the districts in fraction λ of the continuum of states⁶, and party D controls redistricting in the other $1 - \lambda$. Thus, suppose that (after redistrictings) the distribution of median voters in the population is $M(\mu)$. Then party R wins

$$X(\phi) = \int C(\mu + \phi) M(\mu) d\mu$$

districts. The party values the fraction of seats won by the function $W(\cdot)$, which is a weakly increasing function.

3.2 A Special Case

Suppose that we make the strong assumption that C is a uniform distribution, so that $c(\cdot) = k$ for some constant. Then we can rewrite the expected number of seats won by party R as

$$\begin{aligned} X(\phi) &= k \int (\mu + \phi) M(\mu) d\mu \\ &= k(\lambda \bar{\mu}_R + (1 - \lambda) \bar{\mu}_D + \phi). \end{aligned}$$

⁶The assumption that there are a continuum of states is important in that it allows us, by the law of large numbers, to treat C as the proportion of districts won, not just the probability of so doing. We view this, however, as a technical assumption.

and the expected value function for party R as

$$EW = \int W(k\lambda\bar{\mu}_R + (1 - \lambda)\bar{\mu}_D + \phi) b(\phi) d\phi.$$

But here it is clear that each party does best simply to maximize the average of the median voters in the districts in their control. Note that there is no strategic interaction at all between the parties in this special case—or at least the game is dominance solvable. As a result, the share of states λ under the control of party R can have no impact on the optimal gerrymander.

3.3 The General Case

By assumptions made above, all states are the same, so the optimal gerrymander will be identical across them. Denote by $\{\mu_{dR}\}$ and $\{\mu_{dD}\}$ the medians of the districts in states controlled by party R or D , respectively, and denote by N_R and N_D the total number of districts controlled by parties R and D . We then have

$$X(\{\mu_{nR}\}, \{\mu_{nD}\}; \phi) = \lambda \left[\sum_{n=1}^{N_R} C(\mu_{nR} + \phi) \right] + (1 - \lambda) \left[\sum_{n=1}^{N_D} C(\mu_{nD} + \phi) \right].$$

Parties then maximize their value function, as weighted by the different possible outcomes of the aggregate shock ϕ .

We can restate the problem that the parties solve. Proposition 3 shows that the parties act as though they maximize vote shares over a weighted average of different situations.

Proposition 3 *Suppose that $W(x)$ is continuous. As a function of the strategies chosen by the parties $\{\mu_{nD}\}$ and $\{\mu_{nR}\}$, define $\phi^*(\{\mu_{nD}\}, \{\mu_{nR}\}, x)$ such that*

$$\lambda \sum_{n \in R} C(\mu_n + \phi^*) + (1 - \lambda) \sum_{n \in D} C(\mu_n + \phi^*) = x.$$

Then the optimal gerrymander $\{\mu_{nR}\}$ will satisfy the necessary conditions to the problem

$$(1) \quad \max_{\{\mu_{nR}\}} \int \left[W'(x) b(\phi^*(x)) \sum C(\mu_n + \phi^*(\{\mu_{nD}^*\}, \{\mu_{nR}^*\}, x)) \right] dx$$

such that $\{\mu_{nR}\} \in \Omega_R$.

Each party thus acts as though it maximizes the number of seats won across a weighted average of values of ϕ (which does not depend on the strategies).

Note that the alternative maximization above does not involve anything about the districts designed by the opponent party D , conditional on $\phi^*(x)$. If we can specify the set of $\phi^*(x)$ values, then Proposition 3 allows us to restate the maximization problem in a way that does not involve the other party's choices. This simplifies the analysis greatly. Of course, the optimal sets of district medians $\{\mu_{nD}^*\}$ and $\{\mu_{nR}^*\}$ and the set of $\phi^*(x)$ values are jointly determined. But if we can identify variables that shift the $\phi^*(x)$ values, for instance, then we can trace through the implications for the optimal district medians.

This result is a generalized version of Theorem 1 in Gul and Pesendorfer (2007), we focus on the case where parties care only about winning a majority in the legislature. The following Corollary links our result above to theirs.

Corollary 1 Suppose that the party's value function over seats won is

$$W(x) = \begin{cases} 1 & x > \frac{1}{2} \\ \frac{1}{2} & x = \frac{1}{2} \\ 0 & x < \frac{1}{2} \end{cases}.$$

Then

$$\{\mu_{nR}^*\} = \arg \max_{\{\mu_{nR}\}} \sum C\left(\mu_n + \phi^*\left(\frac{1}{2}\right)\right)$$

so that parties simply maximize the share of seats won at one specific value of the aggregate shock, which is the “pivotal value.”

These two results are, at some level, quite intuitive. If, for instance, a party controls very few states, then it must turn out to be an extremely favorable state of the world in order for it to win. And in such a setup, it is natural for the party to simply assume that it receives such a shock when redistricting.

But these results are also far more precise than the preceding intuition might suggest. Suppose, for instance, that two parties control the same number of states, and so the aggregate shock must simply be above average for a given party to win. Corollary 1 shows that parties do not maximize over all such winning values of the shock; rather, they do so only with respect to the one pivotal value at which the parties are evenly matched.

3.4 Comparative Statics

We wish to know whether parties that control the redistricting in more states will act differently in equilibrium than a party that controls fewer states. Given Proposition 3, we can rephrase the problem faced by each as maximizing the vote-share conditional on the pivotal value ϕ^* . Of course, this requires knowing ϕ^* , which is jointly determined with the optimal strategy. But we do know that, as λ increases, ϕ^* will decrease, since a party needs less luck from the aggregate shock because it has more districts gerrymandered to its advantage. Thus, we can solve for the comparative static of λ by solving for the comparative static on ϕ^* . And even though one cannot actual solve for ϕ^* , we can simply assume a value of ϕ^* and then see what changing that value does to the districting scheme.

3.4.1 Two Districts per State

Suppose that there are only two districts per state. Proposition 4 then shows how parties alter their optimal redistricting strategies when they control more states.

Proposition 4 *Assume that*

$$W(x) = \begin{cases} 1 & x > \frac{1}{2} \\ \frac{1}{2} & x = \frac{1}{2} \\ 0 & x < \frac{1}{2} \end{cases},$$

that there are only 2 districts per state and that c is log-concave. Then as λ increases, and party R controls more districts, μ_1^ increases, so that party R includes a larger upper slice in the first district. This increases the disparity between μ_1 and μ_2 .*

The intuition behind this result stems from the fact that parties optimize their districts relative to the marginal value of the aggregate shock. If parties control an equal number of states, then the aggregate must be better than average for that party to win. In this case, both the favorable and unfavorable districts may be in play, since the local shock necessary to tip a district to one party or another is not so big. But if a party controls many states, then the aggregate shock will have to be quite negative for the party to lose the election. And in such a situation, the trade-off between favorable districts and unfavorable districts is much different. Since the aggregate shock is so negative, unfavorable districts are now essentially unwinnable, and so increasing the median voter helps very little. Favorable districts are still very winnable, and so parties choose to increase the median voter in district 1 (in our two-district example) at the cost of lowering that in district 2.

This result implies that the control of redistricting matters crucially for the nature of representation in the legislature. There are two effects. First, parties redistrict so as to maximize their own representation, so more equal control of state districting has a straightforward effect on the balance of representation in the legislature. But Proposition 4 shows that there is another effect as parties change the way they draw districts in states they do control. As one party controls more states, it draws districts with median voters that are further from the overall median voter, thus increasing both the polarization of the legislature and the representation of extreme voters in the population.

3.4.2 Many Districts

When there are many districts per state and the objective function remains a simple step function, the intuition from the previous section still holds. When the share of states apportioned by R increases, there is a force pulling each pair of district medians. But suppose there are three districts; the middle median should move down relative to the upper median, but up relative to the lower median. These competing forces make the direction of movement for this middle median theoretically ambiguous. As before, the uppermost median must increase, while the lowermost median must decrease. Proposition 5 characterizes this situation more generally.

Proposition 5 *Suppose there are N districts per state. Assume that*

$$W(x) = \begin{cases} 1 & x > \frac{1}{2} \\ \frac{1}{2} & x = \frac{1}{2} \\ 0 & x < \frac{1}{2} \end{cases}$$

and that c is log-concave. Then as λ increases so that party R controls more districts, μ_1^ increases and μ_N^* decreases.*

Intuitively, these forces tend to stretch out the distribution of district medians within a given state, but the complexity of the dynamics prevents a more systematic characterization. Table 1 and Figure 2 present a numerical example that illustrates these forces. In this example, we suppose that there is mass one of identical states with five districts each. In each state, both the signal distribution $H(\sigma)$ and the conditional preference distribution $G(\beta | \sigma)$ are Normals with mean 0 and variance 2.5. We assume that $B(\phi)$ and $C(\psi)$ are Normal distributions with mean 0 and variance $\frac{1}{4}$. Each row of Table 1 presents the equilibrium strategy given a share of control λ . As we increase λ , the medians are pulled further apart. Both μ_1^* and μ_2^* increase monotonically in this example; μ_4^* and μ_5^* each decrease as λ increases. The middle median sometimes moves up and sometimes moves

down as party R 's state control increases. In this example, district 2 exhibits a monotonic increase and district 4 a monotonic decrease. We do not believe this to be a general result—all districts other than the top and bottom could move ambiguously—but we do not have a counterexample.

[Table 1 and Figure 2 about here]

3.4.3 Generalizing the Objective Function

The above results have focused on the case when the objective function is a step function with a single discontinuity. With a more complex objective function, each first order condition becomes the ratio

$$\frac{\int W'(x) b(\phi^*(x)) c(\mu_i + \phi^*(x)) dx}{\int W'(x) b(\phi^*(x)) c(\mu_j + \phi^*(x)) dx} = \lim_{\varepsilon \rightarrow 0} \frac{\Delta_j}{\Delta_i}.$$

Analyzing this expression is difficult in general. In order to sign a similar comparative static with respect to λ , this ratio must be weakly monotonic in λ . Intuitively, we need more than log-concavity of c ; instead, we need log-concavity in a weighted average of c . It is certainly not the case that this holds for all increasing functions W ; for instance, if the value function were a step function with two discontinuities, then the effective bimodality of $W'(x)$ can counterbalance the log-concavity of c . We can, however, provide a condition on the objective function under which Proposition 4 generalizes. This involves so-called *Pólya frequency functions* and hence a couple of definitions are in order before we state our result.

Definition 1 Let X and Y be subsets of \mathbb{R} and let $K : X \times Y \rightarrow \mathbb{R}$. We say that K is **totally positive** of order n (TP_n) if $x_1 < \dots < x_n$ and $y_1 < \dots < y_n$ imply:

$$\begin{vmatrix} K(x_1, y_1) & \dots & K(x_1, y_m) \\ \vdots & & \vdots \\ K(x_m, y_1) & \dots & K(x_m, y_m) \end{vmatrix} \geq 0$$

for each $m = 1, \dots, n$.

Total positivity has wide applications in economics. When K is a density TP_2 is equivalent to the monotone likelihood ratio property.

Definition 2 A *Pólya frequency function* of order n (PF_n) is a function of a single real argument $f(x)$ for which $K(x, y) = f(x - y)$ is TP_n , with $-\infty < x, y < \infty$.

Proposition 6 Suppose that there are N districts per state, that $W'(x)$ is PF_2 , that b is the uniform distribution and that c is log-concave. Then as λ increases, and party R controls more districts, μ_1^* increases and μ_N^* decreases.

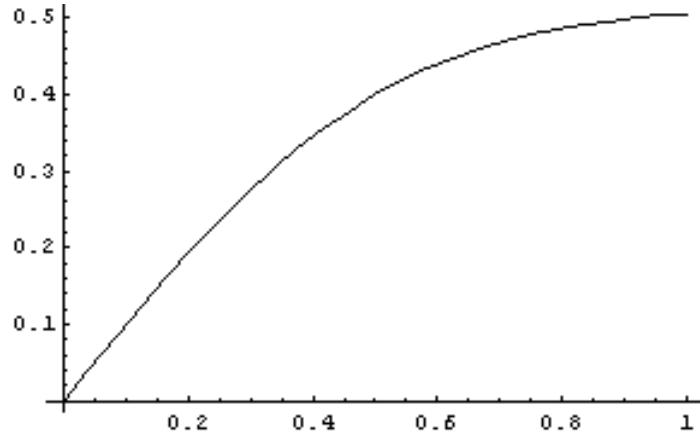
This proof of this result is closely related to the observation that the convolution of two log-concave densities is itself log-concave. $W'(x)$ is clearly not a density, but the appropriate generalization is that it must be PF_2 (a condition which log-concave densities satisfy).

A natural question to ask is what objective functions $W(x)$ have derivatives which satisfy this requirement, and how economically reasonable are they.

As the definition of PF_2 makes clear, there is a wide variety of functions in this class. Here we provide one simple example of interest. It is easy to verify that the function $W'(x) = e^{-\gamma x^2}$ is PF_2 (see Karlin 1968, p.30 for a proof). Integrating over x yields

$$W(x) = \frac{\sqrt{\pi} ERF(x\sqrt{\gamma})}{2\sqrt{\gamma}},$$

where ERF is the “error correction function” (i.e. the integral of the Normal distribution function). The following graph illustrates this function for $\gamma = 3$.



Example of an Objective Function Satisfying the
Conditions of Proposition 6

This objective function is clearly increasing and strictly concave in the number of districts won. It is easy to see that γ parameterizes the degree of concavity, with a higher value of γ corresponding to a more concave objective function. This is only one function in the class of functions satisfying the conditions of Proposition 6, but this demonstrates that there are plausible objective functions in that class.

Another function that satisfies Definition 2, of course, is the Normal cumulative distribution function.⁷ This class of functions have a great deal of intuitive appeal as a continuous legislative value function, since the marginal value of a seat is greatest at 50% and falling as one party has a larger and larger majority.

⁷Of course, $\Phi : \mathbb{R} \rightarrow [0, 1]$ while our $W : [0, 1] \rightarrow \mathbb{R}$. But the range restriction is unimportant, and since Definition 2 is unaffected by scale changes, we can define

$$W(x) = \frac{\Phi(x - \frac{1}{2})}{\Phi(1) - \Phi(0)}$$

for any Normal CDF $\Phi(x)$.

It is also informative to think about objective functions W that do *not* satisfy this definition. One simple example is the double-discontinuity function

$$W(x) = \begin{cases} 0 & x < \frac{1}{3} \\ \frac{1}{4} & x = \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} < x < \frac{2}{3} \\ \frac{3}{4} & x = \frac{2}{3} \\ 1 & x > \frac{2}{3} \end{cases}.$$

Assume that there are two districts per state and that ϕ^* is distributed uniformly, for simplicity. Suppose that C is log-concave. Then by Proposition 3, we know that the ratio

$$(2) \quad DDR = \frac{c(\mu_1 + \phi^*(\frac{1}{3})) + c(\mu_1 + \phi^*(\frac{2}{3}))}{c(\mu_2 + \phi^*(\frac{1}{3})) + c(\mu_2 + \phi^*(\frac{2}{3}))}$$

must be monotone in λ for the comparative static to hold. From log-concavity, we know that the ratio $\frac{c(\mu_1 + \phi^*(x))}{c(\mu_2 + \phi^*(x))}$ is increasing in λ for all x . But it will not generally be the case that the combined ratio in expression (2) is increasing. For instance, suppose that the following values hold for $\lambda_H > \lambda_L$.

	$\frac{c(\mu + \phi^*(\frac{1}{3}))}{c(\mu_2 + \phi^*(\frac{1}{3}))}$	$\frac{c(\mu + \phi^*(\frac{2}{3}))}{c(\mu_2 + \phi^*(\frac{2}{3}))}$	DDR
λ_H	$\frac{8}{2}$	$\frac{100}{100}$	$\frac{108}{102} \approx 1$
λ_L	$\frac{3}{1}$	$\frac{1}{2}$	$\frac{4}{3} > 1$

Intuitively, the double-discontinuity objective function is an extreme example of a function that is *not* PF_2 , since it is the limit of an extremely bimodal function. When “convoluted” with W' , C loses its log-concavity, and so a fall in ϕ^* is no longer enough to guarantee an increase in the value of the higher median.

4 Discussion and Conclusion

We have presented a model of competitive gerrymandering in which two parties control redistricting across many states. After confirming that the “matching slices” strategy from Friedman and Holden (2008) obtains in this richer setting, we showed that this redistricting game can be restated as a control problem, in the manner of Gul and Pesendorfer (2008). We then showed that an increase in the number of states controlled by a party in the redistricting game tends to spread out the distribution of optimal district medians. This shift increases the representation of extreme voters of both parties at the expense of moderates, especially those in the party gaining power.

These results bear on a number of broader topics in American politics. In recent years, Republicans have gained control of a number of key state legislatures, allowing them to design partisan gerrymanders in large states such as Pennsylvania, Florida and Texas. As a result of new partisan gerrymanders in these states in 2002 (and 2004, in Texas), the Republicans increased their majority of representatives from these states from 11 to 32.⁸ Our results imply that this shift in power may well have affected the nature of representation in other states as well.

Our results also speak to the phenomenon of independent redistricting commissions. Such non-partisan bodies handle apportionment in Iowa, Arizona, and now California. Although they failed to pass, ballot initiatives in Florida and Ohio recently considered this change too. Our results imply that there is both a direct and an *indirect* effect of adopting such an institution. That is, with California’s districts constructed by an independent commission then the strategies of Democrats and Republicans should change in other states. In principle, such effects could be large—particularly since the state in question has a large number of districts. Since the change in strategies leads to districts being constructed with less extreme median voters in other states, this may be seen as an additional benefit of independent commissions.

⁸Due to reapportionment, these states collectively gained one representative in 2002.

References

- [1] Friedman, John N., and Richard T. Holden (2008). “Optimal Gerrymandering: Sometimes Pack but Never Crack.” *American Economic Review*, 98(1):113-144.
- [2] Friedman, John N., and Richard T. Holden (2009). “The Rising Incumbent Reelection Rate: What’s Gerrymandering Got to Do With It?” *Journal of Politics*, forthcoming.
- [3] Gilligan, Thomas W., and John G. Matsusaka (1999). “Structural Constraints on Partisan Bias Under the Efficient Gerrymander.” *Public Choice*, 100 (1/2): 65-84.
- [4] Gul, Faruk, and Wolfgang Pesendorfer (2008). “Strategic Redistricting.” Princeton University Working Paper.
- [5] Karlin, Samuel (1968). *Total Positivity*. Stanford University Press, Stanford CA.
- [6] Milgrom, Paul and Chris Shannon (1994). “Monotone Comparative Statics.” *Econometrica*, 62(1): 157-180.
- [7] Owen, Guillermo, and Bernard Grofman (1988). “Optimal Partisan Gerrymandering.” *Political Geography Quarterly*, 7(1): 5-22.
- [8] Schoenberg, I.J. (1947). “On Totally Positive Functions, Laplace Integrals and Entire Functions of the Laguerre-Polya-Schur Type,” in *Proceedings of the National Academy of Sciences* 33: 11-17.
- [9] Schoenberg, I.J. (1951). “On Polya Frequency Functions,” *Journal d’Analyse Mathématique* 1(1): 331-374.

5 Appendix

Proof of Proposition 1 *This result follows the proof of Proposition 6 in FH (2008). Note that objection function, for each district a party p must create, can be factored such that*

$$EV_p = B_s(\mu_{ns}) K_{ns} + (1 - B_s(\mu_{ns})) L_{ns}$$

where $K_{ns} = E[V_p | d_{ns} = 1]$ and $L_{ns} = E[V_p | d_{ns} = 0]$ denote the expected value if party p were to win or lose district n in state s , respectively. Now, fix the districting plan (for both parties) and consider the change in the objective function resulting from small deviation from the existing plan in district n with an offsetting change in district m , with both districts in state s . The derivative of the value function, with respect to this change (which, in shorthand, we denote χ), is

$$\frac{\partial E[V_p]}{\partial \chi} = b_s(\mu_{ns}) (K_{ns} - L_{ns}) \frac{\partial \mu_{ns}}{\partial \chi} - b_s(\mu_{ms}) (K_{ms} - L_{ms}) \frac{\partial \mu_{ms}}{\partial \chi} = 0$$

which must equal 0 for the plan to be optimal. At this point, we note that, but for the constants K_{ns} , L_{ns} , K_{ms} , and L_{ms} , this expression is identical to that in equation (7) of FH (2008). Thus, we can directly apply Lemmas 1 through 3 from that paper, which imply Propositions 1 that paper, which is the result here. Since any optimal strategy must have this form, it must be that all equilibria must be such that each party employs a strategy of this form.

QED

Proof of Proposition 2 *The proof follows exactly along the lines of Proposition 2 from FH (2008). Since all optimal districting schemes have this feature, it must be that all equilibria involve strategies with this feature.*

QED

Proof of Proposition 3 Define the function $\phi^* (\{\mu_{nR}\}, \{\mu_{nD}\}; x)$ such that

$$X (\{\mu_{nR}\}, \{\mu_{nD}\}, \phi^*) = \lambda \sum_{n \in R} C (\mu_n + \phi^*) + (1 - \lambda) \sum_{n \in D} C (\mu_n + \phi^*) = x.$$

The maximization problem for party R can then be written as

$$\max_{\{\mu_{nR}\}} \int W' (x) [1 - B (\phi^* [\{\mu_{nR}\}, \{\mu_{nD}\}; x])] dx$$

such that $\{\mu_{nR}\} \in \Omega_R$.

In words, the party gets $W' (x)$ if the aggregate shock is higher than $\phi^* (x)$, and we must add up across all of the values x . At an optimum it cannot be the case that reallocating voters with positive mass between (say) district i to district j increases the value function and is still within the constraint set. However, consider such a reallocation and denote the increase in the median of district i as Δ_i and the decrease in the median of district j as Δ_j . Since the value function is differentiable it must be that for any two districts i and j in the same state

$$\frac{\int W' (x) b (\phi^* (x)) \frac{\partial \phi^* (x)}{\partial \mu_i} dx}{\int W' (x) b (\phi^* (x)) \frac{\partial \phi^* (x)}{\partial \mu_j} dx} = \lim_{\varepsilon \rightarrow 0} \frac{\Delta_j}{\Delta_i},$$

where the limit is taken such that the profile of switching voters is held constant. But, by our definition of ϕ^* above, we know that

$$\frac{\partial \phi^*}{\partial \mu_i} = \frac{c (\mu_i^* + \phi^* (x))}{\lambda \sum_{n \in R} c (\mu_{nR}^* + \phi^*) + (1 - \lambda) \sum_{n \in D} c (\mu_{nD}^* + \phi^*)}.$$

Therefore the above ratio can be rewritten as

$$\frac{\int W' (x) f (\phi^* [\{\mu_{nR}^*\}, \{\mu_{nD}^*\}; x]) c (\mu_i + \phi^* [\{\mu_{nR}^*\}, \{\mu_{nD}^*\}; x]) dx}{\int W' (x) f (\phi^* [\{\mu_{nR}^*\}, \{\mu_{nD}^*\}; x]) c (\mu_j + \phi^* [\{\mu_{nR}^*\}, \{\mu_{nD}^*\}; x]) dx} = \lim_{\varepsilon \rightarrow 0} \frac{\Delta_j}{\Delta_i},$$

where ϕ^* is that value associated with the equilibrium strategies. But these are the just the necessary conditions to the problem in which the gerrymanderer maximizes the alternative

objective function

$$\max_{\{\mu_{nR}\}} \int W'(x) f(\phi^* [\{\mu_{nR}^*\}, \{\mu_{nD}^*\}; x]) \sum_n C(\mu_n + \phi^* [\{\mu_{nR}^*\}, \{\mu_{nD}^*\}; x]) dx.$$

such that $\{\mu_{nR}\} \in \Omega_R$.

QED

Proof of Corollary 1 Consider the situation in which party R 's value function is

$$\begin{aligned} W_n &= \frac{x^n}{x^n + (1-x)^n} \\ W'_n &= \frac{((x-1)x)^n (\log x - \log(1-x))}{(x^n + (1-x)^n)^2} \end{aligned}$$

Note that, as $n \rightarrow \infty$, W limits to the desired function. By Proposition 3, party R solves the alternative maximization

$$\max_{\{\mu_{dR}\}} \int \left[\frac{W'_n(x)}{W'_n(\frac{1}{2})} b(\phi^*(x)) \sum C(\mu_d + \phi^* (\{\mu_{dD}^*\}, \{\mu_{dR}^*\}, x)) \right] dx$$

such that $\{\mu_{nR}\} \in \Omega_R$.

which is identical to equation (1) above but for scaling by the constant term $W'_n(\frac{1}{2})$. But as $n \rightarrow \infty$, the weights

$$\lim_{n \rightarrow \infty} \frac{W'_n(x)}{W'_n(\frac{1}{2})} \rightarrow \begin{cases} 0 & x \neq \frac{1}{2} \\ 1 & x = \frac{1}{2} \end{cases}.$$

Thus the necessary conditions are simply that

$$\frac{c(\mu_i + \phi^*)}{c(\mu_j + \phi^*)} = \frac{\Delta\mu_j}{\Delta\mu_i}$$

These are the same necessary conditions as if party R simply maximized the number of seats

won at critical value $\phi^* \left(\{\mu_{nD}^*\}, \{\mu_{nR}^*\}, \frac{1}{2} \right)$, which could be written

$$\max_{\{\mu_{nR}\}} \sum_{n \in R} C \left(\mu_n + \phi^* \left(\{\mu_{nD}^*\}, \{\mu_{nR}^*\}, \frac{1}{2} \right) \right)$$

such that $\{\mu_{nR}\} \in \Omega_R$.

QED

Proof of Proposition 4 *Following Corollary 1, there are two FOCs that combine to imply*

$$\frac{c(\mu_1 + \phi^*)}{c(\mu_2 + \phi^*)} = \frac{\Delta\mu_2}{\Delta\mu_1}.$$

Writing $\mu_2(\mu_1)$ one can substitute into the objective function above, so that the FOC becomes

$$\begin{aligned} \mu_1^* &= \arg \max_{\{\mu_1\}} \{C(\mu_1 + \phi) + C(\mu_2(\mu_1) + \phi)\} \\ &\Rightarrow \frac{c(\mu_1 + \phi^*)}{c(\mu_2 + \phi^*)} = -\frac{d\mu_2(\mu_1)}{d\mu_1}. \end{aligned}$$

Of course, $\frac{d\mu_2(\mu_1)}{d\mu_1} < 0$. Then, by the implicit function theorem, we know that $\frac{\partial\mu_1^*}{\partial\phi^*} < 0$ if and only if the LHS is decreasing in ϕ^* , which is true if and only if

$$\frac{c'(\mu_1 + \phi^*)}{c'(\mu_2 + \phi^*)} < \frac{c(\mu_1 + \phi^*)}{c(\mu_2 + \phi^*)}$$

Note that by the equal mass constraint it must be that $\frac{\partial\mu_2}{\partial\phi^*}$ is of the opposite sign as $\frac{\partial\mu_1^*}{\partial\phi^*}$.

Moreover, $\frac{\partial\mu_1^*}{\partial\phi^*}$ depends entirely on whether the ratio $\frac{c'(\psi)}{c(\psi)}$ is increasing or decreasing in ψ .

This ratio being decreasing is precisely the definition of log concavity and hence $\frac{\partial\mu_1^*}{\partial\phi^*} < 0$.

QED

Proof of Proposition 5 *Following Corollary 1, we maximize the objective function*

$$\max_{\{\mu_{nR}\}} \sum_{n \in R} C(\mu_n + \phi^*).$$

Consider a deviation in which one shifts μ_1 upwards by amount $\Delta\mu_1$ and then shifts all other medians down by $\Delta\mu_{-1}$. The no-benefit condition from such a deviation is

$$(3) \quad \frac{c(\mu_1 + \phi^*)}{\sum_{n \neq 1} c(\mu_n + \phi^*)} = \frac{\Delta\mu_{-1}}{\Delta\mu_1}.$$

Note, at this point, that the medians $\{\mu_2, \dots, \mu_N\}$ are chosen optimally. Therefore, we can implicitly differentiate this expression to obtain the impact of ϕ^* on μ_1 , since all deviations within the medians $\{\mu_2, \dots, \mu_N\}$ have a second order impact on the value function, by the Envelope Theorem. From Proposition 4, we know that the ratio $\frac{c(\mu_1 + \phi^*)}{c(\mu_N + \phi^*)}$ is falling with ϕ^* , and therefore we know that the LHS of equation (3) is also decreasing in ϕ^* . Therefore, we know that $\frac{\partial\mu_1^*}{\partial\phi^*} < 0$. A parallel argument establishes that $\frac{\partial\mu_N^*}{\partial\phi^*} > 0$.

QED

Proof of Proposition 6 Karlin (1968, p.30) shows that the convolution $h = f \cdot g$ is PF_n if f and g are PF_n . By a theorem of Schoenberg (1947, 1951), PF_2 of a density is equivalent to log-concavity. The assumption of uniformity of b means that we are left with the term $\int W'(x) c(\mu_i + \phi^*(x)) dx$, which is PF_2 since $W'(x)$ is PF_2 and c is log-concave. Since $W'(x)$ is PF_2 it is integrable and hence continuous. Now the proof of Proposition 5 applies.

QED