Economic Theory Bulletin
https://doi.org/10.1007/s40505-020-00188-3

RESEARCH ARTICLE

Optimal Gerrymandering in a competitive environment

John N. Friedman[1] · Richard Holden[2]

Received: 31 May 2019 / Accepted: 25 June 2020

Abstract
We analyze a model of optimal gerrymandering in which two parties simultaneously redistrict in a competition for influence in a legislature. Parties allocate geographic blocks to districts, in which the median voter determines the winner. The form of the optimal gerrymander involves "slices" of right-wing blocks paired with "slices" of left-wing blocks, as in Friedman and Holden (Am EconRev 98(1):113–144, 2008). We also show that, as one party controls the redistricting process in more states, that party designs districts such that the most extreme districts within its control become more extreme. We show that this comparative static holds for a broad class of objective functions.

Keywords Gerrymandering · redistricting

JEL Classification D72

1 Introduction

A growing literature analyzes *gerrymandering*, the process by which politicians draw the boundaries of their own electoral districts. To simplify the analysis, most of this literature assumes that one party controls the redistricting process (Owen and Grofman 1988; Shershtyuk 1998; Gilligan and Matsusaka 1999; Friedman and Holden 2008). In practice, Republicans and Democrats each control the districting process in a number of states. In this context, the environment is best represented as

Holden acknowledges ARC Future Fellowship FT130101159. We thank Christopher Teh for excellent research assistance.

✉ Richard Holden
richard.holden@unsw.edu.au

John N. Friedman
john_friedman@brown.edu

1 Department of Economics, Brown University, Providence, RI, USA

2 School of Economics, University of New South Wales, Sydney, NSW, Australia

Published online: 10 July 2020

a two-player game rather than a control problem. In this paper, we build on our work in Friedman and Holden (2008) to provide a treatment of the two-player strategic districting game.

Historically, redistricting was primarily a local affair: parties relied upon block captains and local politicians with intimate knowledge of their neighborhood to determine likely voter behavior. In recent years, however, coordination across states has become prevalent. National party organizations have built ever more detailed voter databases;[1] the digitization of districts through TIGERLine files and electronic Census records has made it easier for national officials to participate in local redistricting; and national parties have organized inter-state redistricting efforts, such as the so-called "REDMAP" project. These factors have allowed national strategic concerns to play an important role in redistricting.[2] These concerns are the focus of this paper.

Our analysis has two parts. First, we extend Friedman and Holden (2008) to a multi-state, multi-party environment. We also explicitly consider geographic constraints on redistricting by assuming that parties may only allocate whole "blocks" to districts, rather than voters individually. The key result is that the form of the optimal gerrymander in Friedman and Holden (2008) is the same in the richer environment: when within-block voter distributions are sufficiently concentrated, the party in control forms districts by matching a group of right-wing voter blocks with blocks of left-wing voter blocks, with these "slices" becoming progressively less extreme as the district becomes less favorable to the redistricting party.

Second, we compute comparative statics on optimal district formation with respect to key parameters of the redistricting game. Control over districting can vary substantially between elections: for instance, in 2002, the G.O.P won control of redistricting in seven states, giving them a net gain in control of 95 districts.[3] We show that as one party controls the redistricting in more states, that party redraws electoral boundaries such that the most extreme districts within its control become more extreme. This party stretches the range of district median voters in the states it controls, whereas its opponent compresses the range of district median voters in the states it controls.

The work most closely related to ours is an elegant paper by Gul and Pesendorfer (2010). They characterize the set of redistricting equilibria by restating the game as a control problem involving maximizing the number of seats won at the cutoff values of an aggregate shock to voter preferences. This also allows them to provide the important comparative static of the effect of a change in the number of states districted by a particular party on the optimal gerrymander. One simplifying assumption in Gul and Pesendorfer (2010) 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

[1] For a recount of the Obama campaign's use of data, see for instance https://www.technologyreview.com/s/509026/how-obamas-team-used-big-data-to-rally-voters/.

[2] See, for instance https://www.theatlantic.com/politics/archive/2017/10/gerrymandering-technology-redmap-2020/543888/.

[3] See Friedman and Holden (2009) for a detailed breakdown.

only two voter types, but not with more types (Friedman and Holden 2008, Proposition 3). Our first result on the matching slices strategy contrasts with this. Despite the additional complexity of a continuum of voter types, we also analyze significantly more general objective functions than the simple majoritarian function in Gul and Pesendorfer (2010).

Our results also interact more broadly with the growing literature on districting. A number of papers analyze the impact of majority-minority districts (see Cameron et al. 1996; Epstein and O'Hallaran 1999; Shotts 2001; Bailey and Katz 2005). We show that, even when redistricters are strategically interacting, majority–minority districts are optimal for neither the party favored by minorities, nor the party opposed by them. Gilligan and Matsusaka (2005) and Coate and Knight (2007) analyze districting from a normative perspective. Shotts (2002), Besley and Preston (2007), and Cox and Holden (2011) analyze the interaction between redistricting and policy.

The remainder of this paper is organized as follows. Section 2 shows that the matching slices strategy of Friedman and Holden (2008) is obtained 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. We also recast the model as one in which parties allocate geographic blocks of voters, rather than individuals themselves, to districts.

2.1 The model

There are a total of S states comprising a total of N districts. Each state contains N_s districts. To explicitly account for the geographical constraints on redistricting, we further assume that each district must be comprised of whole "blocks." We use the term "block" here to refer to a generic geographical unit; some states (such as Iowa) require that districts comprise whole counties, while others allow for much finer distinctions.[4] For simplicity, we assume that in each state, there are a continuum of equal-sized infinitessimal blocks. Let P_s denote the total mass of blocks in state s.

Within each block j exists a continuum of voters indexed by $i \in \mathbb{R}$. Voters are ordered monotonically by their preferences for either of two existing electoral parties, D and R. Preferences are characterized by $\beta_i \in \mathbb{R}$. Let voters with a high value

[4] Note that all states require contiguity for districts. We do not model this constraint explicitly due to the significant added complexity required in such a model.

of β_i (i.e., more to the right) prefer party R more relative to party D. Let voters actually vote for R if and only if $\beta_i > 0$.[5]

We assume that each party knows perfectly the distribution of preferences within each block. Without loss of generality, let $\sigma_j \in \mathbb{R}$ denote an index of the distribution of partisan preference within a block, and let σ_j be ordered such that a higher signal implies a more right-wing block. Let $g_s\left(\beta_i|\sigma_j\right)$ denote the distribution of individual preferences within block j in state s, which we refer to as the "conditional preference distribution." The marginal distribution of blocks in state s, or the "block distribution" , is denoted by $h_s\left(\sigma_j\right)$.

The median voter determines the winner in each district. Let $\mu_{n,s}$ denote the preference of the median voter in district n in state s. $\mu_{n,s}$ is further affected by aggregate uncertainty in state s and district n.[6] Let $b_{n,s}$ denote this noise, and suppose that aggregate noise occurs with CDF $B(\cdot)$ and unbounded support. Thus, if the median voter holds *ex ante* bliss point $\mu_{n,s}$, she will vote after receiving the shock as though she has bliss point $\hat{\mu}_{n,s} = \mu_{n,s} - b_{n,s}$, and so the probability that party R wins district n in state s is $B\left(\mu_{n,s}\right)$.

Each state s is to be redistricted by only one party.[7] Each party p has value function $W_p : [0, 1] \to \mathbb{R}$, whose domain is the fraction of districts won in the election. We assume that each W_p is weakly increasing and strictly increasing at least somewhere, and that parties maximize expected payoffs, denoted by EW_p.

Parties act by choosing $\psi_{n,s}\left(\sigma_j\right)$, that is, the distribution of blocks to be placed in district n of state s, in each state that they control. Thus, party R and D choose the following, respectively:

$$\left\{\psi_{n,s}\left(\sigma_j\right)\right\}_{s=1,n=1}^{s=S_R,n=N_s} \qquad \left\{\psi_{n,s}\left(\sigma_j\right)\right\}_{s=S_R+1,n=1}^{s=S,n=N_s}$$

We impose two constraints on the formation of districts. First, each district in state s must contain the same mass of blocks, $\frac{P_s}{N_s}$. Second, each block in state s must appear in exactly one district in state s. We assume that parties move simultaneously,[8] and focus on the Nash equilibria of this game.

To state the optimization problem formally, we define $r_{n,s}$ as a dummy variable equal to one if party R wins the election in district n in state s. Then, party R faces the problem:

$$\max_{\left\{\psi_{n,s}\left(\sigma_j\right)\right\}_{s=1,n=1}^{s=S_R,n=N_s}} EW_R\left(\frac{1}{N}\left(\sum_{s=1}^{S_R}\sum_{n=1}^{N_s} r_{n,s} + \sum_{s=S_R+1}^{S}\sum_{n=1}^{N_s} r_{n,s}\right)\right) \tag{1}$$

[5] 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.

[6] The composition of this uncertainty will be elaborated upon in Sect. 3.

[7] It is trivial to extend these to include a third group of states 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. For the sake of simplicity, we focus on the two-party case.

[8] This assumption matches the reality that 49 states must (by state law) redistrict within a window of about 6 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.

$$s.t. \quad \int_{-\infty}^{\infty} \psi_{n,s}(\sigma_j) d\sigma_j = \frac{P_s}{N_s} \quad \forall n, s$$

$$\sum_{n=1}^{N_s} \psi_{n,s}(\sigma_j) = h_s(\sigma_j) \quad \forall \sigma_j, s$$

$$0 \leq \psi_{n,s}(\sigma_j) \leq h_s(\sigma_j) \quad \forall n, \sigma_j, s,$$

and party D solves a parallel problem where $d_{n,s}$ is a dummy variable equal to one if party D wins the election in district n in state s. Analogously, the choice variables for party D are the districting schemes in states $s \in [S_{R+1}, S]$.

We now make two assumptions about the distribution of voters within each block. First, we require that the block index σ_j is informative about the underlying distribution of individual preferences β_i, in a specific sense.

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

$$\frac{z_s(\beta_i \mid \sigma_j')}{z_s(\beta_i \mid \sigma_j)} < \frac{z_s(\beta_i' \mid \sigma_j')}{z_s(\beta_i' \mid \sigma_j)}, \quad \forall \sigma_j' > \sigma_j, \ \beta_i' > \beta_i, \ s$$

This property is similar to the Monotone Likelihood Ratio Property, and if the signal shifts only the mean of the conditional preference distribution, then this property is equivalent to MLRP.[9] As applied to the distribution of voters within a block, Condition 1 implies that blocks can be categorized easily on a left-to-right basis. In other words, Condition 1 rules out the situation in which a single block contains voters on the extreme in both directions but no moderates. While this assumption cannot literally be true, the literature on the geographic distribution of voter preferences broadly supports this condition. It is difficult to directly measure the strength of partisan preferences; although donors differ from voters in certain respects, campaign contributions are one proxy. Gimpel et al. (2006) show that there is significant geographic concentration in the distribution of party donations across counties. At a more local level, fund-raising maps show distinct clusters of Democratic and Republican support in large cities where blocks are densest. For instance, contributions to Republican candidates in the Boston Metro Area are clearly clustered in area such as Wellesley and Belmont Hill rather than evenly distributed throughout the area.[10]

Second, we require a technical condition on a particular form of unimodality.

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

[10] See `https://web.archive.org/web/20121217195343/http://fundrace.huffingtonpost.com/`.

Condition 2 **(Central Unimodality)** For all s, $g_s\left(\beta_i \mid \sigma_j\right)$ is a unimodal distribution where the mode lies at the median.

Note that, without loss of generality (given Condition 1), we can "re-scale" σ_j such that $\sigma_j = \max_{\sigma_i} g_s\left(\beta_i \mid \sigma_j\right)$. The two parts of Condition 2 essentially require that β_i is distributed "near" σ_j, and not elsewhere.[11]

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 concentrated distribution of voters within blocks, the optimal districting plan in any equilibrium, for each party p, in each state s, can be characterized by breakpoints* $\left\{u_{n,s}\right\}_{n=1}^{N_s}$ *and* $\left\{l_{n,s}\right\}_{n=1}^{N_s}$ *(ordered such that* $u_{1,s} > u_{2,s} > \ldots > u_{N_s-1,s} > l_{N_s-1,s} > \ldots > l_{1,s} \geq -\infty$*) such that*

$$\psi_{1,s} = \begin{cases} h_s\left(\sigma_j\right) & \text{if } \sigma_j < l_{1,s} \text{ or } \sigma_j > u_{1,s} \\ 0 & \text{otherwise} \end{cases},$$

$$\psi_{n,s} = \begin{cases} h_s\left(\sigma_j\right) & \text{if } l_{n-1,s} < \sigma_j < l_{n,s} \text{ or } u_{n-1,s} > \sigma_j > u_{n,s} \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 1 < n < N,$$

$$\text{and } \psi_{N_s,s} = \begin{cases} h_s\left(\sigma_j\right) & \text{if } \sigma_j > l_{N-1,s} \text{ and } \sigma_j < u_{N-1,s} \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, the preferences of median voters in each district of state s under the optimal district plan are such that $\mu_{1,s} > \mu_{2,s} > \ldots > \mu_{N_s,s}$.

This result establishes that "cracking" is not optimal, so that parties find it optimal to instead group the most partisan blocks into one district within a given state. However, parties still may wish to "pack" those least favorable into segregated districts, such that each district comprises of only one continuous "slice " of the marginal block distribution. We now provide conditions under which packing too is not optimal.

Proposition 2 *Suppose that Conditions 1 and 2 hold, and there is a sufficiently concentrated distribution of voters within blocks.Then in any set of equilibrium redistricting strategies, there exists* n *and* $\sigma_j < \sigma_j'$ *such that* $\mu_{n,s} > \mu_{N_s,s}$ *and* $\psi_{n,s}\left(\sigma_j\right) > 0, \psi_{N_s,s}(\sigma_j') > 0$ *for all s.*

We refer to the equilibrium 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

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

of the distribution. The intuition behind this result is as follows.[12] We may deviate from a districting plan that "packs" by redistricting the most left-wing blocks on the farthest left of the block distribution to the district encompassing the furthest right slice, while "sliding " all other districts to the right. While this decreases the likelihood of winning in the (initially) far right district, the likelihood of winning all other districts rise as the median voters' preferences rise in expectation. As the within-block voter distribution becomes more concentrated around its median/mode, the gains from redistricting rise somewhat consistently, while the losses do not rise as quickly. At some point, it becomes optimal to allow the right-most district to encompass two extreme slices at either end of the distribution. We may proceed to apply the same logic to the redistricting of all other remaining single-sliced districts to obtain the *matching slices* result described in Proposition 2, where each district, excluding the middle-most district, will contain two slices.

Figure 1 is an example of such a strategy for an arbitrary marginal block distribution in state s (with five districts, each represented by a different shade of gray, redistricted by party R) that satisfies the conditions in Propositions 1 and 2. The right-hand slice making up each district is larger than the left-hand slice to ensure that party R, whose voters lie more to the right, have a majority in expectation. Furthermore, districts with medians more favorable to party R have smaller right-hand slices. Intuitively, those voters far to the right vote more reliably for party R; thus, the party needs fewer of them in a district in order to guarantee a victory. In the extreme, where voters on the right support party R with probability 1, redistricting would occur such that the right-most slice only contains ε more voters than the left wing.

These results extend those in Friedman and Holden (2008) to the richer setting in which each party controls only a fraction of districts and allocates blocks, rather than voters. To understand intuitively why the original results extend to this broader case, consider the gain to party R from winning a given district, as opposed to losing it. If the value function is non-linear, this value will depend on the set of states controlled by party D and party D's districting plan for those states. But holding all else fixed—which is precisely what happens in a 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 the simpler model by constant terms. While a party may alter the number of right-wing blocks in the upper "slice" of each district, the fundamental strategy in equilibrium, as in Propositions 1 and 2, remains the same.

[12] An analogous discussion of this result is present in Friedman and Holden (2008).

3 Strategic interactions in redistricting

Propositions 1 and 2show that the optimal gerrymander will always take the form of "matching slices." We can therefore abstract to some degree from the precise microfoundation of the constraints on district formation 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. Therefore, we now rewrite the redistricting problem as one with choice variables being the district-level medians $\{\mu_{n,s}\}$. We then capture the constraints from above; let the feasible set of medians for player R be Ω_R. This constraint set Ω_R embodies all of the constraints faced by party R in the fully microfounded problem stated formally above in equation (1). Even though the parties cannot observe each voter's type individually, they may rearrange voters to generate a range of district medians in each state. Define the CDF of all district medians as $M(\mu)$. Denote by $\{\mu_{nR}\}$ and $\{\mu_{nD}\}$ the medians of all districts in states controlled by party R or D, respectively.

To analyze strategic interactions in redistricting, it now becomes important to expand upon the components of aggregate uncertainty, $b_{n,s}$. Suppose that $b_{n,s} = \nu_{n,s} - \phi$, where $\nu_{n,s}$ corresponds to a local shock in state s and district n with CDF $C(\cdot)$ and PDF $c(\cdot)$, and ϕ corresponds to a national shock with CDF $Y(\cdot)$ and PDF $y(\cdot)$, both with unbounded support. Both shocks are independent, have a mean of zero and are symmetrically distributed. Given this, the median voter in district n in state s votes with *ex post* preferences is $\hat{\mu}_{n,s} = \mu_{n,s} - \nu_{n,s} + \phi$.

In order to focus the analysis on complementarities across states, we use the national shock as a summary statistic for the "state" of the election. We therefore assume that each party has control over an infinite number of districts, allowing us to integrate over the distribution of local shocks rather than account for them individually through a combinatorial equation.[13] Define

$$\lambda = \frac{\sum_{s=1}^{S_R} N_s}{\sum_{s=1}^{S} N_s}$$

as the share of districts controlled by party R. This assumption, alongside that of which shocks are independent, allows us to write the share of districts won by party R as

$$X(\phi) = \int C(\mu + \phi)\mathrm{d}M(\mu).$$

The party values the fraction of seats won by the function $W(\cdot)$, which is weakly increasing. We may therefore rewrite the optimization problem faced by party R as

[13] While this is mostly a technical assumption, in practice, both the Republican and Democratic parties control over 150 districts, and so this assumption is not unreasonable on its face. Furthermore, it seems natural for parties to spend considerably more time anticipating possible national shocks, such as a major recession or war, rather than the national influence of many *uncorrelated* local shocks.

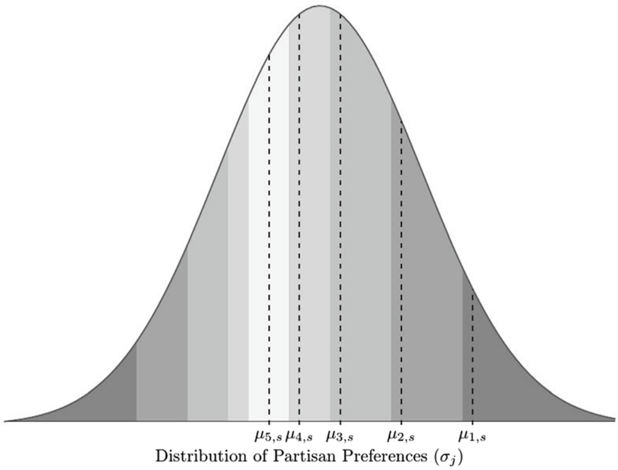

Fig. 1 The form of the optimal gerrymander

$$\max_{\{\mu_{nR}\}} EW = \int W(X(\phi))\mathrm{d}Y(\phi)$$
$$\text{such that } \{\mu_{nR}\} \in \Omega_R. \tag{2}$$

3.1 Optimal Gerrymandering under shocks

In general, the share of districts won by R is given by

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

We may use this expression to restate the problem that the parties solve. To begin, define $\phi^*(\{\mu_{nD}\}, \{\mu_{nR}\}, X)$ as a function of the strategies chosen by the parties, that is, $\{\mu_{nD}\}$and $\{\mu_{nR}\}$ such that

$$\lambda \sum_{nR} C(\mu_{n,s} + \phi^*) + (1-\lambda)\sum_{nD} C(\mu_{n,s} + \phi^*) = X.$$

We may interpret ϕ^* as the "pivotal value" of the shock that, given each party's strategy, just allows R to win X share of districts. Proposition 3 then shows that each party acts as though they maximize vote shares over a weighted average of possible outcomes of the aggregate shock ϕ.

Proposition 3 *Suppose that $W(X)$ is continuous and differentiable.Then the optimal gerrymander $\{\mu^*_{nR}\}$ will satisfy the necessary conditions to the problem*

$$\max_{\{\mu_{nR}\}} \int \left[W'(X)y(\phi^*(X)) \sum C\left(\mu_{n,s} + \phi^*\left(\{\mu^*_{nD}\}, \{\mu^*_{nR}\}, X\right)\right)\right] \mathrm{d}X \qquad (3)$$
$$\text{such that } \{\mu^*_{nR}\} \in \Omega_R,$$

while $\{\mu^*_{nD}\}$ satisfies the necessary conditions to the parallel problem.

The advantage to the alternative maximization above is that it does not involve anything about the districts designed by the opponent party, 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, 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 (2010), who 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

$$\{\mu^*_{nR}\} = \arg\max_{\{\mu_{nR}\}} \sum C\left(\mu_{n,s} + \phi^*\left(\frac{1}{2}\right)\right),$$

and likewise for party D, 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. In such a situation, 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 districts, and so the aggregate shock must simply be above average for party R to win, and vice versa for party D. 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.

3.2 Comparative statics

We now consider whether a party that controls redistricting in more states acts differently in equilibrium to a party that controls fewer states. Proposition 3 has rephrased each party's problem as maximizing vote share conditional on the pivotal value ϕ^*. Of course, this requires knowing ϕ^*, which is jointly determined with the optimal strategy. We will show that, if c is log-concave, then as λ increases, ϕ^* decreases; a party that is advantaged with more districts gerrymandered needs less "luck" from the aggregate shock. Thus, we can solve for the comparative static of λ by solving for the comparative static of ϕ^*. While we cannot actually solve for ϕ^* itself, we can assume a value of ϕ^* and see the effect of changing that value on the optimal districting scheme.

Proposition 4 *Assume that $W(X)$ takes on the form presented in Corollary 1 and that c is log-concave. Recall that the number of districts in state s is N_s. As λ increases so that party R redraws more districts, $\mu^*_{1,s}$ increases and $\mu^*_{N_s,s}$ decreases if state s is controlled by party R, whereas $\mu^*_{1,s}$ decreases and $\mu^*_{N_s,s}$ increases if state s is controlled by party D.*

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 favorable and unfavorable districts may be in focus, since the local shock necessary to tip a district to one party or another is not so big. By contrast, if a party controls many states, then the aggregate shock will have to be very negative for the party to lose the election. Since the aggregate shock is so negative, unfavorable districts are now essentially unwinnable, and so increasing the median voter for the right party in these states is of little help.

This result implies that the control of redistricting matters crucially for the nature of representation in the legislature. There are two main 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 in play changing the way parties draw districts in states they do control. As one party controls more states, it draws districts such that the most extreme districts, that is, the districts with the right-most and left-most medians, become even more extreme. This increases polarization in each state they control, where polarization is measured as the distance between median voters in the most extreme districts. If the positions of candidates are tied to the distribution of voters in their district—for instance, through competitive primary processes—then, for a given party, controlling redistricting in a larger number of states increases the representation of extreme voters and stretches out the distribution of district medians in a given state. The opposite features are exhibited by the party losing control of the redistricting process.

A more detailed view of polarization would consider medians in the intermediate districts. However, the complexity of the problem prevents a more systematic characterization. For instance, suppose a given state controlled by party R includes three districts. As party R gains more control over redistricting in other states, 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.

Table 1 presents a numerical example that highlights these forces. These findings are further illustrated graphically, with respect to a particular block distribution, in Fig. 2. In this example, we suppose that there is a unit mass of identical states with five districts each. In each state, both the block distribution $h_s(\sigma_j)$ and the conditional preference distribution $g_s(\beta_i \mid \sigma_j)$ are normal distributions with mean 0 and variance 2.5. We assume that $Y(\phi)$ and $C(\nu_{n,s})$ are normal distributions with mean 0 and variance $\frac{1}{4}$.

Each row of Table 1 presents the equilibrium strategy of party R given a share of control λ. An increase in λ pulls the most extreme district medians apart; $\mu^*_{1,s}$ increases while $\mu^*_{5,s}$ decreases. The middle median, $\mu^*_{3,s}$, moves up, then down as party R's state control increases. While the intermediate medians, $\mu^*_{2,s}$ and $\mu^*_{4,s}$, display a similar monotonic movement pattern to that of their most extreme counterparts, 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.

It is important to note that Proposition 4 relies on the log-concavity of c. Intuitively, log-concavity ensures that larger preference shocks, in some sense, are less likely than smaller shocks. To show the role of the distribution of c on the redistricting equilibrium, Example 1 considers a highly specialized (and unrealistic) case that entirely eliminates any strategic interaction.

Example 1 Suppose that C is a uniform distribution, so that $c(\cdot) = k$ (a constant). Then each party's optimal gerrymander maximizes the average of the median voters in the districts in their control: the share of states λ under the control of party R has no impact on the optimal gerrymander.

The intuition is almost precisely the opposite to that above. When C is uniform, an aggregate shock of any kind is equally likely: the strategic effect of controlling a larger number of states disappears. This is, in a sense, "the exception that proves the rule."

Note that the intuition here is very similar to that in Gul and Pesendorfer (2010), who emphasize the same comparative static. Given the difference in the information structure between our model and theirs—leading, importantly, to a very different optimal gerrymandering strategy—Proposition 4 demonstrates that the preceding intuition—that, as a party controls more states, its favorable districts are drawn to be made more favorable—is a robust one. However, since the strategies used by the gerrymanderer differ between the two models, it is not *a priori* obvious that this would be the case nor is it clear how to map one result onto the other without performing further analysis.

3.3 Generalizing the objective function

The above results have focused on the case when the objective function is a step function with a single discontinuity, as in Gul and Pesendorfer (2010). With a more complex objective function, each first order condition between two districts i and j no longer depends on the simple ratio $\frac{c(\mu_i+\phi^*)}{c(\mu_j+\phi^*)}$ but rather on the ratio of weighted averages

$$\frac{\int W'(x)y(\phi^*)c\left(\mu_i+\phi^*\right)\mathrm{d}x}{\int W'(x)y(\phi^*)c\left(\mu_j+\phi^*\right)\mathrm{d}x}.$$

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. We can, however, provide a condition on the objective function under which Proposition 5 generalizes. This involves *total positivity*[14] and so-called *Pólya frequency functions*. Hence, some definitions are in order before stating our result.

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

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

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

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 5 *Suppose that $W'(X)$ is PF_2, that y is the uniform distribution and that c is log-concave. Then as the share of districts controlled by R, λ, increases, $\mu^*_{1,s}$ increases and $\mu^*_{N_s,s}$ decreases if state s is controlled by party R, whereas $\mu^*_{1,s}$ decreases and $\mu^*_{N_s,s}$ increases if state s is controlled by party D.*

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

[14] Total positivity has wide applications in economics. For instance, when K is a density TP_2 is equivalent to the monotone likelihood ratio property.

Table 1 Numerical examples of optimal competitive gerrymandering

Share of states controlled	Pr[Win Majority]	District median (probability of winning district)				
		1	2	3	4	5
$\lambda = 0.25$	37.6%	0.69 (84%)	0.43 (73%)	0.31 (67%)	0.04 (52%)	−0.79 (13%)
$\lambda = 0.50$	50.0%	0.87 (89%)	0.60 (80%)	0.49 (76%)	−0.15 (42%)	−1.02 (7%)
$\lambda = 0.75$	62.4%	1.00 (92%)	0.78 (87%)	0.53 (77%)	−0.29 (34%)	−1.20 (4%)
$\lambda = 1.00$	74.7%	1.15 (95%)	0.92 (90%)	0.45 (74%)	−0.36 (31%)	−1.30 (3%)

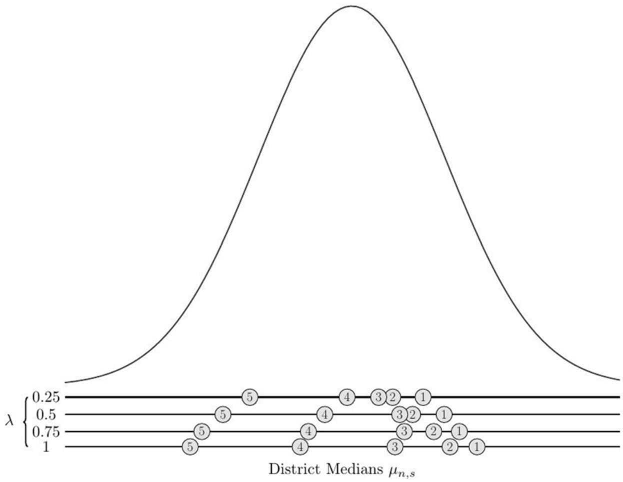

Fig. 2 Graphical illustration of comparative statics; numerical example

A natural question to ask is what objective functions $W(X)$ have derivatives which satisfy this requirement. Note that every PF_2 function is of the form $f(x) = e^{-\psi(x)}$, for convex $\psi(x)$ (Karlin 1968: p. 32). One function that satisfies Definition 2, of course, is the Normal cumulative distribution function.[15] This class of functions has 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. The logistic function $P(x) = \frac{1}{1+e^{-x}}$ is also PF_2.[16] This function has the attractive property that it is first convex and then concave, which seems like a natural objective function for a redistricter.

[15] Of course, $\Phi : \mathbb{R} \to [0, 1]$ while our $W : [0, 1] \to \mathbb{R}$. But the domain restriction is unimportant.

[16] To see this, note that $e^{-\log\left(\frac{1}{1+e^{-x}}\right)} = \frac{1}{1+e^{-x}}$, and that $-\log\left(\frac{1}{1+e^{-x}}\right)$ is convex.

It is also informative to think about objective functions W that do *not* satisfy this definition. An extreme example is that of a two-step double-discontinuity function. We explore a particular specification of this at the end of the appendix section (Example 2).

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 (2010). We then showed that an increase in the number of districts whose boundaries are drawn by a party tends to spread out the distribution of optimal district medians in states controlled by that party, and compress the distribution of optimal district medians in states controlled by the other party.

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. 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 seven states, including California and Arizona; three further states passed ballot initiatives in favor of independent redistricting in 2018, which will be implemented in the near future.[17] Our results imply that there is both a direct and an *indirect* effect of these commissions: with California's districts constructed by an independent commission, 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.

Appendix

Proof of Proposition 1 This result follows the proof of Proposition 7 in Friedman and Holden (2008). Note that the objective function, for each district a party R must create, can be factored, such that

$$EW_p = B\left(\mu_{n,s}\right)K_{n,s} + \left(1 - B\left(\mu_{n,s}\right)\right)L_{n,s},$$

[17] The ballot initiatives were passed in Utah, Colorado and Georgia.

where $K_{n,s} = E\left[W_p | r_{n,s} = 1\right]$ and $L_{n,s} = E\left[W_p | r_{n,s} = 0\right]$ denote the expected value if party R 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 a 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

$$\frac{\partial E\left[W_p\right]}{\partial \chi} = b\left(\mu_{n,s}\right)\left(K_{n,s} - L_{n,s}\right)\frac{\partial \mu_{n,s}}{\partial \chi} - b\left(\mu_{m,s}\right)\left(K_{m,s} - L_{m,s}\right)\frac{\partial \mu_{m,s}}{\partial \chi},$$

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

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

Proof of Proposition 3 Using the definition of ϕ^*, the maximization problem for party R can then be written as

$$\max_{\{\mu_{nR}\}} \int W'(X)\left[1 - Y\left(\phi^*\left[\left\{\mu_{nR}\right\}, \left\{\mu_{nD}\right\};X\right]\right)\right]dX$$
$$\text{such that } \left\{\mu_{nR}\right\} \in \Omega_R.$$

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 of $\Delta\mu_i$ and the decrease in the median of district j of $\Delta\mu_j$. Since the value function is differentiable it must be that for any two districts i and j in the same state

$$\frac{\int W'(X)y(\phi^*(X))\frac{\partial\phi^*(X)}{\partial\mu_i}dX}{\int W'(X)y(\phi^*(X))\frac{\partial\phi^*(X)}{\partial\mu_j}dX} = \lim_{\varepsilon\to 0}\frac{\Delta\mu_j}{\Delta_{\mu i}},$$

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

$$\frac{\partial \phi^*}{\partial \mu_i} = \frac{c\left(\mu_i^* + \phi^*(X)\right)}{\lambda \sum_{nR} c\left(\mu_d^* + \phi^*(X)\right) + (1-\lambda) \sum_{nD} c\left(\mu_d^* + \phi^*(X)\right)}.$$

Therefore the above ratio can be rewritten as

$$\frac{\int W'(X) y\left(\phi^*\left[\left\{\mu_{nR}^*\right\}, \left\{\mu_{nD}^*\right\}; X\right]\right) c\left(\mu_i + \phi^*\left[\left\{\mu_{nR}^*\right\}, \left\{\mu_{nD}^*\right\}; X\right]\right) dX}{\int W'(X) y\left(\phi^*\left[\left\{\mu_{nR}^*\right\}, \left\{\mu_{nD}^*\right\}; X\right]\right) c\left(\mu_j + \phi^*\left[\left\{\mu_{nR}^*\right\}, \left\{\mu_{nD}^*\right\}; X\right]\right) dX} = \lim_{\varepsilon \to 0} \frac{\Delta \mu_j}{\Delta \mu_i},$$

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

$$\max_{\{\mu_{nR}\}} \int W'(X) y\left(\phi^*\left[\left\{\mu_{nR}^*\right\}, \left\{\mu_{nD}^*\right\}; X\right]\right) \sum C\left(\mu_{n,s} + \phi^*\left[\left\{\mu_{nR}^*\right\}, \left\{\mu_{nD}^*\right\}; X\right]\right) dX.$$

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

The proof for party D follows precisely the parallel logic. □

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

$$W_n = \frac{x^n}{x^n + (1-x)^n}$$
$$W_n' = \frac{((x-1)x)^n (\log x - \log(1-x))}{(x^n + (1-x)^n)^2}$$

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

$$\max_{\{\mu_{dR}\}} \int \left[\frac{W_n'(x)}{W_n'\left(\frac{1}{2}\right)} y(\phi^*(x)) \sum C\left(\mu_{n,s} + \phi^*\left(\left\{\mu_{dD}^*\right\}, \left\{\mu_{dR}^*\right\}, x\right)\right) \right] dx$$

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

which is identical to equation (3) above but for scaling by the constant term $W_n'\left(\frac{1}{2}\right)$. But as $n \to \infty$, the weights

$$\lim_{n \to \infty} \frac{W_n'(x)}{W_n'\left(\frac{1}{2}\right)} \to \begin{cases} 0 & x \neq \frac{1}{2} \\ 1 & x = \frac{1}{2} \end{cases}.$$

Thus the necessary conditions are simply that

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

These are the same necessary conditions as if party R simply maximized the number of seats won at critical value $\phi^*\left(\left\{\mu^*_{nD}\right\},\left\{\mu^*_{nR}\right\},\frac{1}{2}\right)$, which could be written

$$\max_{\{\mu_{nR}\}}\sum_{nR}C\left(\mu_{n,s}+\phi^*\left(\left\{\mu^*_{nD}\right\},\left\{\mu^*_{nR}\right\},\frac{1}{2}\right)\right)$$

$$\text{such that }\left\{\mu_{nR}\right\}\in\Omega_R.$$

□

Proof of Proposition 4 First consider a state redistricted by party R. Suppose $N_s=2$. Following Corollary 1, there are two FOCs that combine to imply

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

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

$$\mu_1^*=\arg\max_{\{\mu_1\}}\left\{C\left(\mu_1+\phi^*\right)+C\left(\mu_2\left(\mu_1\right)+\phi^*\right)\right\}$$

$$\Rightarrow\frac{c\left(\mu_1+\phi^*\right)}{c\left(\mu_2+\phi^*\right)}=-\frac{d\mu_2\left(\mu_1\right)}{\mathrm{d}\mu_1}.$$

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

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

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 $\frac{\partial\mu_1^*}{\partial\phi^*}<0$.

To prove the result for $N_s\geq 2$, note that we maximize the objective function

$$\max_{\{\mu_{nR}\}}\sum_{nR}C\left(\mu_n+\phi^*\right).$$

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

$$\frac{c(\mu_1 + \phi^*)}{\sum_{n\neq 1} c(\mu_n + \phi^*)} = \frac{\Delta\mu_{-1}}{\Delta\mu_1}. \tag{4}$$

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 the $N_s = 2$ case 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 (4) is also decreasing in ϕ^* . Therefore, we know that $\frac{\partial\mu_1^*}{\partial\phi^*} < 0$. A parallel argument establishes that $\frac{\partial\mu_N^*}{\partial\phi^*} > 0$. The argument for states controlled by party D follows exactly the same logic. □

Proof of Example 1 We can rewrite the expected number of seats won by party R as

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

and the expected value function for party R as

$$EW = \int W\big(k\big(\lambda\overline{\mu}_R + (1-\lambda)\overline{\mu}_D + \phi\big)\big)\mathrm{d}Y(\phi).$$

This equation shows that $\overline{\mu}_R$ is a sufficient statistic for the impact of party R districts on the aggregate outcome. Therefore each party does best simply to maximize the average of the median voters in the districts in their control. There is no strategic interaction at all between the parties in this special case. As a result, the share of districts λ under the control of party R can have no impact on the optimal gerrymander. □

Proof of Proposition 5 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 y means that we are left with the term $\int W'(x)c\big(\mu_i + \phi^*(x)\big)\mathrm{d}x$, 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 4 applies. □

Example 2 Suppose that the objective function takes on the form of the double-discontinuity function

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

Let $N_s = 2$, and ϕ^* be distributed uniformly for simplicity. Suppose that c is log-concave. By Proposition 3, we know that the ratio

$$DDR = \frac{c\left(\mu_1 + \phi^*\left(\frac{1}{3}\right)\right) + c\left(\mu_1 + \phi^*\left(\frac{2}{3}\right)\right)}{c\left(\mu_2 + \phi^*\left(\frac{1}{3}\right)\right) + c\left(\mu_2 + \phi^*\left(\frac{2}{3}\right)\right)} \tag{5}$$

must be monotonic in λ for the comparative static to hold. By 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 (5) is increasing. For instance, suppose that the following values hold for $\lambda_H > \lambda_L$.

	$\frac{c\left(\mu_1+\phi^*\left(\frac{1}{3}\right)\right)}{c\left(\mu_2+\phi^*\left(\frac{1}{3}\right)\right)}$	$\frac{c\left(\mu_1+\phi^*\left(\frac{2}{3}\right)\right)}{c\left(\mu_2+\phi^*\left(\frac{2}{3}\right)\right)}$	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, it is important to note that 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.

References

Bailey, D., Katz, J.: The impact of majority–minority districts in congressional elections. Cal Tech Working Paper (2005)

Besley, T., Preston, I.: Electoral Bias and Policy Choice. Quarterly Journal of Economics. **112**(4), 1473–1510 (2007)

Cameron, C., Epstein, D., O'Hallaran, S.: Do majority–minority districts maximize substantive black representation in congress? Am. Polit. Sci. Rev. **90**(4), 794–812 (1996)
Coate, S., Knight, B.: Socially optimal districting: a theoretical and empirical exploration. Q. J. Econ. **122**(4), 1409–1471 (2007)
Cox, A.B., Holden, R.T.: Rethinking Racial and Partisan Gerrymandering. Univ. Chicago Law Rev. **78**, 553–604 (2011)
Epstein, D., O'Hallaran, S.: Measuring the impact of majority–minority voting districts. Am. J. Polit. Sci. **43**(2), 367–395 (1999)
Friedman, J.N., Holden, R.T.: Optimal Gerrymandering: sometimes pack but never crack. Am. Econ. Rev. **98**(1), 113–144 (2008)
Friedman, J.N., Holden, R.T.: The rising incumbent reelection rate: What's Gerrymandering got to do with it? J. Polit. **71**(2), 593–611 (2009)
Gilligan, T.W., Matsusaka, J.G.: Structural constraints on Partisan Bias under the efficient Gerrymander. Public Choice. **100**(1/2), 65–84 (1999)
Gilligan, T.W., Matsusaka, J.G.: Public choice principles of redistricting. Public Choice. **129**(3), 381–398 (2005)
Gimpel, J.G., Lee, F.E., Kaminski, J.: The political geography of campaign contributions in American politics. J. Polit. **68**(3), 626–639 (2006)
Gul, F., Pesendorfer, W.: Strategic redistricting. Am. Econ. Rev. **100**(4), 1616–1641 (2010)
Karlin, S.: Total positivity. Stanford Univ. Press Stanford CA **100**(4), 1616–1641 (1968)
Owen, G., Grofman, B.: Optimal Partisan Gerrymandering. Econometrica **7**(1), 5–22 (1988)
Schoenberg, I.J.: On totally positive functions, laplace integrals and entire functions of the Laguerre-Polya-Schur type. Proc. Natl. Acad.Sci. **33**, 11–17 (1947)
Schoenberg, I.J.: On Polya frequency functions. Journal d'Analyse Mathématique. **1**(1), 331–374 (1951)
Shershtyuk, K.V.: How to Gerrymander: a formal analysis. Public Choice. **95**, 27–49 (1998)
Shotts, K.: How to Gerrymander: the effect of majority–minority mandates on Partisan Gerrymandering. Am. J. Polit. Sci. **45**, 120–135 (2001)
Shotts, K.: Gerrymandering, legislative composition, and national policy outcomes. Am. J. Polit. Sci. **46**, 398–414 (2002)