THE ECONOMICS OF PARTISAN GERRYMANDERING

ANTON KOLOTILIN AND ALEXANDER WOLITZKY

Abstract. In the United States, the boundaries of congressional districts are often drawn by political partisans. In the resulting *partisan gerrymandering problem*, a designer partitions voters into equal-sized districts with the goal of winning as many districts as possible. When the designer can perfectly predict how each individual will vote, the solution is to pack unfavorable voters into homogeneous districts and crack favorable voters across districts that each contain a bare majority of favorable voters. We study the more realistic case where the designer faces both aggregate and individual-level uncertainty, provide conditions under which appropriate generalizations of the pack and crack solution remain optimal, and analyze comparative statics. All districting plans that we find to be optimal are equivalent to special cases of *segregate-pair districting*, a generalization of pack and crack where all sufficiently unfavorable voter types are segregated in homogeneous districts, and the remaining types are matched in a negatively assortative pattern. Methodologically, we exploit a mathematical connection between gerrymandering—partitioning voters into districts—and information design—partitioning states of the world into signals.

JEL Classification: C78, D72, D82
Keywords: Gerrymandering, pack and crack, assortative matching, information design

---

Date: 17th September 2020.

Kolotilin: School of Economics, UNSW Business School.
Wolitzky: Department of Economics, MIT.

The title of this paper is typeset in Gerry font (https://www.uglygerry.com/), where each letter is comprised of a current US congressional district. We thank Nikhil Agarwal, Ben Golub, Richard Holden, Hongyi Li, Stephen Morris, and Ben Olken for helpful comments and suggestions. We thank Eitan Sapiro-Gheiler for excellent research assistance. Anton Kolotilin gratefully acknowledges support from the Australian Research Council Discovery Early Career Research Award DE160100964 and from MIT Sloan's Program on Innovation in Markets and Organizations. Alexander Wolitzky gratefully acknowledges support from NSF CAREER Award 1555071 and Sloan Foundation Fellowship 2017-9633.

1. Introduction

Legislative district boundaries are drawn by political partisans under many electoral systems (Bickerstaff, 2020). In the United States, the importance of districting has recently been underscored by the rise of computer-assisted partisan districting (Newkirk, 2017) together with concerted efforts to gain and exploit control of the districting process. These trends culminated in "The Great Gerrymander of 2012" (McGhee, 2020), where the Republican party's Redistricting Majority Project (REDMAP), having previously targeted state-level elections that would give Republicans control of redistricting, aggressively redistricted several states, including Michigan, Ohio, Pennsylvania, and Wisconsin. The resulting districting plans are widely viewed as contributing to the outcome of the 2012 general election, where Republican congressional candidates won a 33-seat majority in the House of Representatives with 49.4% of the two-party vote (McGann, Smith, Latner, and Keena, 2016).[1] In light of these developments—along with the Supreme Court ruling in *Rucho v. Common Cause* (2019) that partisan gerrymanders are not judiciable in federal court—partisan gerrymandering looks likely to remain a prominent feature of American politics for some time. While the Republican party raised over $30 million to fund REDMAP in the 2010 redistricting cycle, it hopes to raise $125 million for REDMAP 2020.[2]

In order to effectively define, measure, assess, and regulate gerrymandering, we must understand how a partisan designer may optimally exercise the power to gerrymander.[3] In the United States, the main constraints a gerrymanderer faces are that districts must contain the same number of voters and must be contiguous. The former constraint is obviously binding and is strictly enforced.[4] The latter constraint is likely often slack,

[1]However, it is difficult to determine how much of the gap between the Republican vote share and seat share is attributable to intentional gerrymandering rather than other features of political geography, such as the concentration of Democratic voters in cities (Chen and Rodden, 2013).

[2]See `http://www.redistrictingmajorityproject.com/` for the 2012 REDMAP summary report and `https://sglf.org/issues/redistricting/` for information on REDMAP 2020.

[3]Of course, asking this question in no way endorses gerrymandering, just as investigating monopolistic behavior does not endorse monopoly. On the contrary, understanding such behavior may be an important step towards curbing it.

[4]In *Karcher v. Daggett* (1983), the Supreme Court rejected a districting plan in New Jersey with less than a 1% deviation from population equality, finding that "there are no *de minimus* population variations, which could practically be avoided, but which nonetheless meet the standard of Article I, Section 2 [of the U.S. Constitution] without justification."

since districts can be designed with extremely irregularly shapes, and is thus ignored in much of the literature on partisan gerrymandering.[5] Following the literature, we study the *partisan gerrymandering problem* of freely partitioning voters into equal-sized districts so as to win as many districts as possible.

When the designer knows exactly how each individual will vote, the solution to this problem is the well-known *pack and crack* districting plan: if the designer's party is supported by a fraction $x^0 < 1/2$ of voters, he creates a fraction $2x^0$ of "cracked" districts which he wins with 50% of the vote, and a fraction $1 - 2x^0$ of "packed" districts where he wins no votes at all. (If instead $x^0 \geq 1/2$, the designer can make all districts identical and win them all.) We instead consider the more realistic case where the designer faces uncertainty both about the total number of votes he will receive (*aggregate uncertainty*) and about which voters will vote for him conditional on the aggregate vote (*individual uncertainty*). We are of course not the first to notice the importance of uncertainty in gerrymandering: in 1922, the political scientist Charles Merriam wrote of gerrymandering that "The shifts in party vote make it a dangerous practice, which sometimes recoils on the heads of those who undertake it," (cited in Owen and Grofman, 1988). But, as we will see, previous analyses focused on specific forms of uncertainty, yielding important but disparate results. We instead develop a general framework that unifies the literature, while also establishing new results on the form of optimal districting plans and new comparative statics.

Our analysis leverages a tight mathematical connection between partisan gerrymandering and information design. In information design (recently surveyed by Bergemann and Morris, 2019 and Kamenica, 2019), a designer partitions states of the world into signals, in order to induce a favorable action by a receiver in as many states of the world as possible. In partisan gerrymandering, a designer partitions voters into equal-sized districts, in order to induce a favorable majority vote in as many districts as possible. For example, pack and crack districting (without uncertainty) is equivalent to optimal information design with a binary state, as in the well-known "prosecutor–judge" example of Kamenica and Gentzkow (2011). In turn, gerrymandering under uncertainty

[5]See Friedman and Holden (2008) for more discussion of these constraints. The title of this paper, typeset in Gerry font (`https://www.uglygerry.com/`), contains many examples of irregularly shaped districts. An additional constraint is that in reality gerrymanderers do not assign individual voters to districts, but rather small groups of voters, such as census blocks. Thus a "voter" in our model should be interpreted as the smallest group of voters that can actually be districted.

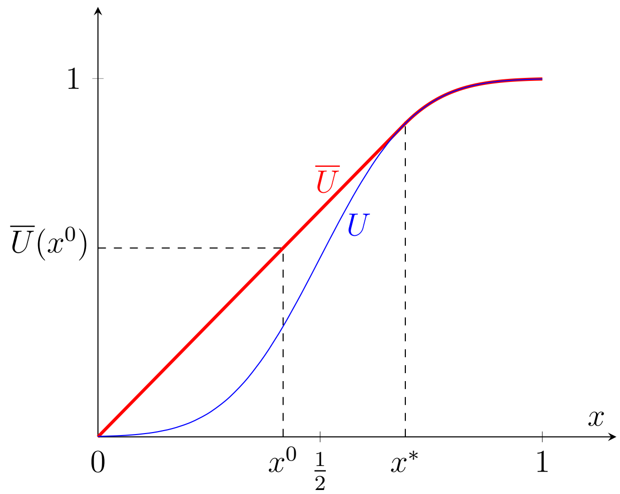


FIGURE 1. An Example of Packing and Cracking under Uncertainty

is akin to information design with an informed receiver, where the receiver's private information is analogous to the aggregate shock facing the gerrymanderer. Figure 1 illustrates a simple example of this connection. Here, each voter is either a "supporter" or an "opponent," and $U(x)$ is the probability that the designer wins a district with x supporters.[6] When the share of supporters in the population x^0 exceeds the critical value x^*, the designer's expected seat share is maximized by creating uniform districts, each with x^0 supporters. When instead $x^0 < x^*$ (as in the figure), the optimal plan creates "cracked" districts with x^* supporters and "packed" districts with 0 supporters, and the designer's expected seat share equals $\overline{U}(x^0)$.[7] This simple graphical analysis immediately reproduces the main insight of the important work of Owen and Grofman (1988).

We consider a standard electoral model with one-dimensional voter types (parameterizing a voter's probability of voting for the designer's party, i.e., a political spectrum ranging from "extreme opponents" to "extreme supporters") and one-dimensional aggregate uncertainty (parameterizing the designer's aggregate vote share), where the designer partitions voters into districts prior to the realization of aggregate uncertainty. We first analyze several benchmark cases that capture the key forces behind optimal

[6]In information design, $U(x)$ would be the probability that the receiver takes the sender's preferred action when the signal induces posterior belief x.

[7]Note that if aggregate uncertainty vanishes, in that U converges to the step function $\mathbf{1}\{x \geq 1/2\}$, then x^* converges to $1/2$, and we recover classical pack and crack districting.

districting (Section 3). Here we show that some districting plans that have previously been viewed as alternatives to pack and crack (such as the "matching extremes" plan of Friedman and Holden, 2008) are better seen as special cases of pack and crack. We also delineate the precise connection between gerrymandering and information design, which forms the basis for the rest of our analysis.

The core of the paper (Sections 4–6) then analyzes optimal districting when the designer aims to maximize his expected seat share. An insight here is that all optimal districting plans identified in our paper and the prior literature are equivalent to special cases of *segregate-pair districting*, where all sufficiently unfavorable voter types are segregated in homogeneous districts, and the remaining types are matched in a negatively assortative pattern.[8] We also consider comparative statics, for example establishing that a designer facing greater aggregate uncertainty creates a more conservative districting plan with more packed districts and fewer, more secure cracked districts, and that a less popular designer benefits more from controlling the districting process. The latter result gives one possible explanation for why in recent decades the Republicans—typically the electorally less popular party—have pursued gerrymandering much more aggressively than the Democrats.

Finally, in Section 7 we consider more general objective functions for the designer, such as weighted averages of expected seat share and the probability of controlling a majority of seats. One result here is that as the designer puts more weight on controlling a majority, he creates more packed districts and fewer cracked districts.

We build closely on three prior papers on optimal partisan gerrymandering: Owen and Grofman (1988), Friedman and Holden (2008), and Gul and Pesendorfer (2010). Owen and Grofman's model is equivalent to the special case of our model with two voter types (e.g., Democrats and Republicans). In this setting, they characterize optimal pack and crack schemes under uncertainty, deriving results like those captured in Figure 1 (see also Corollary 2). Gul and Pesendorfer consider competition between two designers who each control districting in some area and aim to win a majority of seats.[9] A simplified version of their model with a single designer is equivalent to the special case of our model where the probability that a given voter votes for the designer is

[8] As we will see, segregate-pair districting is a kind of hybrid (or mutual generalization) of Friedman and Holden (2008)'s matching extremes plan and Gul and Pesendorfer (2010)'s "p-segregation" plan.

[9] Friedman and Holden (2020) study designer competition in the model of their 2008 paper.

linear in the voter's type.[10] Under their assumptions, optimal districting plan takes a *segregate-pool* form, where voters with types below a threshold are segregated, and those with higher types are pooled into uniform districts.[11] This result is a special case of our Proposition 6, which provides necessary and sufficient conditions for the optimality of segregate-pool districting. Friedman and Holden present a general model that is similar to ours (and in particular allows non-linear vote shares), but their main results focus on the special case where individual uncertainty is small, but non-zero. We characterize optimal districting without individual uncertainty in Proposition 3; the plans characterized by Friedman and Holden are a subset of these. We also give conditions for the same plans to arise with large individual uncertainty (Corollary 4).[12]

The broader literature on gerrymandering and redistricting addresses a wide range of issues, including geographic constraints on gerrymandering (Sherstyuk, 1998; Shotts, 2001; Puppe and Tasnádi, 2009), socially optimal districting (Gilligan and Matsusaka, 2006; Coate and Knight, 2007; Bracco, 2013), district compactness (Chambers and Miller, 2010; Fryer and Holden, 2011; Ely, 2019), the interaction of redistricting and policy choices (Shotts, 2002; Besley and Preston, 2007), measuring gerrymandering (Grofman and King, 2007; McGhee, 2014; Stephanopoulos and McGhee, 2015; Duchin, 2018; Gomberg, Pancs, and Sharma, 2020), and assessing the consequences of redistricting (among many: Gelman and King, 1994; McCarty, Poole, and Rosenthal, 2009; Hayes and McKee, 2009). As the partisan gerrymandering problem interacts with many of these issues, our analysis may facilitate future research in these areas.

2. Model

There is a continuum of voters. Each voter has a type $s \in [0,1]$, which is observed by the designer. Let F denote the population distribution of voter types. The aggregate shock is denoted by $r \in \mathbb{R}$; its distribution is denoted by G. The share of type-s voters who vote for the designer when the aggregate shock takes value r is deterministic and is denoted by $v(s,r) \in [0,1]$. We assume that the distributions F and G are continuously differentiable with strictly positive densities f and g; however, we allow discrete

[10] This linear case is itself a generalization of the binary-type case considered by Owen and Grofman.

[11] Gul and Pesendorfer call this pattern *p-segregation*.

[12] As we discuss in Section 5, Friedman and Holden also develop two lemmas that apply more generally, which are very similar to our Proposition 7.

distributions in some benchmark cases. We also assume that $v(s,r)$ is continuously differentiable, strictly increasing in s, and strictly decreasing in r (outside the benchmark case where $v(s,r) = \mathbf{1}\{s \geq r\}$ for all s and r). That is, higher-type voters are stronger supporters of the designer (in that they vote for him with higher probability for every realization of the aggregate shock r), and higher realizations of the aggregate shock are less favorable for the designer (in that they reduce the probability with which each type of voter votes for him.)

The designer thus faces uncertainty about two types of objects: the aggregate vote share (determined by r), and individual voters' votes conditional on their type and the aggregate vote share (captured by $v(s,r)$, the probability that a type-s voter votes for the designer given aggregate shock r). The latter form of uncertainty may be interpreted as concerning voter "taste shocks," where each type-s voter independently draws a preference parameter $t \in \mathbb{R}$ according to a distribution $Q(t|s)$, and votes for the designer iff $t \geq r$. With this interpretation, we have

$$Q(r|s) = 1 - v(s,r) \text{ for all } (s,r).$$ [13]

We will consider the cases where either or both types of uncertainty are absent or small, as well as the case where both types of uncertainty are substantial.

The designer allocates voters among a continuum of equal-sized districts based on their types s, which determines a distribution P of s in each district. As a district is characterized by the distribution P of voter types it contains, a *districting plan*—which specifies the measure of districts with each voter type distribution P—is a distribution $\mathcal{H}$ over distributions P of s, such that the population distribution of s is given by F: thus, $\mathcal{H} \in \Delta\Delta[0,1]$ and

$$\int P d\mathcal{H}(P) = F.$$

For example, under *uniform districting*, where all districts are the same, $\mathcal{H}$ assigns probability 1 to $P = F$. In the opposite extreme case of *segregation*, where each district consists entirely of one type of voter, every distribution P in the support of $\mathcal{H}$ takes the form $P = \delta_s$ for some $s \in [0,1]$, where δ_s denotes the degenerate distribution on voter type s.

[13]Technically, for $Q(t|s)$ to be a proper probability distribution for all s, this interpretation requires the additional assumption that $\lim_{r\to-\infty} v(s,r) = 1$ and $\lim_{r\to\infty} v(s,r) = 0$ for all s.

The designer wins a district iff he receives a majority of the district vote. Thus, if s has distribution P in a district, the designer wins this district iff r satisfies $\int v(s,r)dP(s) \geq 1/2$. Denote the threshold value of the aggregate shock below which the designer wins a district with voter type distribution P by

$$r^*(P) = \max\left\{r : \int v(s,r)dP(s) \geq 1/2\right\}.$$ [14]

Note that, whenever the designer wins a district with distribution P, he also wins all districts with distributions P' satisfying $r^*(P') \geq r^*(P)$. Our model thus reflects what Grofman and King (2007, p. 12) call "a key empirical generalization that applies to all elections in the U.S. and most other democracies: the statewide or nationwide swing in elections is highly variable and difficult to predict, but the approximate rank order of districts is highly regular and stable."[15]

The designer's utility from winning a measure $m \in [0,1]$ of districts is $W(m)$, where W is an increasing function, and the designer evaluates lotteries over m by expected utility. Much of the literature focuses on two specific designer objective functions: the *majoritarian objective*, where there exists a threshold seat share $k \in (0,1)$ (often $1/2$) such that $W(m) = \mathbf{1}\{m \geq k\}$, and the *proportional objective*, where $W(m) = m$ for all $m \in [0,1]$. That is, under the majoritarian objective the designer maximizes the probability of winning at least k seats, while under the proportional objective the designer maximizes his expected seat share. Outside the majoritarian case, we assume that W is continuously differentiable with strictly positive derivative w.

Under a districting plan $\mathcal{H}$, the designer's utility when the realized value of the aggregate shock is r equals

$$W\left(\int_P \mathbf{1}\{r \leq r^*(P)\}d\mathcal{H}(P)\right).$$

[14] We employ the convention that $r^*(P) = \infty$ if $\int v(s,r)dP(s) \geq 1/2$ for all r, and $r^*(P) = -\infty$ if $\int v(s,r)dP(s) < 1/2$ for all r.

[15] Of course, the rank order of districts does change during rare political realignments. Such realignments could be allowed in a yet more general model. Our model instead allows a very general form of swings, but not realignments.

Thus, the designer's problem is

$$\max_{\mathcal{H}\in\Delta\Delta[0,1]} \int_r W\left(\int_P \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}(P)\right) dG(r)$$
$$\text{s.t.} \int_P P d\mathcal{H}(P) = F.$$

This problem nests the optimal gerrymandering problems of Owen and Grofman (1988), Friedman and Holden (2008), and (with a single designer) Gul and Pesendorfer (2010).[16] Unfortunately it is not very tractable at this level of generality: the designer is choosing a high-dimensional object (a distribution of distributions of voter types) to maximize an objective that is non-linear in the shares of districts of different types $d\mathcal{H}(P)$. We will therefore consider different cases of the model (i.e., different assumptions on F, G, v, and W), focusing on the most realistic cases (including those where aggregate or individual uncertainty is small), cases that nest the prior literature, and cases that yield relevant comparative statics.

We can first note a couple immediate comparative statics, which may help clarify the model. First, if the designer becomes more popular, in that F, G, or v shifts in his favor, then his expected utility increases under an optimal districting plan. Here, F becomes more favorable if $F(s)$ decreases for all s: this lets the designer increase the distribution of voter types s in each district (in the first-order stochastic dominance sense), which increases the cutoff aggregate shock r^* below which the designer wins each district, and hence increases the designer's expected utility. Similarly, v becomes more favorable if $v(s, r)$ increases for all (s, r): for any districting plan, this increases the cutoff aggregate shock r^* below which the designer wins each district. Finally, G becomes more favorable if $G(r)$ increases for all r: this makes more favorable aggregate shocks more likely, and thus increases the probability that the designer wins each district under any districting plan.

Second, under the taste shock interpretation of $v(s, r)$, the designer's expected utility under an optimal districting plan increases if he becomes better-informed about voter preference parameters t. Recall that $(\tilde{F}, \tilde{Q})$ is a *garbling* of (F, Q) if there exists a family of conditional densities $c(\tilde{s}|s)$ such that $\tilde{q}(t|\tilde{s})\tilde{f}(\tilde{s}) = \int c(\tilde{s}|s)q(t|s)f(s)ds$ for all $(\tilde{s}, t)$.

[16] Another apparent difference with Gul and Pesendorfer (2010) is that they consider the majoritarian objective with district-level uncertainty in addition to the aggregate shock. However, once one conditions on the pivotal value of the aggregate shock, the district-level uncertainty in Gul and Pesendorfer plays the same role as the aggregate shock in our model.

Intuitively, $\tilde{s}$ simply adds noise to s and is thus less informative about t. The designer is better-off under (F, Q) because any districting plan that is feasible under $(\tilde{F}, \tilde{Q})$ is also feasible under (F, Q), as the designer can stochastically relabel voter types from s to $\tilde{s}$ according to $c(\tilde{s}|s)$ and then allocate voters among districts based on the relabelled types.[17] Note also that the designer's information does not affect his expected vote share. Therefore, any measure of gerrymandering that increases with the designer's seat share and decreases with his vote share (e.g., partisan bias, mean-median gap, or efficiency gap) increases when the designer becomes better-informed.[18]

3. Benchmark Cases

We begin by considering several benchmark cases. These simple cases convey many of the main ideas in this paper as well as the prior literature. As we discuss in Section 3.6, they also illustrate the main forces governing the designer's problem and can help frame much of our subsequent analysis.

3.1. **Classical Pack and Crack.** We start with the classical case without uncertainty, where it has long been known that optimal districting plans "pack and crack."

Proposition 1. *Assume there is no aggregate or individual uncertainty: there exists r^0 such that $r = r^0$ with certainty, and $v(s, r^0) = \mathbf{1}\{s \geq r^0\}$ for all s. Denote the fraction of the designer's "supporters" by $x^0 = 1 - F(r^0)$. There are two cases.*

(1) *If $x^0 \geq 1/2$, a districting plan is optimal iff it creates measure 1 of districts satisfying $\Pr_P(s \geq r^0) \geq 1/2$. Under such a plan, the designer wins all districts.*
(2) *If $x^0 < 1/2$, a districting plan is optimal iff it creates measure $2x^0$ of "cracked" districts satisfying $\Pr_P(s \geq r^0) = \Pr_P(s < r^0) = 1/2$ and measure $1 - 2x^0$ of "packed" districts satisfying $\Pr_P(s < r^0) = 1$. Under such a plan, the designer wins the cracked districts.*

Case (1) says that a designer with majority support in the population may choose any districting plan where he retains majority support in every district. One such optimal plan is uniform districting. Case (2) is the classical pack and crack solution

[17] This result is essentially Blackwell's theorem (Blackwell, 1953). See also Proposition 4 of Friedman and Holden (2008) and Theorem 4 of Gul and Pesendorfer (2010).

[18] For definitions of these measures see, for example, Duchin (2018).

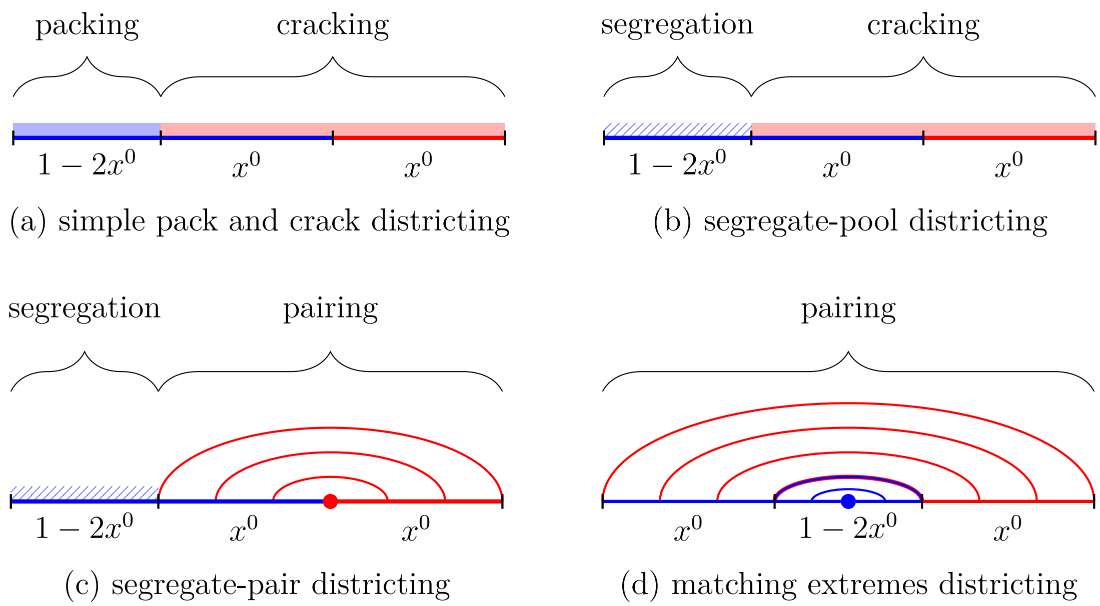


Figure 2. Four Varieties of Pack and Crack Districting

Notes: The designer wins red districts and loses blue ones. Solid shading indicates pooling; hatched shading indicates segregation.

for a designer with minority support: a designer supported by fraction $x^0 < 1/2$ of the population with no uncertainty creates $2x^0$ districts that he wins with 50% of the vote and $1 - 2x^0$ districts where he receives no votes at all.

In the special case where voter types are binary, the classical pack and crack plan is unique: it creates $2x^0$ districts with 50% favorable voters and $1 - 2x^0$ districts with no favorable voters. However, there are many optimal plans when voter types are continuous: for example, some favorable voter types can be assigned to only a subset of cracked districts, and some unfavorable voter types can be assigned only to packed districts. See Figure 2. This seemingly pedantic point will become important when we recognize that optimal plans under a small amount of uncertainty converge to some but not all classical pack and crack plans. Indeed, one of our main themes is that much can be learned about gerrymandering under uncertainty by asking how introducing a little uncertainty refines the large set of pack and crack plans that are optimal under perfect information.

With a finite number N of districts and $x^0 < 1/2$, the designer optimally wins $\lfloor 2x^0 N \rfloor$ districts, where $\lfloor \cdot \rfloor$ denotes the round-down function.[19] Our continuum district assumption lets us ignore such integer problems. We will clarify how our results extend to the finite-district case when appropriate.

If there is almost no aggregate or individual uncertainty—a case that may be fairly realistic in some elections—then a simple variation of a classical pack and crack plan is approximately optimal. Suppose that $r \approx r^0$ with high probability and $v(s,r) \approx \mathbf{1}\{s \geq r\}$ for $s \neq r$. If $1 - F(r^0) > 1/2$, the designer wins all districts with high probability under uniform districting. If $1 - F(r^0) < 1/2$, the designer can win $2(1 - F(r^0 + \varepsilon))/(1 + 2\varepsilon)$ districts with high probability, for arbitrarily small $\varepsilon > 0$, by composing each of these districts of fraction $1/2 + \varepsilon$ voters with $s > r^0 + \varepsilon$ and fraction $1/2 - \varepsilon$ voters with lower types; moreover, this plan is approximately optimal because under any districting plan the designer loses at least $1 - 2(1 - F(r^0 - \varepsilon))/(1 - 2\varepsilon)$ districts with high probability. A similar comment applies to the districting plans characterized in Propositions 2 and 3 in the following subsections: similar variations of these plans are approximately optimal when there is almost no aggregate uncertainty (Proposition 2) or almost no individual uncertainty (Proposition 3).

3.2. **No Aggregate Uncertainty.**

We next consider the case with individual uncertainty but no aggregate uncertainty.

Proposition 2. *Assume there is no aggregate uncertainty: there exists r^0 such that $r = r^0$ with certainty. There are two cases.*

(1) *If $\int v(s,r^0)dF(s) \geq 1/2$, a districting plan is optimal iff it creates measure 1 of districts satisfying $\int v(s,r^0)dP(s) \geq 1/2$. Under such a plan, the designer wins all districts.*

(2) *If $\int v(s,r^0)dF(s) < 1/2$, let s^* be the unique solution to $\int_{s^*}^1 (v(s,r^0) - 1/2)dF(s) = 0$. A districting plan is optimal iff it creates measure $1 - F(s^*)$ of cracked districts satisfying $\Pr_P(s \geq s^*) = 1$ and $\int_{s^*}^1 v(s,r^0)dP(s) = 1/2$, and measure $F(s^*)$ of packed districts satisfying $\Pr_P(s < s^*) = 1$. Under such a plan, the designer wins the cracked districts.*

[19] Obtaining a majority in $\lfloor 2x^0 N \rfloor$ districts requires $\lfloor 2x^0 N \rfloor / 2N$ favorable voters. The remaining $x^0 - \lfloor 2x^0 N \rfloor / 2N$ favorable voters are of no use to the designer and can be assigned to any district.

In Case (1), the designer wins all districts under uniform districting. In Case (2), an optimal plan assigns all voters with type $s > s^*$ to favorable districts, so that the designer wins exactly 50% of the vote in every favorable district, and assigns the remaining voters to unfavorable districts in an arbitrary way. The intuition is that the designer maximizes the measure of favorable districts by assigning only voters with types above s^* to these districts: if any such voters were replaced by voters with types below s^*, then the designer would win less than 50% of the vote in some previously-won district. Note that, since the designer wins 50% of the vote in each favorable district and the total vote share in the unfavorable districts is as small as possible, the pack and crack vote share pattern is approximated as closely as possible given the designer's inability to perfectly predict how each individual will vote. The resulting vote share pattern seems more realistic than the pattern under pack and crack: when aggregate uncertainty is small or absent, the designer can closely target a 50% vote share in favorable districts by adjusting the shares of likely-favorable and likely-unfavorable voters, but he cannot target a 0% vote share in unfavorable districts because every type of voter votes for him with some positive probability.

When individual uncertainty vanishes, $\int v(s, r^0)dF(s) \geq 1/2$ iff $x^0 \geq 1/2$, and if $x^0 < 1/2$ then s^* satisfies $1 - F(s^*) = 2x^0$. Thus, when $x^0 < 1/2$, the optimal plans characterized in Proposition 2 limit to a subset of the pack and crack plans characterized in Proposition 1. In particular, some of these optimal plans limit to the pack and crack plans depicted in Figures 2(a), 2(b), and 2(c), but none of them limits to the pack and crack plan depicted in Figure 2(d). Thus, plans of the forms depicted in Figures 2(a), 2(b), and 2(c) can be optimal with individual uncertainty but no aggregate uncertainty, while plans of the form depicted in Figure 2(d) cannot.

3.3. **No Individual Uncertainty.** We now turn to the case with aggregate uncertainty but no individual uncertainty.

Proposition 3. *Assume there is no individual uncertainty:* $v(s, r) = \mathbf{1}\{s \geq r\}$ *for all* (s, r). *Denote the median voter type by* $s^m = F^{-1}(1/2)$. *A districting plan is optimal iff for each district* $P \in \mathrm{supp}(\mathcal{H})$ *there exists a voter type* $s^P \geq s^m$ *such that* $\Pr_P(s = s^P) = \Pr_P(s < s^m) = 1/2$. *Under such a plan, the designer wins district* P *iff* $r \leq s^P$.

That is, for each voter type s above the median, the designer creates a district consisting of 50% voters with this type and 50% voters with below-median types. The intuition is easy to see with a finite number N of districts. With no individual-level uncertainty, the probability that the designer wins a given district is entirely determined by the median voter type in that district.[20] The strongest district the designer can possibly create is formed by combining the $1/2N$ voters with the highest types with any other voters: that is, it is impossible to create a district where the median voter type is above the $1-1/2N$ quantile of the population distribution. Similarly, it is impossible to create n districts where the median voter type is everywhere above the $1-n/2N$ quantile of the population distribution. But, by creating districts one at time according to the greedy algorithm that always combines the $1/2N$ remaining voters with the highest types with $1/2N$ voters with below-median types, the designer ensures that the median voter type in the n^{th} strongest district is exactly the $1-n/2N$ quantile. So this plan is optimal.[21]

The optimal plans in Proposition 3 are a subset of the classical pack and crack plans characterized by Proposition 1. In particular, if $s^m \geq r^0$ then $\Pr_P(s \geq r^0) \geq 1/2$ for every district P, so the optimal plans in Proposition 3 are optimal in Case (1) of Proposition 1. If instead $s^m < r^0$, then $\Pr_P(s \geq r^0) = 1/2$ for a set of districts P of measure $2x^0$, so the optimal plans in Proposition 3 are optimal in Case (2) of Proposition 1. For example, the pack and crack plan depicted in Figure 2(d) remains optimal when $v(s,r) = \mathbf{1}\{s \geq r\}$ but r is not degenerate, while the plans depicted in Figures 2(a), 2(b), and 2(c) are not optimal in this setting. In general, the optimal plans in Proposition 3 have the feature that, for every realization of aggregate uncertainty r, either the designer wins every district, or the designer wins some districts with exactly 50% of the vote and wins zero votes in all other districts. This is exactly the pack and crack vote share pattern. Optimal districting plans in the absence of individual uncertainty are thus a simple refinement of pack and crack.

[20]This is no longer true once individual uncertainty is introduced. For example, if $v(s,r)$ is linear in s as in Section 4, then the probability that the designer wins a district is determined by the *mean* voter type in that district, as in probabilistic voting models with uniform taste shocks (Hinich, Ledyard, and Ordeshook, 1973).

[21]This argument relies on our assumption that $v(s,r)$ is monotone (hence, single-crossing) in r for each s, so that each voter who votes for the designer at aggregate shock r also votes for him at lower aggregate shocks $r' < r$. Under this assumption, we can order types by the aggregate shock at which they switch from voting for the designer to voting against him.

The main result of Friedman and Holden (2008) characterizes optimal districting plans when individual uncertainty is sufficiently small (but non-zero), under some additional assumptions which we discuss in Section 5. As individual uncertainty vanishes, the plans they characterize limit to a subset of those characterized in Proposition 3, which in turn are a subset of pack and crack plans. The optimal plans in Friedman and Holden's environment are thus a refinement of pack and crack.[22]

3.4. **Majoritarian Objective.** While the preceding benchmarks considered particular forms of designer uncertainty, the next two consider particular designer objectives, starting with the majoritarian objective.

Proposition 4. *Assume there exists $k \in (0,1)$ such that $W(m) = \mathbf{1}\{m \geq k\}$. Let $s^k = F^{-1}(1-k)$ and let r^k be the unique solution to $\int_{s^k}^1 (v(s, r^k) - 1/2)dF(s) = 0$. A districting plan is optimal iff it creates measure k of cracked districts satisfying $\Pr_P(s \geq s^k) = 1$ and $\int_{s^k}^1 v(s, r^k)dP(s) = 1/2$, and measure $1-k$ of packed districts satisfying $\Pr_P(s < s^k) = 1$. Under such a plan, the designer wins the cracked districts iff $r \leq r^k$.*

With the majoritarian objective, the designer maximizes the probability of winning at least fraction k of the districts, or equivalently districts representing fraction k of the voters. The designer thus concentrates the k most favorable voters into k districts; moreover, to ensure that he wins all of these districts for the widest possible range of realizations of aggregate uncertainty, these districts should be equally favorable.[23]

3.5. **Proportional Objective.** Now consider the proportional objective.

Proposition 5. *Assume that $W(m) = m$ for all $m \in [0,1]$. The designer's problem becomes*

$$\max_{\mathcal{H} \in \Delta\Delta[0,1]} \int_P G(r^*(P))d\mathcal{H}(P)$$
$$s.t. \int_P P d\mathcal{H}(P) = F.$$

[22] Note that in every optimal plan in Proposition 3, all voters with the highest type s are assigned to the same district: in Friedman and Holden's words, "one's most ardent supporters should be grouped together." This is what Friedman and Holden mean when they write that "cracking is never optimal" and summarize their findings as "sometimes pack, but never crack."

[23] In the case of binary voter types, this result was established by Owen and Grofman (1988).

This follows because when $W(m) = m$ the designer's objective is a double integral over r and P, and reversing the order of integration gives the desired formulation:

$$\begin{aligned}\int_r W\left(\int_P \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}(P)\right) dG(r) &= \int_r \int_P \mathbf{1}\{r \leq r^*(P)\} d\mathcal{H}(P) dG(r) \\ &= \int_P \int_r \mathbf{1}\{r \leq r^*(P)\} dG(r) d\mathcal{H}(P) \\ &= \int_P G(r^*(P)) d\mathcal{H}(P).\end{aligned}$$

Note that the resulting program is linear in the probabilities $d\mathcal{H}(P)$, unlike in the general case with non-linear W. This is a substantial simplification, which we will impose for the rest of the paper (except for Section 7, where we consider how our results extend to more general designer objectives). Importantly, the designer's problem is now equivalent to a Bayesian persuasion problem, where the designer's utility of inducing posterior distribution P is $G(r^*(P))$.[24] Much of the Bayesian persuasion literature gets traction by assuming that the receiver's utility depends on P only through its mean $\mathbb{E}_P[s]$ (Gentzkow and Kamenica, 2016; Kolotilin, Mylovanov, Zapechelnyuk, and Li, 2017; Kolotilin, 2018; Dworczak and Martini, 2019). This corresponds to assuming that $v(s, r)$ is linear in s, as we do in Section 4.[25]

3.6. **Summary of Benchmark Cases and a Look Ahead.** The benchmark cases analyzed thus far reveal two key forces that determine the structure of the optimal gerrymander: the familiar pack and crack force, which is clearest without aggregate uncertainty; and an "assortative" force, which is clearest without individual uncertainty.

With only individual uncertainty, an optimal plan packs the most extreme opposing voters in losing districts and cracks the remaining voters in barely winning districts (as in Proposition 2 and Figures 2(a), 2(b), and 2(c)). There are many such optimal

[24]From this perspective, it may be interesting to note that the designer's problem with general, non-linear W is equivalent to Bayesian persuasion where the sender has smooth ambiguity preferences, as in Klibanoff, Marinacci, and Mukerji (2005), with ambiguity attitude captured by W.

[25]Without this additional assumption, the designer's problem is essentially equivalent to the *canonical* Bayesian persuasion problem as defined by Kolotilin and Wolitzky (2020), which specializes the general Bayesian persuasion problem of Kamenica and Gentzkow (2011) by assuming that the state and the receiver's action are one-dimensional, the receiver's utility is supermodular and concave in his action, and the sender's utility is independent of the state and increasing in the action.

plans, and the preceeding analysis does not indicate how to select among them once we account for aggregate uncertainty. This result does not rely on our assumptions that s and r are one-dimensional and $v(s,r)$ is monotone, and neither does our analysis in Section 4.

With only aggregate uncertainty, an optimal plan matches each voter above the population median with some voter below the median (as in Proposition 3 and Figure 2(d)). There are again many such optimal plans, and the analysis so far does not indicate how to choose among them once individual uncertainty is introduced. This result does rely on our one-dimensionality and monotonicity assumptions, as does our analysis in Section 5.

In general, it can be shown that there always exists an optimal plan where each district contains at most two voter types.[26] There are thus three questions about how to refine the optimal plans characterized in Propositions 2 and 3.

(1) How should the packed districts be broken up in Proposition 2?

(2) How should the cracked districts be broken up in Proposition 2?

(3) Which below-median types should be matched with which above-median types in Proposition 3?

Assuming a proportional objective for the designer, linear vote shares, and a natural condition on the distribution of the aggregate shock, Section 4 answers question 1: the packed districts should be further segregated into districts that each contain a single voter type. The intuition is that the designer's expected seat share *within districts containing only unfavorable voters* is maximized by segregating the different types of unfavorable voters, so that the designer has a respectable chance of winning the strongest of these districts. In particular, with individual uncertainty and a little aggregate uncertainty, the plans depicted in Figures 2(b) and 2(c) can be optimal, but the plan depicted in Figure 2(a) cannot.

Assuming non-linear vote shares under a natural curvature condition, Section 5 answers questions 2 and 3: cracked districts should be broken up into districts that each contain two voter types, where types are paired in a negatively assortative manner; similarly,

[26]Mathematically, every extreme point of the set of all districting plans has this property. Hence, there always exists an optimal plan with this property; moreover, generically, every optimal plan has this property.

below-median and above-median types should be matched in a negatively assortative manner. To see the intuition, consider two districts that are equally favorable for the designer (in that he wins them with the same probability), of which one is evenly divided between extreme supporters and extreme opponents, while the other consists entirely of moderate voters. In the first district, the designer's vote share will be close to 50% with high probability; in the second, his vote share will fluctuate wildly with the aggregate shock. On the margin, it is therefore optimal for the designer to allocate an extra supporter to the first district (where she is pivotal with high probability) and an extra opponent to the second district (where she is rarely pivotal), which strengthens the first district and weakens the second. An optimal districting plan thus contains of a mix of districts comprised of both extreme supporters and extreme opponents and districts comprised of more moderate voters, where the designer wins districts with more extreme voters with higher probability. In particular, among the plans depicted in Figure 2, only the plan in Figure 2(c) can be optimal with individual uncertainty and a little aggregate uncertainty; and only the plan in Figure 2(d) can be optimal with aggregate uncertainty and a little individual uncertainty.[27]

Finally, note that the plan in Figure 2(d) is also an example of the segregate-pair form depicted in Figure 2(c), where the interval of segregated types is empty. Indeed, all optimal districting plans in the literature are examples of segregate-pair plans. In Section 5, we also give an example where the unique optimal plan takes this form.

4. Linear Vote Shares

In this section, we analyze the general proportional-objective model of Section 3.5 under the assumption that $v(s, r)$ is linear in s:

$$v(s, r) = (1 - s)v(0, r) + sv(1, r) \quad \text{for all } (s, r).$$

Note that this specification nests the case where voter types are binary, $\text{supp}(F) = \{0, 1\}$. It also allows aggregate uncertainty to be small or absent (as in Section 3.2), but it allows individual uncertainty to be small or absent (as in Section 3.3) only when voter types are binary.[28]

[27]The plan in Figure 2(b) can thus be optimal with individual uncertainty and a little aggregate uncertainty if $v(s, r)$ is linear in s, but it cannot be optimal if $v(s, r)$ is even slightly non-linear in s.

[28]One situation where linear vote shares may be especially realistic is when a "voter" in our model corresponds to a small group of voters such as a census tract, individual voters' types are binary, and the type s of a group of voters is the share of its members with favorable types.

The key simplification afforded by linear vote shares is that the threshold value of the aggregate shock $r^*(P)$ below which the designer wins a district with voter type distribution P depends only on the mean voter type in the district, $x = \mathbb{E}_P[s]$. The designer can thus be viewed as choosing a distribution $H(x)$ over mean types x rather than a distribution $\mathcal{H}(P)$ over distributions of types P. With this formulation, the constraint $\int P d\mathcal{H}(P) = F$ simplifies to the requirement that F is a mean-preserving spread of H, which we denote by $F \succsim H$.[29]

Slightly abusing notation, the designer wins all districts where the mean type is at least x iff the aggregate shock is at most $r^*(x)$. The probability of this event is

$$U(x) = G(r^*(x)).$$

Since $r^*(x) = \max\{r : v(x, r) \geq 1/2\}$ and v is continuously differentiable and strictly monotone, the implicit function theorem implies that U is continuously differentiable, with non-negative derivative u; moreover, $U(0) \geq 0$ and $U(1) \leq 1$. We can interpret U as the distribution of a re-scaled aggregate shock z such that the designer wins a district with mean type x iff $x \geq z$; this distribution is in principle empirically measurable and can thus be viewed as an alternative model primitive.[30] With the proportional objective, $U(x)$ may also be viewed as the designer's Bernoulli utility for creating a district with mean type x. In total, the designer's problem becomes

$$\max_{H \in \Delta[0,1]} \int U(x) dH(x)$$
$$\text{s.t. } F \succsim H.$$

We can immediately observe that uniform districting is optimal if U is concave and segregation is optimal if U is convex. Note that the designer has Bernoulli utility function U and chooses a lottery H over outcomes $x \in [0, 1]$ subject to the constraint that H is less risky than F. Thus, if U is concave, the designer is risk averse, and thus chooses the degenerate lottery that yields mean type $\mathbb{E}_F[s]$ with certainty (i.e., uniform

[29]One way to see this is by analogy to statistics, where if a state s is distributed according to F then there exists an experiment such that the distribution of posterior expectations of s is given by H iff F is a mean-preserving spread of H (e.g., Blackwell, 1953; Kolotilin, 2018).

[30]The re-scaled aggregate shock z has density $u = U'$ on $[0, 1]$. Moreover, z is below 0 with probability $U(0)$, in which case the designer wins all districts; and z is above 1 with probability $1 - U(1)$, in which case the designer wins no districts.

districting). If instead U is convex, the designer is risk loving, and thus chooses the riskiest lottery F (i.e., segregation).

A more realistic assumption is that U is *S-shaped*, so the marginal impact of replacing a less favorable voter with a more favorable one on the probability of winning a district is first increasing and then decreasing: formally, this means that there exists an inflection point $x^i \in [0,1]$ such that U is convex on $[0, x^i]$ and concave on $[x^i, 1]$. We say that U is *strictly S-shaped* if it is strictly convex on $[0, x^i]$ and strictly concave on $[x^i, 1]$; this means that the distribution of the re-scaled aggregate shock is unimodal. For example, any weighted average of the case with no aggregate uncertainty (where the marginal impact is ∞ at x^i such that $v(x^i, r^0) = 1/2$ and 0 elsewhere) and the case with uniform aggregate uncertainty (where the marginal impact is constant) induces an S-shaped U.

4.1. **Segregate-Pool Districting.** We will see that U being S-shaped is closely related to the optimality of *segregate-pool districting*, where measure m^* of voters with the highest types are pooled in districts with the same mean type x^* and measure $1 - m^*$ of voters with the lowest types are segregated. That is, for $m^* \in [0,1]$ the cutoff type above which voters are pooled is $s^* = F^{-1}(1 - m^*)$, the mean type in the pooled districts (the *pool mean*) is $x^* = \mathbb{E}_F[s|s \geq s^*]$, and the distribution of mean types is

$$H^*(x) = \begin{cases} F(x), & \text{if } x \in [0, s^*), \\ F(s^*), & \text{if } x \in [s^*, x^*), \\ 1, & \text{if } x \in [x^*, 1]. \end{cases}$$

When the value of s^* is 0 or 1, segregate-pool districting simplifies to uniform districting or segregation, respectively; we thus refer to the case where $s^* \in (0,1)$ as *non-trivial* segregate-pool districting. Note that segregate-pool districting is optimal in the classical case (Section 3.1) as well as in the cases with no aggregate uncertainty (Section 3.2) and the majoritarian objective (Section 3.4): it is the plan illustrated in Figure 2(b).

Under segregate-pool districting with cutoff s^* and pool mean $x^* = \mathbb{E}_F[s|s \geq s^*]$, the designer's expected seat share is

$$\int_0^{s^*} U(x)dF(x) + U(x^*)(1 - F(s^*)).$$

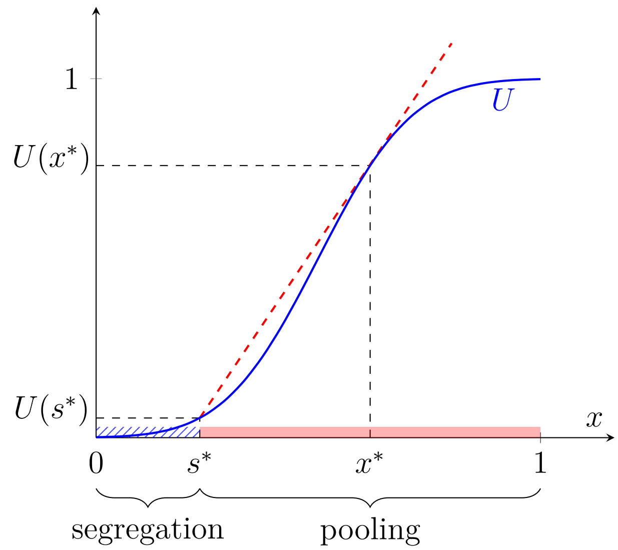


FIGURE 3. Segregate-Pool Districting

The *best* segregate-pool districting plan is the one where s^* is chosen to maximize this expectation. When the optimal value of s^* is interior, it is characterized by the first-order condition

$$u(x^*)(x^* - s^*) = U(x^*) - U(s^*).[31] \tag{1}$$

The intuition for this equation is that a marginal increase in s^* increases the pool mean, which increases the designer's expected seat share by $u(x^*)(1-F(s^*))dx^*/ds^* = u(x^*)(x^* - s^*)f(s^*)$; however, it also decreases the mass of pooled voters, which decreases the designer's expected seat share by $(U(x^*) - U(s^*))f(s^*)$. The first-order condition equates the marginal benefit and marginal cost. See Figure 3.

We now characterize when the best non-trivial segregate-pool districting plan is the optimal plan overall.

Proposition 6. *Segregate-pool districting with cutoff type $s^* \in (0,1)$ is optimal iff*

$$U(x) \leq U(x^*) + u(x^*)(x - x^*) \text{ for all } x \in [s^*, 1], \text{ with equality at } x = s^*, \tag{2}$$

$$\text{and } U \text{ is convex on } [0, s^*]. \tag{3}$$

Moreover, every optimal districting plan has the same distribution H^* of mean types if the inequality is strict for all $x \in (s^*, x^*) \cup (x^*, 1]$ and U is strictly convex on $[0, s^*]$.

[31] This equation is analogous to equation (12) of Gul and Pesendorfer (2010).

Intuitively, condition (3) implies that the designer is risk loving in the pool mean x for $x \in [0, s^*]$, so voters with types below s^* are segregated. In contrast, condition (2) implies that the designer is "on average" risk averse in x for $x \in [s^*, 1]$, so voters with types above s^* are pooled in equally favorable districts. Condition (2) holds with equality at $x = s^*$ by the first-order condition.

The following corollary links segregate-pool districting to unimodality of the re-scaled aggregate shock distribution.

Corollary 1. *If U is S-shaped then segregate-pool districting is optimal. Conversely, if U is not S-shaped then for some distribution $F \in \Delta[0,1]$ segregate-pool districting is suboptimal. Moreover, every optimal districting plan has the same distribution of mean types if U is strictly S-shaped.*

The closest prior result to Proposition 6 and Corollary 1 is Theorem 1 of Gul and Pesendorfer (2010).[32] They consider a more complex model where two competing designers each control districting in some area, but the translation of their Theorem 1 to our setting shows that if U is strictly S-shaped and symmetric about 1/2 then segregate-pool districting is optimal, with pool mean given by (1).[33] As compared to their result, we exactly characterize when segregate-pool districting is optimal, showing that U being S-shaped is sufficient but not necessary, and that symmetry about 1/2 plays no role. In addition, Lemma A2 in Appendix B characterizes optimal districting even when segregate-pool districting is suboptimal.

In the limit as aggregate uncertainty vanishes, the best segregate-pool districting plan converges to the optimal plan characterized in Proposition 2 in which the unfavorable districts are segregated. Thus, districting plans of the form depicted in Figure 2(a) (where unfavorable districts are pooled) and those of the forms depicted in Figures 2(b) and 2(c) (where unfavorable districts are segregated) are all optimal without aggregate uncertainty, but only those in Figures 2(b) and 2(c) remain optimal with a

[32]There are also counterparts to these results in the persuasion literature (e.g., Kolotilin, 2018; Kolotilin, Mylovanov, and Zapechelnyuk, 2019).

[33]Specifically, they assume that $v(1, r)$ is concave in r for $r \in [1/2, 1]$ and satisfies $v(1, 1/2) = 1$, while $v(0, r) = 1 - v(1, 1 - r)$ for all $r \in [0, 1]$. This assumption implies that r^* is strictly concave on $[1/2, 1]$ and symmetric about 1/2. Moreover, they assume that G is also strictly concave on $[1/2, 1]$ and symmetric about 1/2, and thus so is $U(\cdot) = G(r^*(\cdot))$.

small amount of aggregate uncertainty. Note that as G converges to the step function $\mathbf{1}\{r \geq r^0\}$, U converges to the step function $\mathbf{1}\{x \geq x^i\}$, where x^i is the solution to $v(x^i, r^0) = 1/2$. The first-order condition (1) then reduces to the condition that $x^* = x^i$, which yields the same condition for s^* as in Proposition 2.

Note that the plans in Figures 2(b) and 2(c) induce the same distribution of district mean types, and hence may both be optimal even when the optimal distribution of mean types is unique. Here the designer's indifference among different ways of creating favorable districts with the same mean type rests on the assumption of linear vote shares and is not robust to slight non-linearity, as we show in Section 5.

4.2. **Binary Voter Types.** We now discuss the important special case where voter types are binary: $\text{supp}(F) = \{0, 1\}$, with $\Pr_F(s = 1) = x^0$. In this case, the constraint $F \succsim H$ simplifies to $\int x dH(x) = x^0$, where $H \in \Delta[0, 1]$. Moreover, segregate-pool districting with pool mean $x^* \in [x^0, 1]$ corresponds to packing measure $1 - x^0/x^*$ of unfavorable voters into districts with mean type 0 and cracking the remaining x^0/x^* voters into districts with mean type x^*: thus, when $\text{supp}(F) = \{0, 1\}$, we use the term *pack and crack* interchangeably with segregate-pool. Under this districting plan, the distribution of the mean type is

$$H^*(x) = \begin{cases} 1 - \frac{x^0}{x^*}, & \text{if } x \in [0, x^*), \\ 1, & \text{if } x \in [x^*, 1], \end{cases}$$

and the designer's expected seat share is

$$U(0)\left(1 - \frac{x^0}{x^*}\right) + U(x^*)\frac{x^0}{x^*}.$$

In the best pack and crack plan, when the optimal value of x^* is interior, it is characterized by the first-order condition

$$u(x^*)x^* = U(x^*) - U(0).^{34} \tag{4}$$

This is simply equation (1) with $s^* = 0$. For an illustration, refer back to Figure 1 in the Introduction.

Proposition 6 and Corollary 1 specialize to the binary case as follows.

[34] This equation is analogous to equation (17) of Owen and Grofman (1988).

Corollary 2. *Assume that* $\operatorname{supp}(F) = \{0,1\}$ *and* $\Pr_F(s = 1) = x^0$. *Pack and crack with pool mean* $x^* \in (x^0, 1)$ *is optimal iff*

$$U(x) \le U(x^*) + u(x^*)(x - x^*) \text{ for all } x \in [0,1], \text{ with equality at } x = 0.$$

Moreover, the optimal districting plan is unique if the inequality is strict for all $x \notin \{0, x^*\}$. *In particular, pack and crack is optimal if* U *is S-shaped, and it is uniquely optimal if* U *is strictly S-shaped.*

Corollary 2 is related to the analysis of Owen and Grofman (1988). They derive equation (4) and argue that pack and crack is optimal in "clearly far and away the most common case," but do not formalize this point. In contrast, Corollary 2 provides necessary and sufficient conditions for the optimality of pack and crack.

More generally, when voter types are binary optimal districting can be completely characterized using the concavification approach of Aumann and Maschler (1995) and Kamenica and Gentzkow (2011). The concave envelope $\overline{U}$ of U on $[0,1]$ is the pointwise smallest concave function that is everywhere weakly greater than U. Uniform districting is optimal iff $U(x^0) = \overline{U}(x^0)$. If instead $U(x^0) < \overline{U}(x^0)$ then the point $(x^0, \overline{U}(x^0))$ lies on a linear segment of $\overline{U}$ that connects points $(x_L^*, U(x_L^*))$ and $(x_H^*, U(x_H^*))$, where $x_L^* < x^0 < x_H^*$, $U(x_L^*) = \overline{U}(x_L^*)$, and $U(x_H^*) = \overline{U}(x_H^*)$. Thus,

$$\overline{U}(x^0) = U(x_L^*)\frac{x_H^* - x^0}{x_H^* - x_L^*} + U(x_H^*)\frac{x^0 - x_L^*}{x_H^* - x_L^*},$$

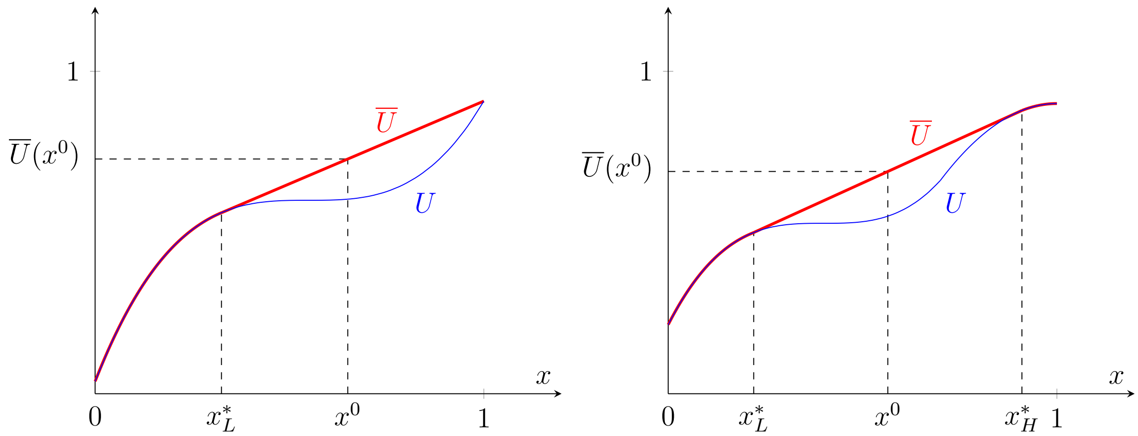
and an optimal districting plan creates measure $(x_H^* - x^0)/(x_H^* - x_L^*)$ of districts with mean type x_L^* and measure $(x^0 - x_L^*)/(x_H^* - x_L^*)$ of districts with mean type x_H^*. Such a plan coincides with pack and crack iff $x_L^* = 0$. Otherwise this plan takes the form of *crack and pack* ($0 < x_L^* < x_H^* = 1$) or *crack and crack* ($0 < x_L^* < x_H^* < 1$), as shown in Figure 4. The next result summarizes this discussion.

Corollary 3. *Assume that* $\operatorname{supp}(F) = \{0,1\}$ *and* $\Pr_F(s = 1) = x^0$. *There are two cases.*

(1) *Uniform districting is optimal iff*

$$U(x) \le U(x^0) + u(x^0)(x - x^0) \text{ for all } x \in [0,1].$$

It is uniquely optimal if the inequality is strict for all $x \ne x^0$.



(a) crack and pack

(b) crack and crack

FIGURE 4. Beyond Pack and Crack with Binary Voter Types

(2) *A districting plan that creates unfavorable districts with mean type* $x_L^* < x^0$ *and favorable districts with mean type* $x_H^* > x^0$ *is optimal iff*

$$U(x) \leq U(x_L^*)\frac{x_H^* - x}{x_H^* - x_L^*} + U(x_H^*)\frac{x - x_L^*}{x_H^* - x_L^*} \text{ for all } x \in [0, 1].$$

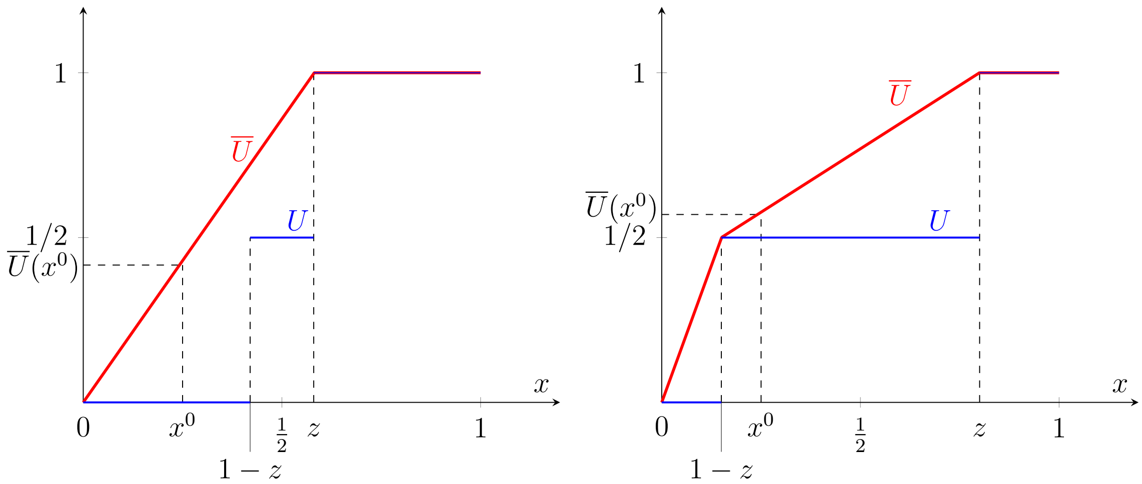
It is uniquely optimal iff the inequality is strict for all $x \notin \{x_L^*, x_H^*\}$.

It may be instructive to consider a simple parameterized example.

Example 1. Suppose the aggregate shock is binary and symmetric about 1/2 with both realizations equally likely, so, for some $z \in [1/2, 1]$,

$$U(x) = \begin{cases} 0, & \text{if } x \in [0, 1-z), \\ \frac{1}{2}, & \text{if } x \in [1-z, z), \\ 1, & \text{if } x \in [z, 1]. \end{cases}$$

Assume that $x^0 < 1/2$, so that uniform districting is never optimal. A simple application of Corollary 3 shows that if $z < 2/3$ (so the amount of aggregate uncertainty is fairly small) then the optimal districting plan is pack and crack with $x_H^* = z$. If instead $z > 2/3$ then there are two cases: when $x^0 < 1 - z$, the optimal plan is pack and crack with $x_H^* = 1 - z$; and when $x^0 > 1 - z$, the optimal plan is crack and crack with $x_L^* = 1 - z$ and $x_H^* = z$. Intuitively, if aggregate uncertainty is small then creating a "safe" district with $x = z$ does not require many more supporters than creating a



(a) small uncertainty, pack and crack

(b) large uncertainty, crack and crack

Figure 5. An Example with Bimodal Aggregate Uncertainty

"risky" district with $x = 1 - z$, so the designer chooses a conservative plan with only safe districts and maximally unfavorable districts. If instead aggregate uncertainty is large then safe districts require many more supporters than risky districts, so a designer with few supporters creates a mix of risky districts and maximally unfavorable districts, while a designer with more supporters creates a mix of risky districts and safe districts. See Figure 5.

An interesting feature of this example is that the designer's expected seat share is non-monotone in the amount of aggregate uncertainty z: it decreases with z when z is small enough that the designer chooses a conservative plan (because creating safe districts gets "more expensive" when z increases), but starts increasing with z once z is large enough that the designer switches to a risky plan (because creating risky districts gets "cheaper" when z increases). This property contrasts with Proposition 10 in Section 6, which shows that the designer's expected seat share is monotone in the amount of aggregate uncertainty when U is S-shaped.[35]

[35]Since aggregate uncertainty is likely unimodal in most elections, we view Proposition 10 as the more realistic case. The current example complements that result by showing that some unimodality-type assumption is required, and by noting that the comparative static can go the other way in an alternative case which is not totally implausible.

5. Non-Linear Vote Shares

We now relax the assumption that $v(s,r)$ is linear in s. We instead impose a natural curvature assumption on v, which will determine some key features of optimal districting plans. Such plans can then be further refined under some additional conditions.

First, we assume that a type-s voter votes for the designer given aggregate shock r iff $s-r+e \geq 0$, where $e \in \mathbb{R}$ is a voter-specific additive noise term distributed according to Q. An interpretation is that s is a signal of the voter's preference parameter $t = s + e$, which is subject to additive noise. Mathematically, this assumption implies that $Q(t|s)$ is translation-invariant and thus can be written as $Q(t-s)$.

Second, we assume that the probability density q of Q is strictly log-concave, in that the derivative of $\ln q$ is strictly decreasing. This implies that the distribution of the noise term e is unimodal, because every distribution with a strictly log-concave density is strictly S-shaped. Many common distributions have strictly log-concave densities (see Table 1 in Bagnoli and Bergstrom 2005). Note that the linear specification of Section 4 is a limit of distributions with strictly log-concave densities. Moreover, in contrast to Section 4, the current setting allows not only aggregate uncertainty but also individual uncertainty to be small or absent (as in Section 3.3), even with continuous voter types.

5.1. **Single-Dipped Districting.** We now show that these assumptions imply that every optimal districting plan partitions voters in a negatively assortative pattern that we call "single-dippedness." A districting plan is *single-dipped* if each district consists of at most two distinct voter types, and for each district P consisting of voter types $s < s''$ and each district P' containing a voter type $s' \in (s, s'')$, we have $r^*(P') < r^*(P)$. That is, more extreme voters are assigned to districts that are more favorable, in that the designer wins them with higher probability. Thus, whenever voters with types s and s'' are assigned to the same district and $s < s' < s''$, the designer's probability of winning the district containing a type $\tilde{s} \in \{s, s', s''\}$ is single-dipped on $\{s, s', s''\}$.[36]

Proposition 7. *Every optimal districting plan is single-dipped.*

[36]Segregation is vacuously single-dipped because no district contains two types $s < s''$. But neither uniform nor segregate-pool districting is single-dipped, because the designer's probability of winning districts containing types $s < s' < s''$ cannot be constant under single-dipped districting.

Proposition 7 implies that, when the taste shock density q is strictly log-concave, the designer should never pool more than two voter types in the same district. This conclusion differs starkly from the optimality of pooling an interval of voter types in Section 4, where Q is linear (so q is constant and thus weakly log-concave). If Q is instead even slightly non-linear, the designer strictly prefers to break up any district that pools an interval of voter types into districts consisting of at most two types in such a way that districts with more extreme voter types are more favorable.

To see the intuition, note that if q is log-concave, then $v(s,r)$ is more convex in s at higher values of r. Hence, if there exists a district P containing voter types $s < s''$ and a district P' containing a voter type $s' \in (s, s'')$ such that $r^*(P) < r^*(P')$, then by reallocating some voters with type s' from district P' to district P and reallocating some voters with types s and s'' from district P to district P', the designer can keep the total mass of voters in each district constant while increasing the threshold vote share $\int v(s, r^*(\tilde{P}))d\tilde{P}(s)$ for both districts $\tilde{P} \in \{P, P'\}$. This in turn increases the threshold aggregate shock for both districts, and hence increases the designer's seat share for every realized aggregate shock. Similarly, if $r^*(P) = r^*(P')$ for two districts P and P' but district P consists of more extreme voter types, then $r^*(P)$ is more sensitive to reallocating voter types than is $r^*(P')$. So it is optimal for the designer to reallocate less favorable voters from P to P' and more favorable voters from P' to P, thus strengthening the district with more extreme voters.[37]

Proposition 7 generalizes Lemmas 1 and 2 in Friedman and Holden (2008) and Proposition 2 in Kolotilin and Wolitzky (2020). It strengthens Kolotilin and Wolitzky's Proposition 2 by showing that *every* optimal districting plan is single-dipped when q is strictly log-concave.[38] In comparison, Friedman and Holden (2008)'s key assumption is that, for all $s < s'$ and $t < t'$,

$$\frac{\partial Q(t|s)}{\partial s}\frac{\partial Q(t'|s')}{\partial s} > \frac{\partial Q(t|s')}{\partial s}\frac{\partial Q(t'|s)}{\partial s},$$

[37]Symmetrically, if the taste shock density q is strictly *log-convex*, then every optimal districting plan is *single-peaked*, so that more extreme voters are assigned to *less* favorable districts. This case seems much less realistic, because it implies that the density of the noise term $e = t - s$ is single-dipped, rather than single-peaked.

[38]It also establishes a stronger notion of single-dippedness which rules out districts P and P' such that P contains types $s < s''$, P' contains a type $s' \in (s, s'')$, and $r^*(P) = r^*(P')$.

which is equivalent to strict log-concavity of q if $Q(t|s)$ takes the form $Q(t-s)$. Friedman and Holden (2008)'s Lemmas 1 and 2 are similar to Proposition 7, but they consider a finite number of districts rather than a continuum and impose an additional assumption that the mode of Q lies at the median.[39]

Single-dippedness is an important assortative feature of a districting plan, but many plans can be single-dipped. Thus, we next consider refinements of single-dippedness.

5.2. **Matching Extremes Districting.** Note that there is a unique single-dipped districting plan among the optimal plans characterized in Proposition 3. This plan takes the "matching extremes" form depicted in Figure 2(d): it assigns voters to districts in a negatively assortative manner, so that the designer's probability of winning the district containing voter type $s \in [0,1]$ is single-dipped on the entire unit interval. Specifically, the least favorable district segregates voters with the median type s^m; the most favorable district consists of 50% voters with type 0 and 50% voters with type 1; and any other district consists of 50% voters with type $s \in (0, s^m)$ and 50% voters with type $s' \in (s^m, 1)$, where s' satisfies $F(s') = 1 - F(s)$.[40] This characterization suggests that, in the limit as individual uncertainty vanishes (i.e., $Q(t-s)$ converges to the step function $\mathbf{1}\{t \geq s\}$), the optimal districting plan approximates matching extremes.

The matching extremes districting plan also plays a central role in Friedman and Holden's analysis. Their main results (Propositions 1 and 2) establish that, when individual uncertainty is sufficiently small, the optimal districting plan approximates a more permissive version of matching extremes, where multiple voter types may be segregated.[41]

We now give conditions under which matching extremes is optimal away from the small individual uncertainty limit. In particular, these conditions are satisfied if the

[39]With a finite number of districts, both Friedman and Holden's Lemmas 1 and 2 and our Proposition 7 imply that every optimal districting plan satisfies a weaker notion of single-dippedness: for any two districts P and P', $r^*(P) \neq r^*(P')$, and if $s, s'' \in \text{supp}(P)$ and $s' \in \text{supp}(P')$ then $r^*(P) > r^*(P')$.

[40]This plan approximates classical pack and crack when the support of F concentrates on $\{0,1\}$ and aggregate uncertainty vanishes. When fraction $x^0 < 1/2$ voters have types close to 1, the designer assigns these voters to approximately $2x^0$ favorable districts, along with the x^0 voters with the *lowest* types. The remaining fraction $1 - 2x^0$ voters—those with the higher types *among those voters with types close to* 0—are in turn assigned to the remaining $1 - 2x^0$ unfavorable districts.

[41]Their discussion and examples focus on the case with only a single segregated "slice" of voter types, but their results allow multiple segregated slices.

aggregate shock r is uniform (G is linear) and noise $e = t - s$ is symmetric (Q is symmetric about 0).

Corollary 4. *Assume that G is weakly concave and the median of Q is weakly higher than the mode. In every optimal districting plan, every two districts P_1 and P_2 are nested, in that P_1 consists of voter types $s_1 \leq s_1'$ and P_2 consists of voter types $s_2 \leq s_2'$ such that either $s_2 \leq s_1 \leq s_1' \leq s_2'$ or $s_1 \leq s_2 \leq s_2' \leq s_1'$.*

It is instructive to compare Corollary 4 with the results in Section 4 where Q is linear. If Q is linear, translation-invariant, and symmetric about 0, then $U = G$ and thus uniform districting is optimal whenever G is weakly concave. Corollary 4 shows that, if Q is even slightly non-linear, then the designer strictly prefers to break up uniform districts into nested districts each containing two voter types in such a way that the designer's probability of winning the district that contains voters with type $s \in [0,1]$ is single-dipped in s on the entire unit interval.

5.3. **Segregate-Pair Districting.** When the distribution of aggregate uncertainty G is not concave, multiple voter types can be optimally segregated. The following example shows that, when G is S-shaped, it may be uniquely optimal to segregate the lowest voter types and pair the remaining types in a negatively assortative manner, as in Figure 2(c). All optimal districting plans discussed in this paper and the prior literature are examples of such *segregate-pair* plans (or, in the case of segregate-pool plans, are equivalent to segregate-pair plans once the pool is broken up). Note that segregate-pair plans limit to the matching extremes plan when the interval of segregated types shrinks to 0. Segregate-pair plans also yield approximately the same expected seat share for the designer as segregate-pool plans when Q is approximately linear. Thus, while it seems difficult to provide general conditions for the optimality of segregate-pair plans outside of cases where matching extremes or segregate-pool plans are also optimal, segregate-pair plans form a general class of districting plans that nest all optimal plans in the literature and are sometimes uniquely optimal.

Example 2. Suppose that Q is symmetric about 0 and $G(r) = Q(2r - 2/3)$ for all $r \in \mathbb{R}$: e.g., $t \sim N(s, \sigma^2)$ and $r \sim N(1/3, (\sigma/2)^2)$ for some $\sigma > 0$. Further, suppose

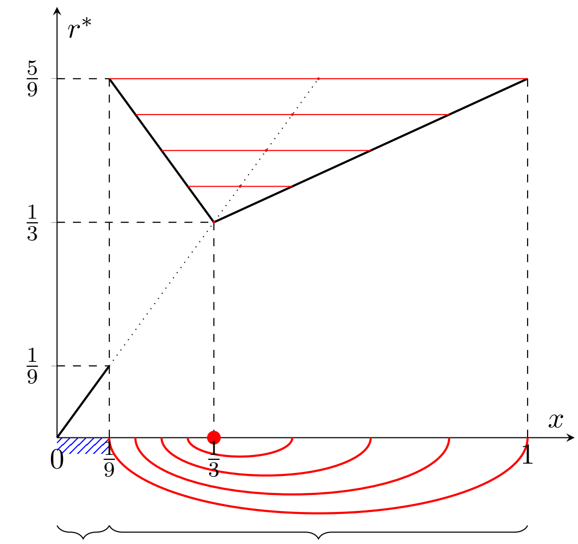


FIGURE 6. Optimal Segregate-Pair Districting in Example 2

that $f(s) = 3f(4/3 - 3s)$ for all $s \in [1/9, 1/3)$: e.g.,

$$f(s) = \begin{cases} \frac{9}{5}, & \text{if } s \in [0, \frac{1}{3}), \\ \frac{3}{5}, & \text{if } s \in (\frac{1}{3}, 1]. \end{cases}$$

In Appendix C, we establish that the unique optimal districting plan (see Figure 6) segregates voters with types $s \in [0, 1/9)$, so that $r^*(s) = s$, and creates districts consisting of 50% voters with type s and 50% voters with type $s' = 4/3 - 3s$ for each $s \in [1/9, 1/3]$, so that $r^*(s) = r^*(s') = (s + s')/2$.

A remarkable feature of this optimal plan is that the non-segregated districts all contain the same share of favorable voters. The non-segregated districts that contain more extreme voters are more favorable (r^* is higher), but this pattern is achieved entirely by making the favorable voters in more favorable districts more extreme than the unfavorable voters in these districts, rather than by allocating a higher share of favorable voters to these districts.[42]

[42] Friedman and Holden (2008) and Cox and Holden (2011) suggest that favorable districts optimally contain a strict majority of favorable voters, and that the size of this majority is smaller in districts containing more extreme voters. The current example shows that this is not always true.

6. Comparative Statics

We now consider some comparative statics of optimal districting. For simplicity, we focus on the setting depicted in Figure 1: voter types are binary (i.e., $\text{supp}(F) = \{0, 1\}$), the designer has the proportional objective, and U is strictly S-shaped, so optimal districting takes the form of pack and crack (by Corollary 2). Most results in this section extend to the more general case of linear vote shares with continuous voter types (where optimal districting takes the form of segregate-pool)—we describe these extensions as we proceed. We defer all proofs as well as detailed statements of some results to Appendix D.

Broadly speaking, we ask three questions:

(1) How do the designer's popularity, the designer's information about voter preferences, and the amount of aggregate uncertainty affect the form of optimal districting, and in particular whether the designer creates a small number of very lopsided favorable districts (a more segregated districting plan) or many less lopsided favorable districts (a less segregated plan)? (Section 6.1.)

(2) How does the amount of aggregate uncertainty affect the designer's expected seat share (i.e., his expected utility)? (Section 6.2.)

(3) How does the designer's popularity affect the *value of gerrymandering*—the designer's benefit from controlling the districting process, relative to some alternative such as uniform districting, segregation, or the districting plan that would emerge if the other party controlled districting? (Section 6.3.)

6.1. **Effects on Segregation.** To define segregation, we again adopt the taste shock interpretation of v. We say that a districting plan H is *more segregated* than another plan $\tilde{H}$ if observing a voter's district under H is Blackwell more informative about her preference parameter t. With binary voter types or linear vote shares, this condition is equivalent to H being a mean-preserving spread of $\tilde{H}$ (Blackwell, 1953). Under pack and crack or segregate-pool districting, voters are more segregated when the measure of segregated districts is higher and the measure of pooled districts m^* is lower, so the designer creates fewer, more lopsided favorable districts. In political science terminology (Tufte, 1973; King and Browning, 1987; Cox and Katz, 2002), such a plan has lower *bias* and lower *responsiveness* than a less segregated plan where favorable districts are more numerous (so the seat share is "usually" more biased towards the designer) but these districts are less secure (so the seat share is more responsive to shifts in the vote share).

We first ask how the distribution of aggregate uncertainty U affects the extent of segregation.

Proposition 8. *The optimal districting plan becomes more segregated if*

(1) *Aggregate uncertainty becomes less favorable, in that U shifts to the right.*

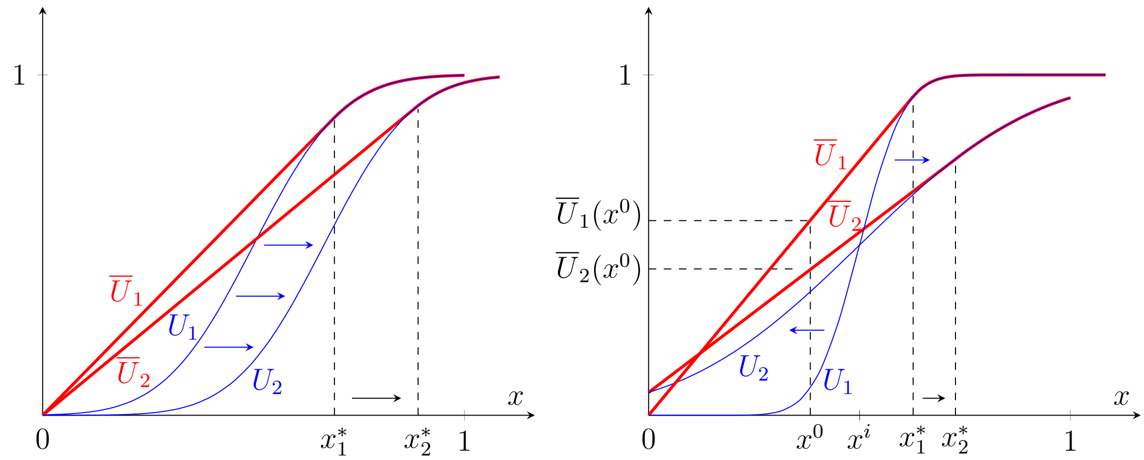
(2) *Aggregate uncertainty increases, in that U stretches horizontally around its mode.*

Proposition 8 holds for both binary and continuous voter types. To see the intuition for Part (1), note that when U shifts to the right, $U(x)$ decreases for all x and $u(x)$ increases for all x above the mode of U. See Figure 7(a). The former effect decreases the marginal cost of segregating unfavorable voters (as shifting voters from pooled districts to segregated districts has a smaller effect on the probability of representing these voters), while the latter effect increases the marginal benefit of segregating unfavorable voters (as the probability of winning the pooled districts is more responsive to the mean voter type in these districts). So the designer segregates more.[43]

The intuition for Part (2) is that an increase in aggregate uncertainty causes the designer to resolve the tradeoff between maintaining a safer margin of support in favorable districts and creating more such districts in favor of the former. See Figure 7(b). For example, as aggregate uncertainty vanishes in that U converges to the step function $\mathbf{1}\{x \geq x^i\}$, the designer creates favorable districts with mean voter type x^i and almost always wins these districts. In contrast, the designer would win these districts only half of the time when aggregate uncertainty is substantial and symmetrically distributed around the mode x^i. So the designer segregates more unfavorable voters to make the favorable districts more secure. This result that the designer packs and cracks "more conservatively" when facing greater aggregate uncertainty seems extremely natural, but we are not aware of formal antecedents in the gerrymandering literature.[44]

[43]This result adapts Theorem 3(ii) of Kolotilin, Mylovanov, and Zapechelnyuk (2019) and is similar to Theorem 2 of Gul and Pesendorfer (2010) and Proposition 3 of Alonso and Câmara (2016).

[44]Friedman and Holden (2008) find a similar pattern in numerical examples of matching extremes plans. Our formulation adapts Theorem 3(iii) of Kolotilin, Mylovanov, and Zapechelnyuk (2019) in the persuasion literature.



(a) U shifts to the right (b) U stretches around its mode x^i

FIGURE 7. Shifting or Stretching the Distribution of Aggregate Uncertainty

We now turn to the effect of the designer's information about voter preferences on segregation. Improving the designer's information about voter preference parameters t is equivalent to taking a mean-preserving spread of F.

Proposition 9. *The optimal districting plan becomes more segregated if the designer becomes more informed about voter preferences.*

To see the intuition, note that as the designer's information improves, the unfavorable voters become "more unfavorable" but the mean voter type remains unchanged. Thus, segregating the same measure of unfavorable voters makes favorable cracked districts stronger. Consequently, a better informed designer can increase the number of favorable districts and make them stronger at the same time. When U is S-shaped, the designer indeed makes favorable districts stronger, which increases segregation.

Proposition 9 is not generally true with continuous voter types. For example, suppose the improvement in the designer's information is such that, for some voters who he previously thought would vote for him with probability 25%, he receives either the bad news that they will vote for him with probability 0% or the good news that they will vote for him with probability 50%. The designer may respond to this information by adding the new "50% favorable" voters to the favorable districts, so segregation does not increase. However, the same logic as in the binary type case dictates that segregation always increases if the designer's new information takes the form of learning whether some voters with intermediate types have either the lowest or highest possible type. This situation may be realistic. For example, if the designer's information is a partition of voters into {Registered Democrat, Registered Republican, Unregistered}, then segregation increases when more voters register for a party.

6.2. **Effects on the Designer's Seat Share.** Recall from Section 2 that the designer's expected utility (with the proportional objective, his expected seat share) increases when he becomes more popular or better-informed. We now ask how his expected seat share depends on the amount of aggregate uncertainty.

Proposition 10. *Assume that $U(0) = 0$. If aggregate uncertainty increases, in that U stretches horizontally around its mode, then the designer's expected seat share decreases.*

The intuition is that aggregate uncertainty makes packing and cracking less effective: the designer must increase his margin of support in favorable districts to maintain the same win probability. Note that the assumption that $U(0) = 0$ says that aggregate uncertainty is not so great that the designer has some chance of winning a district with no supporters whatsoever, so that he continues to lose "packed" districts with probability 1.[45]

Recall from Section 4 that increasing aggregate uncertainty can increase the designer's expected seat share when aggregate uncertainty is bimodal rather than unimodal. For example, the designer may then follow a "crack and crack" strategy where he wins unfavorable districts in the event of a favorable aggregate shock, so greater aggregate uncertainty helps the designer in unfavorable districts. In contrast, when aggregate uncertainty is unimodal, the designer makes unfavorable districts as weak as possible, in which case it is reasonable to assume that the win probability in such districts is approximately 0 over any relevant range of aggregate uncertainty.

[45]Since stretching U around its mode has a discrete negative effect on the win probability in favorable districts, Proposition 10 is robust to slightly relaxing the assumption that $U(0) = 0$, as in the example illustrated in Figure 7(b). However, Proposition 10 is not generally true with continuous voter types: for example, if almost all voter types lie below the mode x^i then optimal districting segregates most of these voters, and hence the positive effect of increasing U below its mode can outweigh the negative effect of decreasing U above its mode.

6.3. **Effects on the Value of Gerrymandering.** Finally, we ask whether a weaker or stronger party benefits more from controlling the districting process. An important motivation for this question is the observation that in recent decades the less popular party in the United States (the Republicans) has pursued gerrymandering much more aggressively than the more popular party (the Democrats).[46]

We measure the value of gerrymandering as the ratio of the designer's expected seat share under an optimal districting plan to his expected seat share under one of three benchmarks: uniform districting (or a single state-wide, multi-member district), segregation (proportional representation with respect to voter types), and pessimal districting (the plan that the other party would impose if it controlled districting). We ask how the value of gerrymandering depends on the share of favorable voters x^0.

Proposition 11. *Assume that $U(0) = 0$. A less popular party benefits more from gerrymandering (i.e., the value of gerrymandering is decreasing in x^0), regardless of whether the benchmark districting plan is uniform, segregated, or pessimal.*[47]

The broad intuition is that, the fewer supporters a party has, the more carefully they must be assigned to districts to translate into seats. The results reported in the proposition are easy to see graphically, given the following observations. The designer's expected seat share is the concave envelope $\overline{U}(x^0)$ of U at x^0 under optimal districting; it is $U(x^0)$ under uniform districting; it is $x^0U(1)$ under segregation (given the assumption that $U(0) = 0$); and it is the *convex* envelope $\underline{U}(x^0)$ of U (i.e., the pointwise largest convex function that is everywhere weakly smaller than U) at x^0 under pessimal districting. Focusing for example on the uniform districting benchmark, the value of gerrymandering is thus the ratio of the concave envelope of U to U itself, which is obviously decreasing in x^0 when U is S-shaped. See Figure 8.[48]

[46]An alternative explanation for this pattern is that, since Democratic voters concentrate in cities, they are easier to segregate. See McGann, Smith, Latner, and Keena (2016) for an interesting, largely skeptical, discussion of this alternative theory.

[47]More precisely, the value of gerrymandering is strictly decreasing for $x^0 < x^*$ and constant for $x^0 > x^*$ under the uniform districting benchmark; it is constant for $x^0 < x^*$ and strictly decreasing for $x^0 > x^*$ under the segregation benchmark; and it is strictly decreasing for all x^0 under the pessimal districting benchmark.

[48]If $U(0) > 0$ then Proposition 11 holds with the same proof if we replace expected seat shares with excess expected seat shares above $U(0)$, given by $\overline{U}(x^0) - U(0)$, $U(x^0) - U(0)$, $(1 - x^0)U(0) + x^0U(1) - U(0)$, and $\underline{U}(x^0) - U(0)$ for optimal districting, uniform districting, segregation, and pessimal

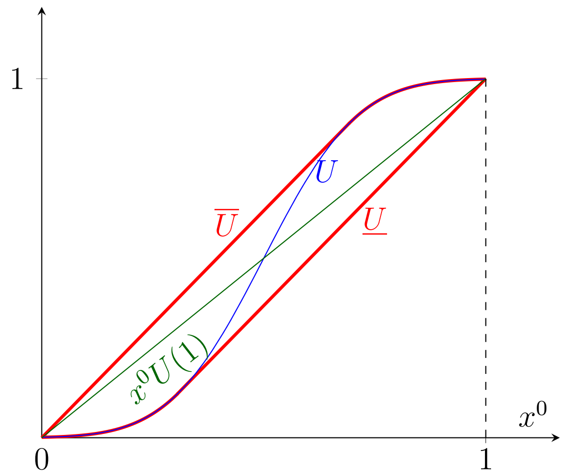


Figure 8. Expected Seat Share under Optimal Districting and Benchmarks

Notes: The value of gerrymandering (i.e., $\overline{U}(x^0)/U(x^0)$, $\overline{U}(x^0)/x^0U(1)$, or $\overline{U}(x^0)/\underline{U}(x^0)$) is decreasing in the share of supporters x^0.

7. General Designer's Objective

We now discuss how our results in Sections 4, 5, and 6 extend when the designer's objective W is non-linear. In general, most of the results in Section 5 hold for any objective W, while the results in Sections 4 and 6, which assume linear W and non-linear U, have remarkably close analogues when U is linear and W is non-linear.

7.1. **Non-Linear Vote Shares.** First, consider the results of Section 5, where $v(s,r)$ is non-linear in s. The result that every optimal districting plan is single-dipped (Proposition 7) continues to hold for general W, with essentially the same proof. Indeed, for any plan that is not single-dipped, the proof of Proposition 7 shows that there exists a reallocation of voters that increases the measure of districts won for every realization of aggregate uncertainty and thus increases the designer's expected utility for any W. Moreover, as we show in Appendix E, when G is weakly concave a "matching extremes" plan remains optimal under the additional assumption that W is weakly convex (which, as we will see, favors pooling over segregation, much like the assumption that G is weakly concave).

districting. Without normalizing the seat shares by subtracting off $U(0)$, the value of gerrymandering would be mechanically non-monotone in x^0 because $\overline{U}(0) = \underline{U}(0)$ and $\overline{U}(1) = \underline{U}(1)$, while $\overline{U}(x^0) > \underline{U}(x^0)$ for all $x^0 \in (0,1)$.

7.2. **Linear Vote Shares with Uniform Aggregate Uncertainty.** Next, consider the results of Section 4, where $v(s, r)$ is linear in s. When both U and W are non-linear, deriving conditions under which simple districting plans are optimal seems challenging, as we highlight in Section 7.4. Here we consider the case where U is linear: that is, $U(x) = (1 - x)U(0) + xU(1)$ for all $x \in [0, 1]$. This condition means that aggregate uncertainty is uniform, so the marginal impact of replacing a less favorable voter with a more favorable one on the probability of winning a district is constant. We will see that this assumption allows an analysis parallel to that of Section 4, where the curvature of W plays an analogous role to the curvature of U in that section, up to a "sign change": while concavity of U favors pooling when W is linear (and convexity of U favors segregation), we will see that convexity of W favors pooling when U is linear (and concavity of W favors segregation).

The designer's problem becomes

$$\max_{H \in \Delta[0,1]} \int W(1 - H(x))dx$$
$$\text{s.t. } F \succsim H.$$

We first argue that the curvature of W determines the form of optimal districting. Under uniform districting, all districts have the same mean type $x^0 = \mathbb{E}_F[x]$; so, for each realization of the aggregate shock, the designer wins either all districts or no districts. This is the riskiest possible distribution of the measure of districts won m. In contrast, segregation induces the safest possible distribution of m. Thus, if W is convex—so the designer is risk loving in m—he prefers uniform districting. If instead W is concave—the designer is risk averse in m—he prefers segregation.

A more realistic assumption is that W is S-shaped, so the designer's marginal utility for winning an additional district is first increasing and then decreasing. For example, any weighted average of the proportional and majoritarian objective functions is S-shaped. We will see that W being S-shaped is closely related to the optimality of segregate-pool districting (much as U being S-shaped relates to the optimality of segregate-pool in Section 4).

Under segregate-pool districting with cutoff s^*, pool mean $x^* = \mathbb{E}_F[s|s \geq s^*]$, and pool measure $m^* = 1 - F(s^*)$, the designer's expected utility is

$$\int_0^{s^*} W(1 - F(x))dx + W(m^*)(x^* - s^*) + W(0)(1 - x^*).$$

When the optimal value of s^* is interior, it is characterized by the first-order condition

$$w(m^*)m^* = W(m^*) - W(0). \tag{5}$$

The intuition for this equation is that a marginal increase in s^* increases the pool mean, which increases the designer's expected utility by $(W(m^*) - W(0))dx^*/ds^* = (W(m^*)-W(0))(x^*-s^*)f(s^*)/m^*$; however, it also decreases the share of pooled voters, which decreases the designer's expected utility by $w(m^*)(x^* - s^*)f(s^*)$. The first-order condition equates the marginal benefit and marginal cost. While this condition closely parallels condition (1) in Section 4, it is important to bear in mind that these equations involve different variables: the cutoff s^* and pool mean x^* in (1), and the pool measure m^* in (5).

We now characterize when the best non-trivial segregate-pool districting plan is the optimal plan overall.

Proposition 12. *Segregate-pool districting with pool measure $m^* \in (0,1)$ is optimal iff*

$$W(m) \leq W(m^*) + w(m^*)(m - m^*) \text{ for all } m \in [0, m^*], \text{ with equality at } m = 0, \tag{6}$$

$$\text{and } W \text{ is concave on } [m^*, 1]. \tag{7}$$

Moreover, every optimal districting plan has the same distribution H^* of mean types if the inequality is strict for all $m \in (0, m^*)$ and W is strictly concave on $[m^*, 1]$.

Intuitively, condition (7) implies that the designer is risk averse in the measure of won districts m for $m \in [m^*, 1]$, so voters with types below s^* are segregated. In contrast, condition (6) implies that the designer is "on average" risk loving in m for $m \in [0, m^*]$, so voters with types above s^* are pooled in equally favorable districts. Condition (6) holds with equality at $m = 0$ by the first-order condition.

Corollary 5. *If W is S-shaped then segregate-pool districting is optimal. Moreover, every optimal districting plan has the same distribution of mean types if W is strictly S-shaped.*

When voter types are binary ($\text{supp}(F) = \{0, 1\}$, with $\Pr_F(s = 1) = x^0$), segregate-pool districting with pool measure $m^* \in [x^0, 1]$ corresponds to packing measure $1 - m^*$ of unfavorable voters into districts with mean type 0 and cracking the remaining m^*

voters into districts with mean type x^0/m^*. Under this districting plan, the designer's expected utility is

$$W(m^*)\frac{x^0}{m^*} + W(0)\left(1 - \frac{x^0}{m^*}\right).$$

In the best pack and crack plan, when the optimal value of x^* is interior, it is again characterized by the first-order condition (5).

Proposition 12 and Corollary 5 specialize to the binary case as follows.

Corollary 6. *Assume that* $\text{supp}(F) = \{0, 1\}$ *and* $\Pr_F(s = 1) = x^0$*. Pack and crack with pool measure* $m^* \in (x^0, 1)$ *is optimal iff*

$$W(m) \leq W(m^*) + w(m^*)(m - m^*) \text{ for all } m \in [0, 1], \text{ with equality at } m = 0.$$

Moreover, the optimal districting plan is unique if the inequality is strict for all $m \in (0, m^*) \cup (m^*, 1]$*. In particular, pack and crack is optimal if* W *is S-shaped, and it is uniquely optimal if* W *is strictly S-shaped.*

As earlier papers focused exclusively on the majoritarian and proportional objectives, there are no counterparts to Proposition 12 or Corollaries 5 and 6 in the literature.[49] In addition, Lemma A11 in Appendix E characterizes optimal districting even when segregate-pool districting is suboptimal.

7.3. **Comparative Statics.** The comparative statics results of Section 6 also have analogues in the non-linear W/linear-U case, which we summarize in Proposition 13.

Proposition 13. *Assume that* $\text{supp}(F) = \{0, 1\}$*,* U *is linear, and* W *is strictly S-shaped.*

(1) *The optimal districting plan becomes more segregated if*

   (a) *The designer becomes more satisfied with fewer seats, in that* W *shifts to the left.*

[49]There are also no counterparts to these results in the persuasion literature, where non-linear W corresponds to a form of non-expected utility for the sender, as observed in Footnote 24. Mathematically, Proposition 6/Corollary 1 and Proposition 12/Corollary 5 are very different. As we show in Appendix E.2, the latter results are related to Myerson (1981)'s ironing procedure and its recent extension by Kleiner, Moldovanu, and Strack (2020).

(b) *The designer's objective becomes closer to majoritarian, in that W compresses horizontally around its mode.*

(c) *The designer becomes more informed about voter preferences.*

(2) *Assume that $W(0) = 0$. The designer's expected utility increases if his objective becomes closer to majoritarian.*

(3) *Assume that $W(0) = 0$. A less popular party benefits more from gerrymandering, regardless of whether the benchmark districting plan is uniform, segregated, or pessimal.*

Notably, Parts (1a) and (1b) show that changing W in the same way U is changed in Proposition 8 has the opposite effect on the degree of segregation of the optimal districting plan. This sign change occurs because the pool measure m^*, determined by the first-order condition on W, is inversely related to the pool mean $x^* = x^0/m^*$, determined by the first-order condition on U. Intuitively, if the designer is more satisfied with fewer seats or his objective is closer to majoritarian, he creates fewer favorable districts but makes them more secure.

7.4. **Non-Uniform Aggregate Uncertainty.** Finally, consider the case where $v(s, r)$ is linear in s, but both U and W are non-linear. The form of optimal districting is now driven by the curvatures of U and W. In particular, if U is convex and W is concave, then segregation is optimal, whereas if U is concave and W is convex, then uniform districting is optimal. More generally, the following proposition extends the characterization of segregate-pool districting in Propositions 6 and 12.

Proposition 14. *If segregate-pool districting with pool mean $x^* \in [\mathbb{E}[s], 1]$ is optimal under both $(\tilde{U}(x), \tilde{W}(x)) = (U(x), x)$ and $(\tilde{U}(x), \tilde{W}(x)) = (x, W(x))$, then it is also optimal under $(U(x), W(x))$.*

Moreover, restricting attention to segregate-pool districting, we can provide intuitive bounds on the optimal pool mean when U and W are S-shaped.

Corollary 7. *Suppose that U and W are strictly S-shaped. Let $x_U^* \in [\mathbb{E}[s], 1]$ and $x_W^* \in [\mathbb{E}[s], 1]$ be the pool means of the optimal segregate-pool districting plans under $(\tilde{U}(x), \tilde{W}(x)) = (U(x), x)$ and $(\tilde{U}(x), \tilde{W}(x)) = (x, W(x))$. In the best segregate-pool districting plan under $(U(x), W(x))$, the pool mean x^* lies between x_U^* and x_W^*.*

However, even when U and W are S-shaped, segregate-pool districting is not necessarily optimal. The following example shows that when voter types are binary and U and W are both S-shaped, an optimal districting plan may pack some districts with unfavorable voters while creating multiple cracked districts with distinct mean types that lie between x_U^* and x_W^*.

Example 3. Suppose that $x^0 = 3/8$,

$$U(x) = \begin{cases} 0, & \text{if } x \in [0, \frac{1}{2}), \\ \frac{3}{8} + x, & \text{if } x \in [\frac{1}{2}, \frac{5}{8}), \\ 1 & \text{if } x \in [\frac{5}{8}, 1], \end{cases} \quad \text{and} \quad W(m) = \begin{cases} 0, & \text{if } m \in [0, \frac{1}{2}), \\ \frac{1}{2} + \frac{m}{2}, & \text{if } m \in [\frac{1}{2}, 1]. \end{cases}$$

It is not hard to show that, in the best pack and crack plan, the designer creates measure 3/5 of cracked districts with mean voter type 5/8 and always wins these districts, while always losing the remaining districts. But there exists a strictly better plan (which is in fact optimal) that creates measure 1/2 of districts with mean type 5/8, measure 1/8 of districts with mean type 1/2, and measure 3/8 of districts with mean type 0. Intuitively, the designer's "top priority" is to win measure 1/2 of districts for sure, which requires a mean type of 5/8 in these districts. But once this priority is achieved, the designer's next objective is to maximize his expected seat share, which calls for creating as many districts as possible with mean type 1/2 from among the remaining favorable voters.

8. Conclusion

This paper has developed a simple and general model of optimal partisan gerrymandering. Our analysis unifies, clarifies, and generalizes the path-breaking but disparate analyses of Owen and Grofman (1988), Friedman and Holden (2008), and Gul and Pesendorfer (2010). We also establish a range of results without precedents in the literature, which address previously unexplored districting plans (such as *segregate-pair districting* which nests all optimal plans considered in the literature), comparative statics, and designer objectives beyond maximizing expected seat share or the probability of obtaining a legislative majority. Methodologically, we unite the gerrymandering and information design literatures, showing how powerful results developed to study information design can shed light on partisan gerrymandering.[50]

We hope our model can serve as a basis for research on various aspects of redistricting. What are the implications of our general gerrymandering model for political competition and the resulting public policies? What are the comparative statics of popular measures of gerrymandering, such as partisan bias, mean-median gap, or efficiency gap—and can our model be used to suggest new gerrymandering measures? What are the effects of introducing additional, realistic constraints on gerrymandering, such as geographic constraints or legal constraints to provide representation for communities of interest? How does competitive gerrymandering—where each party controls districting in some areas—play out? Finally, and perhaps most importantly, how should society best regulate gerrymandering?

References

Alonso, R., and O. Câmara (2016): "Political Disagreement and Information in Elections," *Games and Economic Behavior*, 100, 390–412.

Aumann, R. J., and M. Maschler (1995): *Repeated Games with Incomplete Information*. MIT press.

Bagnoli, M., and T. Bergstrom (2005): "Log-Concave Probability and Its Applications," *Economic Theory*, 26(2), 445–469.

Bergemann, D., and S. Morris (2019): "Information Design: A Unified Perspective," *Journal of Economic Literature*, 57(1), 44–95.

Besley, T., and I. Preston (2007): "Electoral Bias and Policy Choice: Theory and Evidence," *Quarterly Journal of Economics*, 122(4), 1473–1510.

Bickerstaff, S. (2020): *Election Systems and Gerrymandering Worldwide*. Springer.

Blackwell, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Bracco, E. (2013): "Optimal Districting with Endogenous Party Platforms," *Journal of Public Economics*, 104, 1–13.

[50] Conversely, information design can also learn from gerrymandering. In Kolotilin and Wolitzky (2020), we develop a theory of assortative information disclosure related to the analysis of Section 5 in the current paper, which in turn builds on Friedman and Holden (2008). Similarly, the analysis of gerrymandering with non-linear designer objectives developed here in Section 7 may hold insights for information design with non-expected sender utility.

Chambers, C. P., and A. D. Miller (2010): "A Measure of Bizarreness," *Quarterly Journal of Political Science*, 5, 27–44.

Chen, J., and J. Rodden (2013): "Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures," *Quarterly Journal of Political Science*, 8, 239–269.

Coate, S., and B. Knight (2007): "Socially Optimal Districting: A Theoretical and Empirical Exploration," *Quarterly Journal of Economics*, 122, 1409–1471.

Cox, A. B., and R. R. Holden (2011): "Reconsidering Racial and Partisan Gerrymandering," *University of Chicago Law Review*, 78, 553–604.

Cox, G. W., and J. N. Katz (2002): *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. Cambridge University Press.

Dizdar, D., and E. Kováč (2020): "A Simple Proof of Strong Duality in the Linear Persuasion Problem," *Games and Economic Behavior*, 122, 407–412.

Duchin, M. (2018): "Gerrymandering Metrics: How to Measure? What's the Baseline?," Tufts University.

Dworczak, P., and G. Martini (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127(5), 1993–2048.

Ely, J. (2019): "A Cake-Cutting Solution to Gerrymandering," Northwestern University.

Friedman, J. N., and R. Holden (2020): "Optimal Gerrymandering in a Competitive Environment," *Economic Theory Bulletin*, pp. 1–21.

Friedman, J. N., and R. T. Holden (2008): "Optimal Gerrymandering: Sometimes Pack, but Never Crack," *American Economic Review*, 98(1), 113–44.

Fryer, R. G., and R. Holden (2011): "Measuring the Compactness of Political Districting Plans," *Journal of Law and Economics*, 54, 493–535.

Gelman, A., and G. King (1994): "A Unified Method of Evaluating Electoral Systems and Redistricting Plans," *American Journal of Political Science*, 38, 514–554.

Gentzkow, M., and E. Kamenica (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review, Papers & Proceedings*, 106, 597–601.

Gilligan, T. W., and J. G. Matsusaka (2006): "Public Choice Principles of Redistricting," *Public Choice*, 129, 381–398.

Gomberg, A., R. Pancs, and T. Sharma (2020): "Electoral Maldistricting," ITAM.

Grofman, B., and G. King (2007): "The Future of Partisan Symmetry as a Judicial Test for Partisan Gerrymandering After LULAC v. Perry," *Election Law Journal*, 6, 2–35.

Gul, F., and W. Pesendorfer (2010): "Strategic Redistricting," *American Economic Review*, 100(4), 1616–41.

Hayes, D., and S. C. McKee (2009): "The Participatory Effects of Redistricting," *American Journal of Political Science*, 53, 1006–1023.

Hinich, M. J., J. O. Ledyard, and P. C. Ordeshook (1973): "A Theory of Electoral Equilibrium: A Spatial Analysis Based on the Theory of Games," *Journal of Politics*, 35(1), 154–193.

Kamenica, E. (2019): "Bayesian Persuasion and Information Design," *Annual Review of Economics*, 11, 249–272.

Kamenica, E., and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

King, G., and R. X. Browning (1987): "Democratic Representation and Partisan Bias in Congressional Elections," *American Political Science Review*, 81(4), 1251–1273.

Kleiner, A., B. Moldovanu, and P. Strack (2020): "Extreme Points and Majorization: Economic Applications," Yale University.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73(6), 1849–1892.

Kolotilin, A. (2018): "Optimal Information Disclosure: A Linear Programming Approach," *Theoretical Economics*, 13, 607–636.

Kolotilin, A., T. Mylovanov, and A. Zapechelnyuk (2019): "Censorship as Optimal Persuasion," University of St Andrews.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964.

Kolotilin, A., and A. Wolitzky (2020): "Assortative Information Disclosure," MIT.

McCarty, N., K. T. Poole, and H. Rosenthal (2009): "Does Gerrymandering Cause Polarization?," *American Journal of Political Science*, 53, 666–680.

McGann, A. J., C. A. Smith, M. Latner, and A. Keena (2016): *Gerrymandering in America: The House of Representatives, the Supreme Court, and the Future of Popular Sovereignty.* Cambridge University Press.

McGhee, E. (2014): "Measuring Partisan Bias in Single-Member District Electoral Systems," *Legislative Studies Quarterly*, 39(1), 55–85.

——— (2020): "Partisan Gerrymandering and Political Science," *Annual Review of Political Science*, 23, 171–185.

Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Newkirk, V. R. (2017): "How Redistricting Became a Technological Arms Race," *The Atlantic*, 28 October.

Owen, G., and B. Grofman (1988): "Optimal Partisan Gerrymandering," *Political Geography Quarterly*, 7(1), 5–22.

Puppe, C., and A. Tasnádi (2009): "Optimal Redistricting under Geographical Constraints: Why "Pack and Crack" Does Not Work," *Economics Letters*, 105, 93–96.

Shaked, M., and J. G. Shanthikumar (2007): *Stochastic Orders*. Springer.

Sherstyuk, K. (1998): "How to Gerrymander: A Formal Analysis," *Public Choice*, 95, 27–49.

Shotts, K. W. (2001): "The Effect of Majority-Minority Mandates on Partisan Gerrymandering," *American Journal of Political Science*, pp. 120–135.

——— (2002): "Gerrymandering, Legislative Composition, and National Policy Outcomes," *American Journal of Political Science*, pp. 398–414.

Stephanopoulos, N. O., and E. M. McGhee (2015): "Partisan Gerrymandering and the Efficiency Gap," *University of Chicago Law Review*, 82, 831–900.

Tufte, E. R. (1973): "The Relationship between Seats and Votes in Two-Party Systems," *American Political Science Review*, 67(2), 540–554.

ONLINE APPENDIX (NOT FOR PUBLICATION)

Appendix A. Proofs for Section 3

Proof of Proposition 1. (1) This case is trivial, as the designer wins all districts if he creates measure 1 of districts satisfying $\Pr_P(s \geq r^0) \geq 1/2$ and loses some positive measure of districts otherwise.

(2) Since the designer wins a district P iff $\Pr_P(s \geq r^0) \geq 1/2$, a districting plan can be described by a distribution $H \in \Delta[0,1]$ over $x = \Pr_P(s \geq r^0)$. The designer's utility for any feasible H is

$$W\left(\int \mathbf{1}\{x \geq \tfrac{1}{2}\} dH(x)\right) \leq W\left(\int 2x dH(x)\right) = W(2x^0), \tag{A1}$$

where the inequality holds because $W(m)$ is increasing in m and $\mathbf{1}\{x \geq \frac{1}{2}\} \leq 2x$ for all $x \in [0,1]$, and the equality holds because $\int x dH(x) = x^0$ for any feasible H, by the law of iterated expectations. Thus, any plan that creates measure $2x^0$ of cracked districts satisfying $\Pr_P(s \geq r^0) r^0) = 1/2$ and measure $1 - 2x^0$ of packed districts satisfying $\Pr_P(s < r^0) = 1$ is optimal. Moreover, any other plan creates a positive measure of districts with $\Pr_P(s \geq r^0) \notin \{0, 1/2\}$ (i.e., $\mathrm{supp}(H) \not\subseteq \{0, 1/2\}$), so that the inequality in (A1) is strict, because W is strictly increasing and $\mathbf{1}\{x \geq \frac{1}{2}\} = 2x$ iff $x \in \{0, 1/2\}$. So any such plan is suboptimal. $\square$

Proof of Proposition 2. (1) This case is trivial, as the designer wins all districts if he creates measure 1 of districts satisfying $\int v(s, r^0) dP(s) \geq 1/2$ and loses some positive measure of districts otherwise.

(2) Since $v(s, r^0)$ is continuously differentiable and strictly increasing in s, without loss of generality, we can redefine s as $v(s, r^0)$. (With this change of variables, the density of F does not have to be strictly positive on $[0,1]$, but this is irrelevant for our proof.) Since the designer wins a district P iff $\mathbb{E}_P[s] \geq 1/2$, a districting plan can be described by a distribution $H \in \Delta[0,1]$ over $x = \mathbb{E}_P[s]$. The designer's utility for any feasible H is

$$W\left(\int \mathbf{1}\{x \geq \tfrac{1}{2}\} dH(x)\right) = W\left(1 - H(\tfrac{1}{2}_-)\right),$$

where $H(1/2_-)$ is the left limit of H at $1/2$. Since $W(m)$ is strictly increasing in m, the proposition will follow easily from the following lemma.

Lemma A1. *Let $s^\dagger \in (0,1)$ and define $x^\dagger = \mathbb{E}_F[s|s \geq s^\dagger]$. For any feasible H, we have $1 - H(x^\dagger_-) \leq 1 - F(s^\dagger)$ with equality iff*

$$H(x) = \begin{cases} F(s^\dagger), & \text{if } x \in [s^\dagger, x^\dagger), \\ 1, & \text{if } x \in [x^\dagger, 1]. \end{cases} \tag{A2}$$

Proof. For any feasible H, the distribution F is a mean-preserving spread of H (see Footnote 29), so

$$\int_{s^\dagger}^1 (1 - F(x))dx \geq \int_{s^\dagger}^1 (1 - H(x))dx. \tag{A3}$$

Note that

$$\begin{aligned} \int_{s^\dagger}^1 (1 - F(x))dx &= -(1 - F(s^\dagger))s^\dagger + \int_{s^\dagger}^1 s dF(s) \\ &= (1 - F(s^\dagger))(x^\dagger - s^\dagger), \end{aligned} \tag{A4}$$

where the first equality is by integration by parts and the second equality is by $\mathbb{E}_F[s|s \geq s^\dagger] = x^\dagger$. Since H is a distribution function, we have

$$\begin{aligned} \int_{s^\dagger}^1 (1 - H(x))dx &= \int_{s^\dagger}^{x^\dagger} (1 - H(x))dx + \int_{x^\dagger}^1 (1 - H(x))dx \\ &\geq \int_{s^\dagger}^{x^\dagger} (1 - H(x^\dagger_-))dx + \int_{x^\dagger}^1 (1 - H(1))dx \\ &= (1 - H(x^\dagger_-))(x^\dagger - s^\dagger), \end{aligned} \tag{A5}$$

with equality iff $H(x) = H(x^\dagger_-)$ for $x \in [s^*, x^\dagger)$ and $H(x) = 1$ for $x \in [x^\dagger, 1]$. Combining (A3)–(A5) proves the lemma. □

Lemma A1 evaluated at $s^\dagger = s^*$ such that $x^\dagger = \mathbb{E}_F[s|s \geq s^*] = 1/2$ implies that H is optimal iff it satisfies (A2), which means that a districting plan creates measure $1 - F(s^*)$ cracked districts satisfying $\mathbb{E}_P[s] = 1/2$. □

Proof of Proposition 3. For a districting plan $\mathcal{H}$, define H as $H(r) = \Pr_{\mathcal{H}}(r^*(P) \leq r)$ for all r. That is, the designer wins measure $1 - H(r_-)$ of districts when the realized aggregate shock is r. For each realization r, the designer wins a district P iff it contains at least measure $1/2$ voters with types $s \geq r$ (i.e., $\Pr_P(s \geq r) \geq 1/2$). Since the population has measure $1 - F(r)$ voters with types $s \geq r$, the designer wins at most measure $2(1 - F(r))$ districts, so $1 - H(r_-) \leq 2(1 - F(r))$. Taking into account that the designer can win at most measure 1 districts implies that any feasible H satisfies $H(r_-) \geq H^\star(r)$ where

$$H^\star(r) = \begin{cases} 0, & \text{if } r \leq s^m, \\ 1 - 2(1 - F(r)), & \text{if } r > s^m. \end{cases}$$

Thus, the designer's expected utility for any feasible H is

$$\int W\left(1 - H(r_-)\right) dG(r) \leq \int W\left(1 - H^\star(r)\right) dG(r),$$

with strict inequality if $H(r) \neq H^\star(r)$ for some r (and thus on some interval (r, r') with $r' > r$ by right-continuity of H), because $W(m)$ is strictly increasing in m and $G(r)$ is strictly increasing in r. Thus, a districting plan $\mathcal{H}$ is optimal iff it induces $H^\star$, which means that (almost) every district P that the designer wins iff the aggregate shock is at most r satisfies $\Pr_P(s = r) = \Pr_P(s < s^m) = 1/2$. □

Proof of Proposition 4. Since $v(s, r^k)$ is continuously differentiable and strictly increasing in s, without loss of generality, we can redefine s as $v(s, r^k)$, similarly to the variable change in the proof of Proposition 2. Also, for a districting plan $\mathcal{H}$, let $H \in \Delta[0, 1]$ denote the induced distribution over $x = \mathbb{E}_P[s]$. By Lemma A1 evaluated at $s^\dagger = s^k$ and $x^\dagger = 1/2$, for any feasible H, when the realized aggregate shock is r^k, the designer wins at most measure $1 - F(s^k) = k$ districts, and he wins exactly measure k districts iff H satisfies (A2). Moreover, since $v(s, r)$ is strictly decreasing in r, the designer wins a strictly smaller measure than k when the realized aggregate shock is $r > r^k$. Finally, since the designer cares only about winning measure k districts, it follows that H is optimal iff it satisfies (A2), which means that a districting plan creates measure k cracked districts satisfying $\mathbb{E}_P[s] = 1/2$. □

Appendix B. Details and Proofs for Section 4

This appendix contains the proofs of the results in Section 4. It also generalizes some of these results and in particular characterizes optimal districting when it does not take the form of segregate-pool districting.

We first present an important duality result, established by Dworczak and Martini (2019) and Dizdar and Kováč (2020) in the persuasion context, which provides necessary and sufficient conditions for the optimality of a candidate districting plan.

Lemma A2. *Assume that* $F \in \Delta[0,1]$*. A distribution of district mean types* $H \in \Delta[0,1]$*, such that* $F \succsim H$*, is optimal iff there exists a continuous convex function* $\widehat{U}$ *such that* $\widehat{U}(x) \geq U(x)$ *for all* $x \in [0,1]$*, and*

$$\int U(x)dH(x) = \int \widehat{U}(x)dF(x). \tag{A6}$$

The "if" part of the result is straightforward: if such a function $\widehat{U}$ exists then, for any feasible distribution of district mean types $\tilde{H}$, we have

$$\int U(x)d\tilde{H}(x) \leq \int \widehat{U}(x)d\tilde{H}(x) \leq \int \widehat{U}(x)dF(x) = \int U(x)dH(x), \tag{A7}$$

and hence H is optimal. (Here the first inequality follows because $U(x) \leq \widehat{U}(x)$ for all x, and the second follows because $\widehat{U}$ is convex and F is a mean-preserving spread of any feasible distribution $\tilde{H}$.) Theorem 1 in Dizdar and Kováč (2020) establishes the converse under regularity conditions on U weaker than those we have imposed.

We will also use the following result, which is an immediate implication of Proposition 2 in Dworczak and Martini (2019).

Lemma A3. *Assume that* $F \in \Delta[0,1]$ *has a strictly positive density on* $[0,1]$*. Let* $\widehat{U} \geq U$ *be a continuous convex function that satisfies* (A6) *for some* $H \in \Delta[0,1]$ *such that* $F \succsim H$*. If* $\widehat{U}$ *is strictly convex on* $[0,s^*]$ *for some* $s^* \in (0,1]$*, then* $H(x) = F(x)$ *for all* $x \in [0,s^*]$*.*

To prove Lemma A3, first note that all inequalities in (A7) must hold with equality for $\tilde{H} = H$; so integrating (A7) by parts twice yields

$$0 = \int_0^1 \widehat{U}(x)dF(x) - \int_0^1 \widehat{U}(x)dH(x) = \int_0^1 \left(\int_0^x F(s)ds - \int_0^x H(s)ds\right) d\widehat{U}'(x), \tag{A8}$$

where $\widehat{U}'$ is the right derivative of $\widehat{U}$, which is non-decreasing given that $\widehat{U}$ is convex. Next, note that the integrand in (A8) is non-negative,

$$\int_0^x F(s)ds \geq \int_0^x H(s)ds \quad \text{for all } x \in [0,1],$$

because F is a mean-preserving spread of H. Thus, for (A8) to hold, the integrand must be zero almost everywhere where $\widehat{U}'$ is strictly increasing. Since $\widehat{U}$ is strictly convex on $[0,s^*]$ and F has a density, this implies that

$$\int_0^x F(s)ds = \int_0^x H(s)ds \quad \text{for all } x \in [0,s^*],$$

and thus $H(x) = F(x)$ for all $x \in [0, s^*]$.

We are now ready to establish a generalized version of Proposition 6, which characterizes conditions for the optimality of trivial and non-trivial segregate-pool districting plans. This result is a slight generalization of Proposition 3 in Kolotilin (2018) in the persuasion context, which does not provide conditions for the optimal H to be unique.

Proposition A6. *Assume that $F \in \Delta[0,1]$ has a strictly positive density on $[0,1]$.*

(1) *Segregate-pool districting with cutoff type $s^* \in (0,1]$ is optimal iff*

$$U(x) \leq U(x^*) + u(x^*)(x - x^*) \text{ for all } x \in [s^*, 1], \text{ with equality at } x = s^*,$$
$$\text{and } U \text{ is convex on } [0, s^*].$$

(2) *Uniform districting (i.e., $s^* = 0$) is optimal iff*

$$U(x) \leq U(x^*) + u(x^*)(x - x^*) \text{ for all } x \in [0,1].$$

Moreover, every optimal districting plan has the same distribution H^* of mean types if the inequality is strict for all $x \in (s^*, x^*) \cup (x^*, 1]$ and U is strictly convex on $[0, s^*]$.

To prove the "if" part, consider the function

$$\widehat{U}(x) = \begin{cases} U(x), & \text{if } x \in [0, s^*), \\ U(x^*) + u(x^*)(x - x^*), & \text{if } x \in [s^*, 1]. \end{cases} \tag{A9}$$

It is easy to see that $\widehat{U}$ is a continuous convex function that satisfies $\widehat{U} \geq U$. Moreover, (A6) holds because $\widehat{U}(x) = U(x)$ for all $x \leq s^*$ and

$$U(x^*)(1 - F(s^*)) = \int_{s^*}^{1} \widehat{U}(x) dF(x). \tag{A10}$$

Thus, by Lemma A2, segregate-pool districting H^* is optimal.

To prove the "only if" part, suppose that H^* is optimal. Then, by Lemma A2, there exists continuous convex $\widehat{U}$ such that $\widehat{U} \geq U$, $U(x) = \widehat{U}(x)$ for $x \in [0, s^*)$, and (A10) holds. Since $\widehat{U} \geq U$, $\widehat{U}$ is convex, and F has a strictly positive density on $[s^*, 1]$, Jensen's inequality applied to (A10) implies that $\widehat{U}(x)$ is linear in x on $[s^*, 1]$ and $\widehat{U}(x^*) = U(x^*)$, so $\widehat{U}$ is given by (A9). This, in turn, implies that U must satisfy the conditions of Proposition A6.

Finally, to show that H^* is uniquely optimal under the strict conditions, notice that we have shown that H^* and $\widehat{U}$ satisfy the conditions of Lemma A3. Thus, every optimal distribution H satisfies $H(x) = F(x)$ for all $x \in [0, s^*]$. Taking into account that every optimal H satisfies all inequalities in (A7) with equality gives

$$\int_{s^*}^{1} U(x)dH(x) = \int_{s^*}^{1} \widehat{U}(x)dF(x),$$

which implies that the support of H does not contain any x in $(s^*, x^*) \cup (x^*, 1]$, because $U(x) < \widehat{U}(x)$ for any such x and H has the same mean as F on $[s^*, 1]$, given that $F \succsim H$ and $H(s^*) = F(s^*)$. Thus, $H = H^*$, completing the proof of Proposition A6.

Corollary 1 follows easily from our Proposition A6 and Theorem 1 in Kolotilin, Mylovanov, and Zapechelnyuk (2019). If U is S-shaped, then it is easy to see from Figure 3 that there exists a cutoff type $s^* \in [0, 1]$ that satisfies the conditions of Proposition A6, so segregate-pool districting is optimal. Moreover, the induced distribution H^* of mean types is uniquely optimal if U is strictly S-shaped, because the strict conditions of Proposition A6 are satisfied. Finally, if U is not S-shaped, then there exist non-empty intervals (s_1, s_2) and (s_3, s_4) with $s_3 \geq s_2$ such that U is strictly concave on $[s_1, s_2]$, linear on $[s_2, s_3]$, and strictly convex on $[s_3, s_4]$. Then, as Kolotilin, Mylovanov, and Zapechelnyuk show, there exists a distribution $F \in \Delta[s_1, s_4]$ such that the uniquely optimal distribution H corresponds to non-trivial *pool-segregate districting*, where measure $m \in (0, 1)$ of voters with the lowest types are pooled in districts with the same mean type and the remaining voters are segregated.

Corollary 2 is a straightforward adaptation of Proposition 6 and Corollary 1 to the binary case. A couple remarks about uniqueness are in order. First, with binary voter types, the distribution H^* uniquely determines the districting plan $\mathcal{H}^*$. Second, the optimal distribution H^* is unique because the strict conditions of Corollary 2 together with (A6) imply that the support of H^* belongs to $\{0, x^*\}$, which uniquely determines H^*.

Corollary 3, except for the uniqueness part, follows immediately from the discussion preceding it. In Case (1), if the inequality is strict for all $x \neq x^0$, then $\text{supp}(H^*) = \{x_0\}$ which uniquely determines H^*. Similarly, in Case (2), if the inequality is strict for all $x \notin \{x_L^*, x_H^*\}$, then $\text{supp}(H^*) = \{x_L^*, x_H^*\}$, which again uniquely determines H^*. Conversely, if the inequality holds with equality at some other $x_O^* \notin \{x_L^*, x_H^*\}$, then there exists an optimal distribution H_O^* with $\text{supp}(H_O^*) = \{x_L^*, x_H^*, x_O^*\}$.

Appendix C. Details and Proofs for Section 5

This appendix contains the proofs of the results in Section 5. For simplicity, some proofs in this Appendix are written for the case of discrete voter types.

Proof of Proposition 7. The proposition rests on the following two lemmas.

Lemma A4. *No optimal districting plan has districts P and P' such that district P contains voter types $s < s''$, district P' contains a voter type $s' \in (s, s'')$, and $r^*(P) < r^*(P')$.*

Proof. Suppose for contradiction that such districts P and P' exist and let $r^*(P) = r$ and $r^*(P') = r'$, with $r' > r$. Consider a perturbation that shifts mass $\alpha = (v(s'', r) - v(s', r))\varepsilon$ of voters with type s and mass $\gamma = (v(s', r) - v(s, r))\varepsilon$ of voters with type s'' from P to P' and shifts an equal mass $\beta = \alpha + \gamma = (v(s'', r) - v(s, r))\varepsilon$ of voters with type s' from P' to P, for a sufficiently small $\varepsilon > 0$. Since $v(s, r)$ is strictly increasing in s, these masses are strictly positive and thus the perturbation is well-defined. Since this perturbation does not change the mass of voters in P and P', to show that it strictly increases the designer's expected utility, it suffices to show that $r^*(P)$ does not change and $r^*(P')$ strictly increases. First, $r^*(P)$ does not change because $\int v(s, r)dP(s)$ does not change, as

$$-v(s, r)\alpha + v(s', r)\beta - v(s'', r)\gamma = 0.$$

Second, $r^*(P')$ strictly increases because $\int v(s, r')dP'(s)$ strictly increases, as

$$\begin{aligned}
&v(s, r')\alpha - v(s', r')\beta + v(s'', r')\gamma \\
&= [(v(s'', r') - v(s', r'))(v(s', r) - v(s, r)) - (v(s'', r) - v(s', r))(v(s', r') - v(s, r'))]\varepsilon \\
&= \left[\int_{s'}^{s''}\int_{s}^{s'} \frac{\partial v(\tilde{s}', r')}{\partial s}\frac{\partial v(\tilde{s}, r)}{\partial s} d\tilde{s}d\tilde{s}' - \int_{s'}^{s''}\int_{s}^{s'} \frac{\partial v(\tilde{s}', r)}{\partial s}\frac{\partial v(\tilde{s}, r')}{\partial s} d\tilde{s}d\tilde{s}'\right]\varepsilon \\
&= \left[\int_{s'}^{s''}\int_{s}^{s'} \left\{q(r' - \tilde{s}')q(r - \tilde{s}) - q(r - \tilde{s}')q(r' - \tilde{s})\right\} d\tilde{s}d\tilde{s}'\right]\varepsilon > 0,
\end{aligned}$$

where the inequality holds because the integrand is strictly positive for $r' > r$ and $\tilde{s}' > \tilde{s}$ by strict log-concavity of q. □

Lemma A5. *No optimal districting plan has a district that contains voter types $s < s' < s''$.*

Proof. Suppose for contradiction that there exists a district P that has a strictly positive mass of voters with each of the types $s < s' < s''$. Let $r^*(P) = r$. Suppose we split district P into two identical equal-sized districts P_1 and P_2. Then consider a perturbation that shifts mass $\alpha = (v(s'', r) - v(s', r))\varepsilon$ of voters with type s and mass $\gamma = (v(s', r) - v(s, r))\varepsilon$ of voters with type s'' from P_1 to P_2 and shifts an equal mass $\beta = \alpha + \gamma = (v(s'', r) - v(s, r))\varepsilon$ of voters with type s' from P_2 to P_1, for a sufficiently small $\varepsilon > 0$. Notice that $r^*(P_2) = r^*(P_1) = r$ because

$$v(s, r)\alpha - v(s', r)\beta + v(s'', r)\gamma = 0.$$

Consider now an additional perturbation that moves an infinitesimal mass dm of voters with type s from P_2 to P_1 and moves the same mass dm of voters with type s'' from P_1 to P_2.

By the Implicit Function Theorem, $r^*(P_2) = r + dr_2 + o(dr_2)$ and $r^*(P_1) = r - dr_1 + o(dr_1)$ where

$$dr_2 = \frac{(v(s'', r) - v(s, r))}{\int q(r - s)dP_2(s)}dm \quad \text{and} \quad dr_1 = -\frac{(v(s'', r) - v(s, r))}{\int q(r - s)dP_1(s)}dm.$$

To show that this perturbation strictly increases the designer's expected utility, it suffices to show that $dr_2 > dr_1$ or equivalently $\int q(r - s)dP_2(s) < \int q(r - s)dP_1(s)$, which holds because

$$\begin{aligned}
&q(r - s)\alpha - q(r - s')\beta + q(r - s'')\gamma \\
&= [q(r - s)(Q(r - s') - Q(r - s'')) - q(r - s')(Q(r - s) - Q(r - s'')) \\
&\quad + q(r - s'')(Q(r - s) - Q(r - s'))]\varepsilon \\
&= [(q(r - s) - q(r - s'))(Q(r - s') - Q(r - s'')) \\
&\quad - (q(r - s') - q(r - s'')(Q(r - s) - Q(r - s'))]\varepsilon \\
&= \left[\int_{r-s'}^{r-s} q'(e')de' \int_{r-s''}^{r-s'} q(e)de - \int_{r-s''}^{r-s'} q'(e)de \int_{r-s'}^{r-s} q(e')de'\right]\varepsilon \\
&< \frac{q'(r - s')}{q(r - s')}\left[\int_{r-s'}^{r-s} q(e')de' \int_{r-s''}^{r-s'} q(e)de - \int_{r-s''}^{r-s'} q(e)de \int_{r-s'}^{r-s} q(e')de'\right]\varepsilon = 0,
\end{aligned}$$

where the inequality follows from strict log-concavity of q, which requires that the derivative of $\ln q$ is strictly decreasing and thus

$$\frac{q'(e)}{q(e)} > \frac{q'(r - s')}{q(r - s')} > \frac{q'(e')}{q(e')} \quad \text{for } e < r - s' < e'. \qquad \square$$

By Lemmas A4 and A5, to show that every optimal districting plan is single-dipped it suffices to show that for each district P consisting of voter types $s < s''$ and each district P' containing a voter type $s' \in (s, s'')$, we have $r^*(P') \neq r^*(P)$. Suppose for contradiction that $r^*(P') = r^*(P)$. Then merging districts P and P' into one district would also be optimal, but the merged district would contain voter types $s < s' < s''$, contradicting Lemma A5. $\square$

Proof of Corollary 4. The corollary rests on the following two lemmas.

Lemma A6. *If a single-dipped districting plan has a district that consists of voter types $s < s'$, then there is a segregated type $s^* \in (s, s')$.*

Proof. Let $\mathcal{H}$ be a single-dipped districting plan that has a district consisting of voters with types $s < s'$. Suppose first that there exists a district $P_1 \in \text{supp}(\mathcal{H})$ such that $\text{supp}(P_1) = \{s_1, s_1'\}$ and $s < s_1 < s_1' < s'$. Let $P^* \in \text{supp}(\mathcal{H})$ be a district that minimizes $r^*(P)$ over all districts $P \in \text{supp}(\mathcal{H})$ such that $\text{supp}(P) \subset [s_1, s_1']$. We claim that $\text{supp}(P^*) = s^*$ for some $s^* \in [s_1, s_1']$. Indeed, if instead $\text{supp}(P^*) = \{s^*, s^{*\prime}\}$ with $s_1 \leq s^* < s^{*\prime} \leq s_1'$, then any district $P_2 \in \text{supp}(\mathcal{H})$ such that $s_2 \in \text{supp}(P_2)$ for some $s_2 \in (s^*, s^{*\prime})$ would satisfy $\text{supp}(P_2) \subset [s^*, s^{*\prime}]$ and $r^*(P_2) < r^*(P^*)$, because $\mathcal{H}$ is single-dipped, contradicting the definition of P^*.

Suppose now that there does not exist $P_1 \in \text{supp}(\mathcal{H})$ such that $\text{supp}(P_1) = \{s_1, s_1'\}$ and $s < s_1 < s_1' < s'$. Thus, any $P^* \in \text{supp}(\mathcal{H})$ such that $s^* \in \text{supp}(P)$ for some $s^* \in (s_1, s_2)$ satisfies $\text{supp}(P^*) \subset \{s, s^*, s'\}$, because $\mathcal{H}$ is single-dipped. Since the distribution F of s has a density, it follows from Bayes's rule that for (almost) all such P^*, we have $\text{supp}(P^*) = s^*$. $\square$

Lemma A7. *If G is weakly concave and the median of Q is weakly higher than the mode, then every optimal districting plan segregates at most one voter type.*

Proof. Suppose for contradiction that exist two districts P_1 and P_2 that segregate voters with types $s_1 < s_2$. Consider a perturbation that moves an infinitesimal mass dm of voters from P_1 to P_2.

Recalling that $r^*(P)$ is defined by $\int Q(r^*(P) - s)dP(s) = 1/2$, the Implicit Function Theorem implies that $r^*(P_2) = r^*(\delta_{s_2}) - dr + o(dr)$ where

$$dr = \frac{Q(r^*(\delta_{s_2}) - s_1) - Q(r^*(\delta_{s_2}) - s_2)}{q(r^*(\delta_{s_2}) - s_2)} dm. \tag{A11}$$

To show that this perturbation strictly increases the designer's expected utility, it suffices to show that

$$[G(r^*(\delta_{s_2})) - G(r^*(\delta_{s_1}))]dm - g(r^*(\delta_{s_2}))dr > 0, \tag{A12}$$

which holds because

$$\frac{G(r^*(\delta_{s_2})) - G(r^*(\delta_{s_1}))}{g(r^*(\delta_{s_2}))(r^*(\delta_{s_2}) - r^*(\delta_{s_1}))} \geq 1 > \frac{Q(r^*(\delta_{s_2}) - s_1) - Q(r^*(\delta_{s_2}) - s_2)}{q(r^*(\delta_{s_2}) - s_2)(r^*(\delta_{s_2}) - r^*(\delta_{s_1}))},$$

where the first inequality holds because G is assumed to be weakly concave and the second inequality holds because Q is strictly concave above $r^*(\delta_{s_2}) - s_2$, as follows from the following two observations. First, $r^*(\delta_{s_2}) - s_2 = e^{med}$ where e^{med} is the median of Q, so that $Q(e^{med}) = 1/2$. Second, the distribution Q, which has a strictly log-concave density, is strictly concave above its mode and is thus also above its median, since the median is assumed to be weakly higher than the mode. □

Since the optimal districting plan is single-dipped by Proposition 7, it suffices to show that there are no districts P_1 and P_2 such that $\mathrm{supp}(P_1) = \{s_1, s_1'\}$, $\mathrm{supp}(P_2) = \{s_2, s_2'\}$, and $s_1 \leq s_1' \leq s_2 \leq s_2'$ with at least one inequality being strict. If there were such districts, then there would exist segregated types $s_1^* < s_2^*$ by Lemma A6, contradicting Lemma A7. □

Proof of Example 2. We first present a weak duality result, established in Lemma 1 of Kolotilin (2018) in the persuasion context. This result provides sufficient conditions for the optimality of a candidate districting plan. To state it, we represent a districting plan as a joint distribution J of voter type s and a threshold aggregate shock r below which the designer wins the district containing voter type s.

Let $S = [0, 1]$ denote the set of voter types. Also, let $R = [0, 1]$ denote the set of feasible threshold aggregate shocks: for each $P \in \Delta[0, 1]$, the value $r^*(P)$ belongs to $[r^*(\delta_0), r^*(\delta_1)]$, which is equal to $[0, 1]$ because Q is symmetric about 0. Thus, any feasible joint distribution J belongs to $\Delta(S \times R)$.

Lemma A8. *A joint distribution $J \in \Delta(S \times R)$ of voter type s and the threshold aggregate shock r, such that*

$$\text{the marginal distribution of } s \text{ is } F, \quad \text{and} \tag{A13}$$

$$\int_{S\times\tilde{R}} (v(s,r) - 1/2) dJ(s,r) = 0 \text{ for every measurable set } \tilde{R} \subset R, \tag{A14}$$

is optimal if there exist bounded measurable functions η and ν such that

$$\eta(s) - (v(s,r) - 1/2)\nu(r) \geq G(r) \text{ for all } (s,r) \in S \times R, \quad \text{and} \tag{A15}$$

$$\int_{S\times R} G(r) dJ(s,r) = \int_S \eta(s) dF(s). \tag{A16}$$

A joint distribution J is *feasible* iff it satisfies (A13) and (A14), where (A13) requires that the population distribution of s is given by F, and (A14) requires that the threshold aggregate shock in each district P is $r^*(P)$.

The proof of the lemma is simple: if such functions η and ν exist then, for any feasible joint distribution $\tilde{J}$, we have

$$\begin{aligned}\int_{S\times R} G(r) d\tilde{J}(s,r) &\leq \int_{S\times R} (\eta(s) - (v(s,r) - 1/2)\nu(r)) d\tilde{J}(s,r) \\ &= \int_S \eta(s) dF(s) - \int_{S\times R} (v(s,r) - 1/2)\nu(r) d\tilde{J}(s,r) \\ &= \int_S \eta(s) dF(s) \\ &= \int_{S\times R} G(r) dJ(s,r),\end{aligned} \tag{A17}$$

and hence J is optimal. Here the inequality follows from (A15), whereas the equalities follow from (A13), (A14), and (A16).

To simplify notation, we redefine s as $s - 1/3$ and r as $r - 1/3$, so that $S = R = [-1/3, 2/3]$ and $G(r) = Q(2r)$ for all $r \in \mathbb{R}$. With this notation, our candidate districting plan J segregates voter types $s \in [-1/3, -2/9)$, and pairs each voter type $s \in [-2/9, 0]$ with $s' = -3s \in [0, 2/3]$, so that

$$r^*(s) = \begin{cases} s, & \text{if } s \in [-\frac{1}{3}, -\frac{2}{9}), \\ -s, & \text{if } s \in [-\frac{2}{9}, 0), \\ \frac{s}{3}, & \text{if } s \in [0, \frac{2}{3}]. \end{cases}$$

The optimality of J will follow easily from the following lemma.

Lemma A9. *Functions η and ν given by*

$$\eta(s) = \begin{cases} Q(2s), & \text{if } s \in [-\frac{1}{3}, 0), \\ 3Q(\frac{2}{3}s) - 1, & \text{if } s \in [0, \frac{2}{3}], \end{cases} \quad \text{and} \quad \nu(r) = \begin{cases} \frac{2q(2r)}{q(0)}, & \text{if } r \in [-\frac{1}{3}, 0), \\ 2, & \text{if } r \in [0, \frac{2}{3}]. \end{cases}$$

satisfy (A15) with equality for $(s, r) \in Y$ and with strict inequality for $(s, r) \notin Y$ where

$$Y = \{(s, r^*(s)) : s \in [-\tfrac{1}{3}, \tfrac{2}{3}]\} \cup \{(s, s) : s \in [-\tfrac{1}{3}, 0]\}.$$

Proof. Denote the excess share (above 1/2) of type-s voters who vote for the designer when the aggregate shock is r by

$$V(s - r) = Q(s - r) - \tfrac{1}{2},$$

where the equality holds because $v(s, r) = 1 - Q(r - s)$ and Q is symmetric about 0. Clearly, V is also symmetric about 0, so that $V(e) = -V(-e)$ for all $e \in \mathbb{R}$.

It is easy to see that η and ν satisfy (A15) with equality for all $(s, r) \in Y$. Indeed, for $s \in [-1/3, 0]$, we have

$$\eta(s) - V(0)\nu(s) = Q(2s) = G(s).$$

Similarly, for $s \in [-2/9, 0]$, we have

$$\eta(s) - V(s - r^*(s))\nu(r^*(s)) = \tfrac{1}{2} + V(2s) - 2V(2s) = Q(-2s) = G(r^*(s)).$$

Finally, for $s \in [0, 2/3]$, we have

$$\eta(s) - V(s - r^*(s))\nu(r^*(s)) = \tfrac{1}{2} + 3V(\tfrac{2}{3}s) - 2V(\tfrac{2}{3}s) = G(r^*(s)).$$

We now verify that η and ν satisfy (A15) with strict inequality for all $(s, r) \notin Y$. There are a few cases to consider.

(1) For $r \in [0, 2/3]$ and $s \in [r, 2/3]$, the condition (A15) simplifies to

$$3V(\tfrac{2}{3}s) \geq 2V(s - r) + V(2r),$$

which holds with strict inequality for $s \neq 3r$ because V is strictly concave on $(0, \infty)$.

(2) For $r \in [0, 2/3]$ and $s \in [0, r)$, the condition (A15) simplifies to

$$3V(\tfrac{2}{3}s) + 2V(r - s) \geq V(2r) + 4V(0),$$

which always holds with strict inequality because V is strictly concave on $(0, \infty)$.

(3) For $r \in [0, 2/3]$ and $s \in [-1/3, 0)$, the condition (A15) simplifies to

$$2V(r-s) \geq V(2r) + 4V(-2s),$$

which holds with strict inequality for $s \neq -r$ because V is strictly concave on $(0, \infty)$.

(4) For $r \in [-1/3, 0)$ and $s \in [0, 2/3]$, the condition (A15) simplifies to

$$3V(\tfrac{2}{3}s) + V(-2r) \geq \nu(r)V(s-r) + 2V(0),$$

which always holds with strict inequality because $\nu(r) < 2$ and V is strictly positive and strictly concave on $(0, \infty)$.

(5) For $r \in [-1/3, 0)$ and $s \in (r, 0)$, the condition (A15) simplifies to

$$V(-2r) \geq V(-2s) + \nu(r)V(s-r),$$

which is equivalent to

$$\frac{Q(-2r) - Q(-2s)}{q(-2r)(2s-2r)} \geq \frac{Q(s-r) - Q(0)}{q(0)(s-r)}.$$

This always holds with strict inequality because Q is strictly concave on $(0, \infty)$ and thus the left-hand side is strictly greater than one whereas the right-hand side is strictly less than one.

(6) For $r \in [-1/3, 0)$ and $s \in [-1/3, r)$, the condition (A15) simplifies to

$$V(-2r) + \nu(r)V(r-s) \geq V(-2s),$$

which is equivalent to

$$\frac{Q(r-s) - Q(0)}{q(0)(r-s)} \geq \frac{Q(-2s) - Q(-2r)}{q(-2r)(2r-2s)}.$$

This always holds with strict inequality because

$$\begin{aligned}
\frac{Q(-2s) - Q(-2r)}{q(-2r)(2r-2s)} &= \frac{1}{2r-2s}\int_0^{2(r-s)} \frac{q(y-2r)}{q(-2r)} dy \\
&< \frac{1}{2r-2s}\int_0^{2(r-s)} \frac{q(y)}{q(0)} dy \\
&= \frac{Q(2r-2s) - Q(0)}{q(0)(2r-2s)} \\
&< \frac{Q(r-s) - Q(0)}{q(0)(r-s)},
\end{aligned}$$

where the first inequality holds because $q(y+c)/q(y)$ is strictly decreasing in c for a strictly log-concave q, and the second inequality holds because Q is strictly concave on $(0,\infty)$. $\square$

By Lemma A9, η and ν satisfy (A15) with equality for all $(s,r) \in \operatorname{supp}(J)$, which consists of points $(s, r^*(s))$ for $s \in S$. Taking into account (A17) implies that (A16) holds, and thus J is optimal by Lemma A8.

Moreover, by (A17), any optimal plan $\tilde{J}$ must satisfy

$$\int_{S\times R} (\eta(s) - (v(s,r) - 1/2)\nu(r) - G(r))d\tilde{J}(s,r) = 0.$$

By (A15), the integrand is non-negative, so, for the equality to hold, $\operatorname{supp}(\tilde{J})$ must be a subset of Y, where the integrand is zero. But it is easy to see that J is the unique plan satisfying this property. Indeed, since $(s', r') \in Y$ for $s' \in [0, 2/3]$ iff $r' = s'/3$, the plan $\tilde{J}$ must allocate each voter with type s' to a district where the threshold aggregate shock is $r^*(s') = s'/3$. Moreover, since $(\tilde{s}, s'/3) \in Y$, for $s' \in [0, 2/3]$, iff $\tilde{s} = s'$ or $\tilde{s} = -s'/3$, each district of the plan $\tilde{J}$ where the threshold aggregate shock is $s'/3$ must consist of $f(s')$ voters with type s' and $f(s')$ voters with type $-s'/3$. Since, by assumption, $|f(-s'/3)d(-s'/3)| = f(s')ds'$ for $s' \in [0, 2/3]$, the plan $\tilde{J}$ must match all existing voters with types s' and $-s'/3$. Finally, the plan $\tilde{J}$ must segregate the remaining voters with types $s \in [-1/3, -2/9)$ because $(s,r) \in Y$ for $s \in [-1/3, -2/9)$ iff $r = s$. So $\tilde{J}$ must coincide with J. $\square$

Appendix D. Details and Proofs for Section 6

We assume that U is a continuously differentiable and weakly increasing function from $\mathbb{R}$ to $[0, 1]$. To allow for the possibility that the designer loses "packed" districts with certainty (as in Propositions 9 and 11), we assume that $U(x) = 0$ for all $x \leq \underline{x}$ and some $\underline{x} < 1$. Finally, we assume that U is strictly S-shaped on $(\underline{x}, \infty)$. Specifically, there exists an inflection point $x^i \in (\underline{x}, 1)$ such that U is strictly convex on $(\underline{x}, x^i)$ and strictly concave on (x^i, ∞).

We first note that a simple generalization of Corollary 1 gives the existence of a unique optimal plan that takes the form of pack and crack if $\operatorname{supp}(F) = \{s_L, s_H\}$ with $s_H > \underline{x}$, which is assumed in all propositions in this appendix. Moreover, we henceforth also assume, except in Proposition A9, that $\operatorname{supp}(F) = \{0, 1\}$ with $\Pr_F(s = 1) = x^0$.

We now present a simple condition that in the binary case is equivalent to the statement that one districting plan is more segregated than another.

Lemma A10. *Let distributions H and $\tilde{H}$ have the same mean x^0 and at most binary supports $\{x_L, x_H\}$ and $\{\tilde{x}_L, \tilde{x}_H\}$, with $x_L \leq x_H$ and $\tilde{x}_L \leq \tilde{x}_H$. Then H is a mean-preserving spread of $\tilde{H}$ iff $x_L \leq \tilde{x}_L \leq \tilde{x}_H \leq x_H$.*

Proof. By definition, H is a mean-preserving spread of $\tilde{H}$ iff

$$\int_{-\infty}^{x} (H(s) - \tilde{H}(s))ds \geq 0 \text{ for all } x \in \mathbb{R} \text{ with equality at } x \to \infty. \tag{A18}$$

Thus, $x_L \leq \tilde{x}_L$, otherwise

$$\int_{-\infty}^{x} H(s)ds = 0 < \int_{-\infty}^{x} \tilde{H}(s)ds \text{ for all } x \in (\tilde{x}_L, x_L),$$

and, similarly, $x_H \geq \tilde{x}_H$. Conversely, let $x_L \leq \tilde{x}_L \leq \tilde{x}_H \leq x_H$. Then (A18) holds with equality for $x \geq x_H$, because H and $\tilde{H}$ have the same mean. Moreover, (A18) holds for $x \leq \tilde{x}_L$, because $H(x) \geq \tilde{H}(x) = 0$ for $x < \tilde{x}_L$. Similarly, (A18) holds for $x \geq \tilde{x}_H$, because $H(x) \leq \tilde{H}(x) = 1$ for $x \geq \tilde{x}_H$ and (A18) holds with equality for $x \geq x_H$. Finally, (A18) holds for $x \in (\tilde{x}_L, \tilde{x}_H)$ because $\int_{-\infty}^{x}(H(s) - \tilde{H}(s))ds$ is linear in x on $(\tilde{x}_L, \tilde{x}_H)$ and (A18) holds for $x \in \{\tilde{x}_L, \tilde{x}_H\}$, as shown above. □

A formal version of Proposition 8 takes the following form.

Proposition A8.

(1) *Let x^*_λ be the pool mean of the optimal pack and crack plan under $U_\lambda(x) = U(x - \lambda)$ where $\lambda < 1 - \underline{x}$. Then x^*_λ is increasing in λ, strictly so if $x^*_\lambda \in (x^0, 1)$.*

(2) *Let x^*_ζ be the pool mean of the optimal pack and crack plan under $U_\zeta(x) = U(x^i + (x - x^i)/\zeta)$ where $\zeta > 0$. Then x^*_ζ is increasing in ζ, strictly so if $x^*_\zeta \in (x^0, 1)$.*

In the persuasion context, Kolotilin, Mylovanov, and Zapechelnyuk (2019) establish Proposition 8 for a general $F \in \Delta[0, 1]$. To see why Part (1) holds, consider $x^*_\lambda \in (x^0, 1)$ and notice that by Corollary 2, we have

$$U_\lambda(0) = U_\lambda(x^*_\lambda) - u_\lambda(x^*_\lambda)x^*_\lambda.$$

Because U is strictly S-shaped on $(\underline{x}, \infty)$, it is easy to see that for $\lambda' < \lambda$ and $x' \geq x^*_\lambda$,

$$U_{\lambda'}(0) < U_{\lambda'}(x') - u_{\lambda'}(x')x'.$$

Thus, the equality can be restored only at $x^*_{\lambda'} < x^*_\lambda$. The argument for Part (2) is analogous, except that the corresponding strict inequality holds for $\zeta' \in (0, \zeta)$ and $x' \geq x^*_\zeta$, because U is both strictly S-shaped and increasing on $(\underline{x}, \infty)$.

A formal version of Proposition 9 takes the following form.

Proposition A9. *Let distributions F and $\tilde{F}$ have the same mean x^0 and binary supports $\{s_L, s_H\}$ and $\{\tilde{s}_L, \tilde{s}_H\}$ with $s_L < \tilde{s}_L < \tilde{s}_H < s_H$ and $\tilde{s}_H > \underline{x}$. Moreover, let x^* and $\tilde{x}^*$ be the pool mean of the optimal pack and crack plan under F and $\tilde{F}$. Then $x^* > \tilde{x}^*$, unless $x^* = \tilde{x}^* = x^0$.*

Proof. The proposition trivially holds if $x^* = s_H$. By Corollaries 2 and 3, if $x^* < s_H$, then

$$U(x) \leq U(x^*) + u(x^*)(x - x^*) \text{ for all } x \in [s_L, s_H].$$

Because U is strictly S-shaped on $(\underline{x}, \infty)$, it is easy to see that for $\tilde{x} \geq x^*$ and $\tilde{s}_L \in (s_L, \tilde{x})$, we have

$$U(\tilde{s}_L) < U(\tilde{x}) + u(\tilde{x})(\tilde{s}_L - \tilde{x}).$$

Thus, in the optimal pack and crack plan, either $\tilde{x}^* < x^*$ or $\tilde{x}^* = x^* = x^0$. □

A formal version of Proposition 10 takes the following form.

Proposition A10. *Let $\overline{U}_\zeta(x^0)$ be the expected seat share of the optimal pack and crack plan under $U_\zeta(x) = U(x^i + (x - x^i)/\zeta)$ where $\zeta \in (0, x^i/(x^i - \underline{x})]$, so that $U_\zeta(0) = 0$. Then $\overline{U}_\zeta(x^0)$ is strictly decreasing in ζ.*

Proof. In the optimal pack and crack plan under ζ, the pool mean is $x^*_\zeta > x^i$ and

$$\overline{U}_\zeta(x^0) = U_\zeta(0)\frac{x^*_\zeta - x^0}{x^*_\zeta} + U_\zeta(x^*_\zeta)\frac{x^0}{x^*_\zeta} = U_\zeta(x^*_\zeta)\frac{x^0}{x^*_\zeta}.$$

Then for $\zeta' \in (0, \zeta)$, we have

$$\overline{U}_{\zeta'}(x^0) \geq U_{\zeta'}(0)\frac{x^*_\zeta - x^0}{x^*_\zeta} + U_{\zeta'}(x^*_\zeta)\frac{x^0}{x^*_\zeta} = U_{\zeta'}(x^*_\zeta)\frac{x^0}{x^*_\zeta} > U_\zeta(x^*_\zeta)\frac{x^0}{x^*_\zeta} = \overline{U}_\zeta(x^0),$$

where the first inequality is by definition of $\overline{U}_{\zeta'}$, and the second is by $x^*_\zeta > x^i$ and strict monotonicity of U on (x^i, ∞). $\square$

A formal version of Proposition 11 takes the following form.

Proposition A11. *Assume that $\underline{x} \geq 0$, so that $U(0) = 0$.*

(1) *$\overline{U}(x^0)/U(x^0)$ is decreasing in x^0, strictly so if $x^0 \leq x^*$.*

(2) *$\overline{U}(x^0)/(x^0 U(1))$ is decreasing in x^0, strictly so if $x^0 \geq x^*$.*

(3) *$\overline{U}(x^0)/\underline{U}(x^0)$ is strictly decreasing in x^0 for all x^0.*

Proof. Since U is strictly S-shaped on $(\underline{x}, 1)$ and $U(0) = 0$,

$$\overline{U}(x^0) = \begin{cases} \frac{x^0}{x^*} U(x^*), & \text{if } x^0 \leq x^*, \\ U(x^0), & \text{if } x^0 > x^*, \end{cases}$$

$$\underline{U}(x^0) = \begin{cases} U(x^0), & \text{if } x^0 < x_*, \\ \frac{1-x^0}{1-x_*} U(x_*) + \frac{x^0 - x_*}{1-x_*} U(1), & \text{if } x^0 \geq x_*, \end{cases}$$

where $x^* > x^i$ and $x_* < x^i$ are uniquely determined by

$$x^* = \arg\max_{x \in [x^i, 1]} \frac{x^0}{x} U(x),$$

$$x_* = \arg\min_{x \in [0, x^i]} \frac{1 - x^0}{1 - x} U(x) + \frac{x^0 - x}{1 - x} U(1).$$

(1) Clearly, $\overline{U}(x^0)/U(x^0) = 1$ for $x^0 \in [x^*, 1)$. Moreover, for $x^0 \in (0, x^*]$,

$$\frac{\overline{U}(x^0)}{U(x^0)} = \frac{U(x^*)}{x^*} \frac{x^0}{U(x^0)}$$

is strictly decreasing in x^0, because U is strictly S-shaped.

(2) Clearly, $\overline{U}(x^0)/(x^0 U(1)) = U(x^*)/(x^* U(1))$ is constant for $x^0 \in (0, x^*]$. Moreover, for $x \in [x^*, 1)$,

$$\frac{\overline{U}(x^0)}{x^0 U(1)} = \frac{1}{U(1)} \frac{U(x^0)}{x^0}$$

is strictly decreasing in x^0, because U is strictly S-shaped.

(3) First, $\overline{U}(x^0)/\underline{U}(x^0) = \overline{U}(x^0)/U(x^0)$ is strictly decreasing in x^0 for $x^0 \in (0, x_*]$, by Part (1). Next, for $x^0 \in [x_*, x^*]$,

$$\frac{\overline{U}(x^0)}{\underline{U}(x^0)} = \frac{\frac{x^0}{x^*}U(x^*)}{\frac{1-x^0}{1-x_*}U(x_*) + \frac{x^0-x_*}{1-x_*}U(1)}$$

is strictly decreasing in x^0 because $x_*U(1) > \underline{U}(x_*) = U(x_*)$ by definition of the convex envelope. Finally, for $x^0 \in [x^*, 1)$,

$$\frac{\overline{U}(x^0)}{\underline{U}(x^0)} = \frac{U(x^0)}{\frac{1-x^0}{1-x_*}U(x_*) + \frac{x^0-x_*}{1-x_*}U(1)} = \frac{U(x^0)}{\frac{x^0}{x^*}U(x^*)} \frac{\frac{x^0}{x^*}U(x^*)}{\frac{1-x^0}{1-x_*}U(x_*) + \frac{x^0-x_*}{1-x_*}U(1)}$$

is strictly decreasing in x^0 because both terms are strictly decreasing in x^0, as shown above. □

Appendix E. Details and Proofs for Section 7

E.1. **Details for Section 7.1.** Corollary 4 generalizes as follows.

Corollary A4. *Assume that W is weakly convex, q is strictly log-concave, G is weakly concave, and the median of Q is weakly higher than the mode. In every optimal districting plan, every two districts P_1 and P_2 are nested, in that P_1 consists of voter types $s_1 \leq s_1'$ and P_2 consists of voter types $s_2 \leq s_2'$ such that either $s_2 \leq s_1 \leq s_1' \leq s_2'$ or $s_1 \leq s_2 \leq s_2' \leq s_1'$.*

The proof again rests on Lemmas A6 and A7. The proof of Lemma A6 is the same. To see heuristically why Lemma A7 holds, suppose that there exist two districts P_1 and P_2 that segregate voters with types $s_1 < s_2$ and consider a perturbation that pools mass dm^2 of voters with type s_1 and mass dm of voters with type s_2. This perturbation decreases $H(r)$ for $r \in [r^*(\delta_{s_1}), r^*(\delta_{s_2}) - dr]$ by dm^2 and increases $H(r)$ for $r \in [r^*(\delta_{s_2}) - dr, r^*(\delta_{s_2})]$ by dm where dr is given by (A11). Thus, this perturbation increases the designer's expected utility by

$$\left[\int_{r^*(\delta_{s_1})}^{r^*(\delta_{s_2})} w(1 - H(r))dG(r)\right] dm^2 - w(1 - H(r^*(\delta_{s_2}))g(r^*(\delta_{s_2}))dmdr$$
$$\geq w(1 - H(r^*(\delta_{s_2}))\left[(G(r^*(\delta_{s_2})) - G(r^*(\delta_{s_1})))dm - g(r^*(\delta_{s_2}))dr\right] dm > 0,$$

where the first inequality is by convexity of W, so $w(1 - H(r)) \geq w(1 - H(r^*(\delta_{s_2})))$ for $r < r^*(\delta_{s_2})$, and the second inequality is by (A12).

E.2. **Proofs for Section 7.2.** We first fully characterize the optimal districting plan. For simplicity, we state the result for the case where F has a strictly positive density, but it is straightforward to extend it to the general case (see, e.g., Corollary A6 for the binary case). This result uses Myerson (1981)'s ironing procedure and is similar to a recent and independent result of Kleiner, Moldovanu, and Strack (2020) (their Proposition 2). Let $\overline{W}$ denote the concave envelope of W.

Lemma A11. *Assume that $F \in \Delta[0,1]$ has a strictly positive density. Let $\{(\underline{s}_i, \overline{s}_i)\}_{i \in I}$ be a collection of disjoint nonempty intervals such that $\overline{W}(m)$ is linear in m on each interval $(1-F(\overline{s}_i), 1-F(\underline{s}_i))$ and $\overline{W}(m) = W(m)$ elsewhere. A districting plan H that pools voters with types $s \in (\underline{s}_i, \overline{s}_i)$, for each $i \in I$, in districts with the same mean type $x_i = \mathbb{E}_F[s | s \in (\underline{s}_i, \overline{s}_i)]$ and segregates the remaining voters is optimal. The designer's expected utility under this plan is*

$$\int W(1-H(x))dx = \int \overline{W}(1-F(x))dx. \tag{A19}$$

Proof. For any $\tilde{H} \in \Delta[0,1]$, define the right-continuous inverse of $\tilde{H}$ as

$$\tilde{H}^{-1}(p) = \begin{cases} \inf\{x \in [0,1] : p < \tilde{H}(x)\}, & \text{if } p \in [0,1), \\ 1, & \text{if } p = 1. \end{cases}$$

By definition, $\tilde{H}^{-1}$ is a non-negative, non-decreasing, and right-continuous function that satisfies $\tilde{H}^{-1}(1) = 1$, so $\tilde{H}^{-1} \in \Delta[0,1]$. Moreover, $\tilde{H}^{-1} \succsim F^{-1}$ iff $F \succsim \tilde{H}$, because

$$F \succsim \tilde{H} \iff \int_0^p F^{-1}(\tilde{p})d\tilde{p} \le \int_0^p \tilde{H}^{-1}(\tilde{p})d\tilde{p} \text{ for all } p \iff \tilde{H}^{-1} \succsim F^{-1}, \tag{A20}$$

where the equivalences are by Theorems 3.A.5 and 3.A.1 in Shaked and Shanthikumar (2007).

For any feasible $\tilde{H}$, we have

$$\begin{aligned} \int W(1-\tilde{H}(x))dx &\le \int \overline{W}(1-\tilde{H}(x))dx = \int \overline{W}(1-p)d\tilde{H}^{-1}(p) \\ &\le \int \overline{W}(1-p)dF^{-1}(p) = \int \overline{W}(1-F(x))dx, \end{aligned} \tag{A21}$$

where the equalities are by variable change, the first inequality is by $\overline{W} \ge W$, and the second inequality is by concavity of $\overline{W}$ and $\tilde{H}^{-1} \succsim F^{-1}$. Thus, to verify the optimality of H, it suffices to show that (A19) holds. Note that H is given by

$$H(x) = \begin{cases} F(s), & \text{if } x \notin \cup_{i\in I}(\underline{s}_i, \overline{s}_i), \\ F(\underline{s}_i), & \text{if } x \in (\underline{s}_i, x_i), \\ F(\overline{s}_i), & \text{if } x \in [x_i, \overline{s}_i). \end{cases}$$

Denote $\overline{X} = [0,1] \setminus \cup_{i\in I}(\underline{s}_i, \overline{s}_i)$, $\underline{m}_i = 1 - F(\overline{s}_i)$, and $\overline{m}_i = 1 - F(\underline{s}_i)$, so

$$\int W(1-H(x))dx = \int_{\overline{X}} W(1-F(x))dx + \sum_{i\in I}(W(\overline{m}_i)(x_i - \underline{s}_i) + W(\underline{m}_i)(\overline{s}_i - x_i)).$$

Notice that

$$\int_{\overline{X}} \overline{W}(1-F(x))dx = \int_{\overline{X}} W(1-F(x))dx,$$

because $W(1-F(x)) = \overline{W}(1-F(x))$ for all $x \in \overline{X}$. Moreover, for each $i \in I$,

$$\begin{aligned} \int_{\underline{s}_i}^{\overline{s}_i} \overline{W}(1-F(x))dx &= \int_{\underline{s}_i}^{\overline{s}_i} \left[W(\underline{m}_i)\frac{\overline{m}_i - (1-F(x))}{\overline{m}_i - \underline{m}_i} + W(\overline{m}_i)\frac{1-F(x) - \underline{m}_i}{\overline{m}_i - \underline{m}_i} \right] dx \\ &= W(\underline{m}_i)(\overline{s}_i - x_i) + W(\overline{m}_i)(x_i - \underline{s}_i), \end{aligned} \tag{A22}$$

where the first equality holds because $\overline{W}(m)$ is linear in m on $[\underline{m}_i, \overline{m}_i]$ with $\overline{W}(m) = W(m)$ for $m \in \{\underline{m}_i, \overline{m}_i\}$, and the second equality holds by integration by parts,

$$\int_{\underline{s}_i}^{\overline{s}_i} (1-F(x))dx = (1-F(x))x|_{\underline{s}_i}^{\overline{s}_i} + \int_{\underline{s}_i}^{\overline{s}_i} x dF(x) = \underline{m}_i\overline{s}_i - \overline{m}_i\underline{s}_i + (\overline{m}_i - \underline{m}_i)x_i.$$

So H satisfies (A19) and is thus optimal. □

We will also use the following result, which is analogous to Lemma A3.

Lemma A12. *Assume that $F \in \Delta[0,1]$ has a strictly positive density on $[0,1]$. If $\overline{W}$ is strictly concave on $[m^*, 1]$ for some $m^* \in [0,1)$, then every optimal H satisfies $H(x) = F(x)$ for all $x \in [0, s^*]$, where $s^* = F^{-1}(1-m^*)$.*

Proof. Note that all inequalities in (A21) must hold with equality for every optimal H; so integrating (A21) by parts twice yields

$$\begin{aligned} 0 &= \int_0^1 \overline{W}(1-p)dF^{-1}(p) - \int \overline{W}(1-p)dH^{-1}(p) \\ &= \int_0^1 \left(\int_0^p H^{-1}(\tilde{p})d\tilde{p} - \int_0^p F^{-1}(\tilde{p})d\tilde{p} \right) d\overline{W}'(1-p), \end{aligned} \tag{A23}$$

where $\overline{W}'(1-p)$ is the right derivative of $\overline{W}$ at $1-p$, which is non-decreasing in p given that $\overline{W}$ is concave. The integrand in (A23) is non-negative by (A20). Thus, for (A8) to hold, the integrand must be zero almost everywhere where $\overline{W}'(1-p)$ is strictly increasing in p. Since $\overline{W}$ is strictly concave on $[m^*, 1]$ and F has a density, this implies that

$$\int_0^p H^{-1}(\tilde{p})d\tilde{p} = \int_0^p F^{-1}(\tilde{p})d\tilde{p} \quad \text{for all } p \in [0, 1-m^*],$$

and thus $H(x) = F(x)$ for all $x \in [0, s^*]$. □

We are now ready to establish a generalized version of Proposition 12, which characterizes when trivial and non-trivial segregate-pool districting plans are optimal.

Proposition A12. *Assume that $F \in \Delta[0,1]$ has a strictly positive density on $[0,1]$.*

(1) *Segregate-pool districting with pool measure $m^* \in [0,1)$ is optimal iff*

$$W(m) \leq W(m^*) + w(m^*)(m - m^*) \text{ for all } m \in [0,1], \text{ with equality at } m = 0,$$
$$\text{and } W \text{ is concave on } [m^*, 1].$$

(2) *Uniform districting (i.e., $m^* = 1$) is optimal iff*

$$W(m) \leq W(0)(1-m) + W(1)m \text{ for all } m \in [0,1].$$

Moreover, every optimal districting plan has the same distribution H^* of mean types if the inequality is strict for all $m \in (0, m^*)$ and W is strictly concave on $[m^*, 1]$.

Proof. It is easy to see that in Case (1) the concave envelope of W is

$$\overline{W}(m) = \begin{cases} W(m^*) + w(m^*)(m - m^*), & \text{if } m \in [0, m^*), \\ W(m), & \text{if } m \in [m^*, 1], \end{cases} \tag{A24}$$

and in Case (2) it is

$$\overline{W}(m) = W(0)(1-m) + W(1)m \text{ for all } m \in [0,1].$$

Thus, by Lemma A11, segregate-pool districting H^* is optimal.

Conversely, suppose that H^* is optimal. Then, by Lemma A11, all inequalities in (A21) must hold with equality for $H = H^*$, so that

$$\int_0^{s^*} W(1 - H^*(x))dx = \int_0^{s^*} W(1 - F(x))dx = \int_0^{s^*} \overline{W}(1 - F(x))dx \tag{A25}$$

and

$$\int_{s^*}^{1} W(1-H^*(x))dx = W(m^*)(x^*-s^*)+W(0)(1-x^*) = \int_{s^*}^{1} \overline{W}(1-F(x))dx. \quad \text{(A26)}$$

Clearly, (A25) implies that $W(1-F(x)) = \overline{W}(1-F(x))$ for all $x \in [0, s^*]$ or equivalently $W(m) = \overline{W}(m)$ for all $m \in [m^*, 1]$, since F has a density. By definition, $\overline{W}$ is concave and thus so is W on $[m^*, 1]$. Moreover, (A26) holds with equality if $\overline{W}(m) = W(0)(1-m) + W(m^*)m$ for all $m \in [0, m^*]$ by (A22). This implies that $W(m) \leq W(0)(1-m) + W(m^*)m$ for all $m \in [0, m^*]$, otherwise the right-hand side in (A26) would be strictly larger than the left-hand side, given that F has a density.

Finally, under the strict conditions, $\overline{W}$ satisfies the conditions of Lemma A12. Thus, every optimal H satisfies $H(x) = F(x)$ for all $x \in [0, s^*]$. Taking into account that every optimal H satisfies all inequalities in (A21) with equality gives

$$\int_{s^*}^{1} W(1-H(x))dx = \int_{s^*}^{1} \overline{W}(1-F(x))dx,$$

which implies that $1-H(x) \in \{1-F(s^*), 0\}$ for all $x \in [s^*, 1]$, because $W(m) < \overline{W}(m)$ for all $m \in (0, m^*)$. Taking into account that $F \succsim H$ and $H(x) = F(x)$ for all $x \in [0, s^*]$ yields $H = H^*$. □

Corollary 5 follows easily from Proposition A12. If W is S-shaped, then it is easy to see that there exists $m^* \in [0, 1]$ that satisfies the conditions of Proposition A12, so segregate-pool districting is optimal. Moreover, the induced distribution H^* of mean types is uniquely optimal if W is strictly S-shaped, because the strict conditions of Proposition A12 are satisfied.

Corollary 6 is a special case of the following result which fully characterizes optimal districting in the case of binary voter types.

Corollary A6. *Assume that* $\operatorname{supp}(F) = \{0, 1\}$ *and* $\Pr_F(s = 1) = x^0$.

(1) *If* $W(x^0) < \overline{W}(x^0)$, *so that there exist* $m_L, m_H \in [0, 1]$ *such that* $m_L < x^0 < m_H$ *and*

$$W(m) \leq W(m_L)\frac{m_H - x^0}{m_H - m_L} + W(m_H)\frac{x^0 - m_L}{m_H - m_L} \text{ for all } m \in [0, 1], \quad \text{(A27)}$$

then, and only then, a districting plan that creates measure $1 - m_H$ *of districts with mean type* $x_L = 0$, *measure* $m_H - m_L$ *of districts with mean type* $x_M =$

$(x^0 - m_L)/(m_H - m_L)$, and measure m_L of districts with mean type $x_H = 1$ is optimal.

(2) *If $W(x^0) = \overline{W}(x^0)$, so that*

$$W(m) \leq W(x^0) + w(x^0)(m - x^0) \text{ for all } m \in [0,1], \tag{A28}$$

then, and only then, segregation (i.e., $m_L = m_H = x^0$) is optimal.

Moreover, the optimal districting plan is unique if the inequality is strict for all $m \notin \{m_L, m_H\}$. In particular, pack and crack (i.e., $m_L = 0$) is optimal if W is S-shaped, and it is uniquely optimal if W is strictly S-shaped.

Proof. Notice that the proof that (A21) holds for any feasible H holds for a general F. The distribution of mean types in Case (1) is

$$H(x) = \begin{cases} 1 - m_H, & \text{if } x \in \left[0, \frac{x^0 - m_L}{m_H - m_L}\right), \\ 1 - m_L, & \text{if } x \in \left[\frac{x^0 - m_L}{m_H - m_L}, 1\right), \\ 1, & \text{if } x = 1, \end{cases}$$

and it is $H = F$ in Case (2). In Case (1), we have

$$\begin{aligned} \int W(1 - H(x))dx &= W(m_H)\frac{x^0 - m_L}{m_H - m_L} + W(m_L)\frac{m_H - x^0}{m_H - m_L} \\ &= \overline{W}(x^0) = \int \overline{W}(1 - F(x))dx, \end{aligned}$$

showing that H is optimal. Similarly, H is optimal in Case (2), because

$$\int W(1 - H(x))dx = W(x^0) = \overline{W}(x^0) = \int \overline{W}(1 - F(x))dx.$$

Conversely, if H is optimal, then the above equalities hold, which implies inequalities (A27) and (A28). Moreover, if the inequalities are strict for all $m \notin \{m_L, m_H\}$, then every $\tilde{H}$ that satisfies the above equalities must satisfy $1 - \tilde{H}(x) \in \{m_L, m_H\}$ for all $x \in [0, 1)$, implying that H is uniquely optimal and that that the optimal districting plan is unique, given that voter types are binary. Finally, the part with S-shaped and strictly S-shaped W holds by the same argument as in Corollary 5. □

E.3. **Proofs for Section 7.3.** We assume that U is linear on its domain, and W is a continuously differentiable and weakly increasing function from $\mathbb{R}$ to $[0,1]$. Further, we assume that $W(m) = 0$ for all $m \leq \underline{m}$ and some $\underline{m} < 1$. Finally, we assume that W is strictly S-shaped on $(\underline{m}, \infty)$. Specifically, there exists an inflection point $m^i \in (\underline{m}, 1)$ such that W is strictly convex on $(\underline{m}, m^i)$ and strictly concave on (m^i, ∞).

We first note that a simple generalization of Corollary 6 gives that there exists a unique optimal plan that takes the form of pack and crack if $\mathrm{supp}(F)$ is binary, which is assumed throughout this appendix. Moreover, we henceforth also assume, except in Part (1c), that $\mathrm{supp}(F) = \{0, 1\}$ with $\Pr_F(s = 1) = x^0$.

A formal version of Proposition 13 takes the following form.

Proposition A13.

(1) (a) *Let m^*_λ be the pool measure of the optimal pack and crack plan under $W_\lambda(m) = W(m - \lambda)$ where $\lambda < 1 - \underline{m}$. Then m^*_λ is increasing in λ, strictly so if $m^*_\lambda \in (x^0, 1)$.*

(b) *Let m^*_ζ be the pool measure of the optimal pack and crack plan under $W_\zeta(m) = W(m^i + (m - m^i)/\zeta)$ where $\zeta > 0$. Then m^*_ζ is increasing in ζ, strictly so if $m^*_\zeta \in (x^0, 1)$.*

(c) *Let distributions F and $\tilde{F}$ have the same mean x^0 and binary supports $\{s_L, s_H\}$ and $\{\tilde{s}_L, \tilde{s}_H\}$ with $s_L < \tilde{s}_L < \tilde{s}_H < s_H$. Moreover, let x^* and $\tilde{x}^*$ be the pool mean of the optimal pack and crack plan under F and $\tilde{F}$. Then $x^* > \tilde{x}^*$, unless $x^* = \tilde{x}^* = x^0$.*

(2) *Let $\overline{W}_\zeta(x^0)$ be the concave envelope of W_ζ on $[0,1]$ at x^0 where $W_\zeta(m) = W(m^i + (m - m^i)/\zeta)$ with $\zeta \in (0, m^i/(m^i - \underline{m})]$, so that $W_\zeta(0) = 0$. Then $\overline{W}_\zeta(x^0)$ is strictly decreasing in ζ.*

(3) *Assume that $\underline{m} \geq 0$, so that $W(0) = 0$.*

(a) *$\overline{W}(x^0)/W(x^0)$ is decreasing in x^0, strictly so if $x^0 \leq x^*$.*

(b) *$\overline{W}(x^0)/(x^0 W(1))$ is decreasing in x^0, strictly so if $x^0 \geq x^*$.*

(c) *$\overline{W}(x^0)/\underline{W}(x^0)$ is strictly decreasing in x^0 for all x^0.*

The proofs of all parts, except Part (1c), are the same as the proofs of Propositions A8–A11. In Part (1c), the optimal plan pools measure $m^\dagger$ of voters with the highest types and segregates the remaining voters where $m^\dagger \in [0,1]$ is such that $W(m) < \overline{W}(m)$ for all $m \in (0, m^\dagger)$ and $W(m) = \overline{W}(m)$ for all $m \in [m^\dagger, 1]$. Taking into account that $\Pr_F(s = s_H) = (x^0 - s_L)/(s_H - s_L)$ gives

$$x^* = \begin{cases} \frac{x^0-(1-m^\dagger)s_L}{m^\dagger}, & \text{if } \frac{x^0-s_L}{s_H-s_L} \in [0, m^\dagger), \\ s_H, & \text{if } \frac{x^0-s_L}{s_H-s_L} \in [m^\dagger, 1], \end{cases}$$

and similarly for $\tilde{x}^*$. Thus, $x^* > \tilde{x}^*$, unless $m^\dagger = 1$, so that $x^* = \tilde{x}^* = x^0$.

Part (2) implies that the designer's expected utility $U(0)W_\zeta(1) + (U(1) - U(0))W_\zeta(x^0)$ is strictly decreasing in ζ, because $W_\zeta(1)$ is strictly decreasing in ζ. Similarly, if $U(0) = 0$, Part 3 implies that the ratio of the designer's expected utility under an optimal districting plan to his expected utility under any of the three benchmarks (i.e., (a) segregation, (b) uniform districting, and (c) pessimal districting) is decreasing in the share of favorable voters.

E.4. **Proofs for Section 7.4.**

Proof of Proposition 14. To simplify notation, without loss of generality, we normalize $U(0) = W(0) = 0$ and $U(1) = W(1) = 1$.

Consider first the case where uniform districting (i.e., H^* assigns probability one to $x^* = \mathbb{E}_F[s]$) is optimal under both $(\tilde{U}(x), \tilde{W}(x)) = (U(x), x)$ and $(\tilde{U}(x), \tilde{W}(x)) = (x, W(x))$. Then by Part (2) of Proposition A6,

$$U(x) \leq U(x^*) + u(x^*)(x - x^*) \text{ for all } x \in [0,1], \tag{A29}$$

and by Part (2) of Proposition A12,

$$W(m) \leq m \text{ for all } m \in [0,1]. \tag{A30}$$

For any feasible H, we have

$$\begin{aligned}\int W(1-H(x))dU(x) &\le \int (1-H(x))dU(x)\\ &= \int U(x)dH(x)\\ &\le U(x^*)\\ &= \int W(1-H^*(x))dU(x),\end{aligned}$$

The inequalities are by (A30) and (A29). The first equality is by integration by parts. The last equality is by definition of H^*. So H^* is optimal.

Consider now the case where segregate-pool districting with pool mean $x^* \in [\mathbb{E}[s], 1]$ is optimal under both $(U(x), x)$ and $(x, W(x))$. Then by Part (1) of Proposition A6, $U(x) \le \widehat{U}(x)$ for all $x \in [0,1]$ where $\widehat{U}$ is a convex function given by (A9). Moreover, by Part (1) of Proposition A12, $W(m) \le \overline{W}(m)$ for all $m \in [0,1]$ where $\overline{W}$ is a concave function given by (A24).

For any feasible H, we have

$$\begin{aligned}\int W(1-H(x))dU(x) &\le \int \overline{W}(1-H(x))dU(x)\\ &\le \int \overline{W}(1-H(x))d\widehat{U}(x)\\ &= \int \overline{W}(1-p)d\widehat{U}(H^{-1}(p))\\ &\le \int \overline{W}(1-p)d\widehat{U}(F^{-1}(p))\\ &= \int \overline{W}(1-F(x))d\widehat{U}(x)\\ &= \int_0^{s^*} W(1-F(x))dU(x) + W(m^*)(U(x^*)-U(s^*))\\ &= \int W(1-H^*(x))dU(x).\end{aligned}$$

The first inequality holds because $W \le \overline{W}$. The second inequality holds because $\overline{W}(1-p)$ is decreasing in p and $U \le \widehat{U}$ (i.e., U first-order stochastically dominates $\widehat{U}$). The first and second equalities hold by variable change. The third inequality holds because $\overline{W}(1-p)$ is decreasing and concave in p and $\widehat{U} \circ H^{-1}$ is higher than $\widehat{U} \circ F^{-1}$ in the increasing convex order by Theorem 4.A.8(a) in Shaked and Shanthikumar (2007), given that $\widehat{U}$ is an increasing convex function and $H^{-1} \succsim F^{-1}$ by (A20). The third equality holds because $\widehat{U}$ is given by (A9) with $u(x^*)(x^* - s^*) = U(x^*) - U(s^*)$ and $\overline{W}$ is given by (A24) with $\overline{W}(0) = W(0)$. The last equality holds by definition of H^*. So H^* is optimal. □

Proof of Corollary 7. In the best segregate-pool districting plan under $(U(x), W(x))$, the cutoff type is

$$s^* \in \arg\max_{s^\dagger \in [0,1]} \int_0^{s^\dagger} W(1 - F(x))dU(x) + W(m^\dagger)(U(x^\dagger) - U(s^\dagger)),$$

where $x^\dagger = \mathbb{E}_F[s|s \geq s^\dagger]$ and $m^\dagger = 1 - F(s^\dagger)$. The derivative of the objective function with respect to $s^\dagger$ is

$$\begin{aligned} &W(m^\dagger)u(x^\dagger)\frac{f(s^\dagger)}{m^\dagger}(x^\dagger - s^\dagger) + w(m^\dagger)(-f(s^\dagger))(U(x^\dagger) - U(s^\dagger)) \\ &= f(x^\dagger)w(m^\dagger)u(x^\dagger)(x^\dagger - s^\dagger)\left[\frac{W(m^\dagger)}{w(m^\dagger)m^\dagger} - \frac{U(x^\dagger) - U(s^\dagger)}{u(x^\dagger)(x^\dagger - s^\dagger)}\right]. \end{aligned}$$

Since U is strictly S-shaped,

$$\frac{U(x^\dagger) - U(s^\dagger)}{u(x^\dagger)(x^\dagger - s^\dagger)} \gtreqless 1 \iff x^\dagger \gtreqless x_U^*.$$

Similarly, since W is strictly S-shaped,

$$\frac{W(m^\dagger)}{w(m^\dagger)m^\dagger} \lesseqgtr 1 \iff x^\dagger \gtreqless x_W^*.$$

Thus, the objective function is strictly increasing in $x^\dagger$ for $x^\dagger \leq \min\{x_U^*, x_W^*\}$ and is strictly decreasing in $x^\dagger$ for $x^\dagger \geq \max\{x_U^*, x_W^*\}$, showing that the optimal x^* lies between x_U^* and x_W^*. □

Proof of Example 3. Since U and W are S-shaped, pack and crack is optimal under both $(\tilde{U}(x), \tilde{W}(x)) = (U(x), x)$ and $(\tilde{U}(x), \tilde{W}(x)) = (x, W(x))$. The corresponding optimal pool means are $x_U^* = 1/2$ and $x_W^* = 3/4$. Thus, in the best pack and crack plan, the pool mean is $x^* \in [1/2, 3/4]$. In fact, we can narrow down the range of x^* further to the interval $[1/2, 5/8]$, because the aggregate shock never exceeds $5/8$. Thus, the pool mean of the best pack and crack plan solves

$$\max_{x \in [\frac{1}{2}, \frac{5}{8}]} (\tfrac{1}{2} + \tfrac{1}{2}\tfrac{3/8}{x})(\tfrac{3}{8} + x) = \tfrac{3}{8} + \tfrac{1}{2} \max_{x \in [\frac{1}{2}, \frac{5}{8}]} (x + (\tfrac{3}{8})^2 \tfrac{1}{x})$$

It is easy to see that the objective function is convex and is thus maximized at either $x^* = 1/2$ or $x^* = 5/8$. Since the designer's expected utility is higher under $x^* = 5/8$,

$$(\tfrac{1}{2} + \tfrac{1}{2}\tfrac{3/8}{5/8})(\tfrac{3}{8} + \tfrac{5}{8}) = \tfrac{4}{5} > (\tfrac{1}{2} + \tfrac{1}{2}\tfrac{3/8}{1/2})(\tfrac{3}{8} + \tfrac{1}{2}) = (\tfrac{7}{8})^2,$$

the optimal pool mean is $x^* = 5/8$. Thus, in the best pack and crack plan, the designer always wins measure $m^* = 3/5$ of districts with mean voter type $x^* = 5/8$, while always losing the remaining districts.

There exists a strictly better plan that creates measure $m_2 = 1/2$ of districts with mean type $x_2 = 5/8$, measure $m_1 = 1/8$ of districts with mean type $x_1 = 1/2$, and measure $m_0 = 3/8$ of districts with mean type $x_0 = 0$. Under this plan, the designer wins measure $m_2 + m_1 = 5/8$ of districts with probability $U(1/2) = 7/8$ and measure $m_2 = 1/2$ of districts with probability $U(5/8) - U(1/2) = 1/8$. Thus, the designer's expected utility is strictly higher under this plan than the best pack and crack plan,

$$(\tfrac{1}{2} + \tfrac{1}{2}\tfrac{5}{8})\tfrac{7}{8} + (\tfrac{1}{2} + \tfrac{1}{2}\tfrac{1}{2})\tfrac{1}{8} = \tfrac{103}{128} > \tfrac{4}{5}.$$

Finally, we outline the proof that this plan is optimal. Since the designer always loses districts with mean type below $1/2$ and always wins districts with mean type above $5/8$, an optimal districting plan never creates districts with mean types in $(0, 1/2) \cup (5/8, 1]$. Let $x^\dagger \in [1/2, 5/8]$ be the maximum district mean type of an optimal districting plan, so that $H(x^\dagger) = 1$. Since the designer gets zero utility from winning less than measure $1/2$ of districts, the mean type should be at least $x^\dagger$ in at least measure $1/2$ of districts, so that $H(x^\dagger_-) \leq 1/2$. Now for a fixed such $x^\dagger \in [1/2, 5/8]$, a constrained optimal plan solves a linear program. Since the ratio of the probability of winning a district to the district mean type x (given by $(3/8 + x)/x$) is decreasing in x on $[1/2, x^\dagger]$, the constrained optimal plan creates as many districts as possible with mean type $1/2$. Thus, it creates measure $1/2$ of districts with mean type $x^\dagger$, measure $3/4 - x^\dagger$ of districts with mean type $1/2$, and measure $x^\dagger - 1/4$ of districts with mean type 0. The designer's expected utility under this plan is

$$(\tfrac{1}{2} + \tfrac{1}{2}(\tfrac{5}{4} - x^\dagger))\tfrac{7}{8} + (\tfrac{1}{2} + \tfrac{1}{2}\tfrac{1}{2})(x^\dagger - \tfrac{1}{2}) = \tfrac{39}{64} + \tfrac{5}{16}x^\dagger.$$

Since this value is increasing in $x^\dagger$, the unconstrained optimal plan has $x^\dagger = 5/8$. $\square$