Competitive Fair Redistricting*

Felix J. Bierbrauer† Mattias Polborn‡

August 18, 2022

Abstract

We consider the process of political redistricting after the decennial census. We call rules for redistricting *fair* if each party cannot be blocked from implementing a district map that guarantees that it wins a majority in the legislature whenever it wins the popular vote. We show that, through a system that involves both political parties in the redistricting process, fairness can be achieved.

Keywords: Gerrymandering, legislative elections, redistricting.
JEL classification: D72, C72.

*We benefitted from conversations with Peter Cramton, Tigran Polborn, Axel Ockenfels, and Ashutosh Thakur. Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC2126/1 – 390838866 is gratefully acknowledged.
†University of Cologne, E-mail: bierbrauer@wiso.uni-koeln.de
‡Vanderbilt University and University of Cologne, E-mail: mattias.polborn@vanderbilt.edu

1 Introduction

After each decennial census, legislative districts in the United States have to be redrawn to ensure that each legislator represents the same number of residents. The task of redistricting falls usually to the current state legislature, a body composed of individuals who have a high degree of self-interest in the outcome. In fact, the party in control of the redistricting process can usually manipulate the district map for the following decade through "gerrymandering" in such a way that it remains in control of the legislature.

To see how this works, consider the following example. The polity consists of a large number of *precincts*, where precincts are indivisible geographic units (i.e., a legislative district consists of several precincts, and each precinct is in exactly one district). There are two types of precincts, some are leaning to the Republican party, some are leaning to the Democrats. The margins of victory at the precinct-level fluctuate e.g. because of political scandals or the ups and downs in the popularity of political leaders. Specifically, suppose that one-half of precincts are *Republican leaning* and return a Republican vote share that is between 60 and 70 percent. The other half is *Democratic-leaning* and returns a Republican vote share between 30 and 40 percent. Formally, the Republican vote share in a Republican-leaning precinct is $0.6+0.1\omega$, and $0.3+0.1\omega$ in a Democratic-leaning precinct, where ω is a random variable taking values in $[0,1]$.

Suppose that Republicans are in control of redistricting. Observe that a district is guaranteed to be won by the Republican candidate if it consists of at least 2/3 Republican-leaning precincts because then, even in the worst case, the Republican vote share is $(2/3) \times 60\% + (1/3) \times 30\% = 50\%$. Clearly, Republicans can endow 75 percent of districts with 2/3 Republican-leaning and 1/3 Democratic-leaning precincts, packing the remaining Democratic-leaning precincts into the remaining 25 percent of districts.

Even if the party in control of redistricting faces additional constraints (for example, can only draw districts that are geographically contiguous), it remains true that the ability to gerrymander usually allows a party to insulate itself from most electoral shifts. For example, Republicans took over the Pennsylvania state legislature in the 2010 Republican wave election and used the opportunity to create a very favorable district map for themselves. For example, even though Democratic candidates received 55 percent of the popular vote in the 2018 elections across all districts, versus 44.4% for Republican candidates, Republicans still controlled 110 out of 203 seats in the Pennsylvania House of Representatives. Many other examples exist, including ones in which the partisan advantage was on the Democrats' side.

Because of these problems, there is substantial backlash against existing gerrymanders and also the institutions that allow for it to happen. District assignments engineered by legislatures can be challenged in courts, and some state supreme courts have granted injunctive relief against maps considered to be so unfair that they violate democratic principles in the respective state's constitution.[1] Any judicial solution of the redistricting problem faces the problem of drawing a necessarily somewhat arbitrary line between "still legal" and "sufficiently outrageous to be illegal." In fact, a very good predictor of what maps courts will approve of is the party affiliation of judges.

In this paper, we therefore explore an alternative approach. We start from the premise that, in a democracy, it is desirable that a party that wins the popular vote wins a majority of seats in the legislature. Can we find "fair" rules for redistricting in which both parties participate and that deliver this result? We consider a system as fair if either party has a strategy that ensures a district map with the property that it wins a majority in the legislature whenever it wins the popular vote. We construct such a fair redistricting system in which both parties participate; an example illustrating this system is given in Section 2.

In this system, a party's ability to ensure a fair outcome for itself holds irrespective of the other party's objective or behavior, and is thus very robust. To be clear, on the assumption that the parties' sole objective is to maximize the probability of winning a majority of seats, the strategies that we characterize are actually equilibrium strategies. A party may however deviate from such a strategy when there are other considerations that are traded off against the probability of winning a majority of seats (e.g. maximizing the expected vote share in the legislature, ensuring a representation of different party factions in the legislature, ensuring a representation of ethnic minorities, protecting various incumbents). Even if one party deviates from its equilibrium strategy, the other party can still get the desired outcome if it sticks to its equilibrium strategy. What defines the fairness of the redistricting system is that parties *can* choose to ensure that, whenever they win the popular vote, they get a majority in the legislature, not that they actually do so.

The basic idea is to design an institution so that the parties can keep each other in

[1] However, in the 2004 *Vieth v. Jubilirer* decision, the US Supreme Court has refused to rule against partisan gerrymanders, arguing that "partisan gerrymandering claims were nonjusticiable because there was no discernible and manageable standard for adjudicating political gerrymandering claims."

check. This is similar in spirit to the classical problem of how to fairly divide a cake between two children – we let one child cut the cake in two pieces and the other one choose which one she wants to have. With this procedure, every child has a strategy that ensures getting at least fifty percent of the cake. Consequently, there is no need for general rules and constraints under which only one child chooses both their own and the other child's piece.

There is a large literature on gerrymandering, both empirical and theoretical. However, most of the existing theoretical literature is on "optimal" gerrymandering from the point of view of the party in control of the gerrymandering process; that is, how to cheat democracy most effectively if given the opportunity to do so. Only very few papers deal with the question of how one could implement a better redistricting system. The earliest such paper is William Vickrey's (1961) paper arguing that "the process [of redistricting] should be completely mechanical so that, once set up, there is no room at all for human choice."[2] Similarly, Ely (2019) proposes a mechanism designed to prevent weirdly-shaped districts. Like our paper, his mechanism relies on the participation of both parties in the redistricting process, and he also appeals to the cake-division problem. There are also important differences: Ely takes convexity as the key desideratum. Our analysis, by contrast, focuses on the alignment of election outcomes with the popular vote, and it abstracts from spatial considerations.

Our redistricting institution can be interpreted as a dynamic Colonel Blotto game (for applications of static divide-the-dollar or Colonel Blotto games, see, for instance, Myerson (1993), Lizzeri and Persico (2001, 2005), Laslier and Picard (2002), Konrad (2009) and Kovenock and Roberson (2020)). To the best of our knowledge, using a dynamic version of this class of games is novel in the literature on mechanism design and implementation theory.[3]

The proof of our main result uses results from the game theoretic analysis of zero-sum games. More specifically, we define a fictitious zero-sum game in which one of the parties gets a payoff of 1 when it has "enough" supporters in half of the districts, with the implication that it wins a majority of seats whenever it wins the popular vote.

[2]He proposes an algorithm that produces geographically-compact districts, but does not study whether elections governed by the generated map have any desirable properties.

[3]Groseclose and Snyder (1996) study coalition formation within a legislature on the assumption that there are two competing vote-buyers. While they also look at a sequential mechanism, their focus is positive rather than normative in that they seek an explanation for the frequent occurrence of supermajorities – as opposed to minimal winning coalitions.

Otherwise the payoff is zero. We then show that the equilibrium payoff for this party is one. By the min-max-theorem due to von Neumann (1928),[4] this implies that the party has a successful strategy – in the sense of winning a majority of districts, conditional on winning the popular vote – *for every strategy of the opposing party.*[5]

More peripherally related to our paper is the theoretical literature on gerrymandering which takes the existing redistricting institution as given and analyzes how the party in control of redistricting optimally exerts its power. The initial paper analyzing how an optimal gerrymander involves "packing" (i.e., concentrating likely opponents in few districts) and "cracking" (distribute one's most likely supporters evenly over the remaining majority of districts) is Owen and Grofman (1988). For an excellent review of this literature in a very general framework, see Kolotilin and Wolitzky (2020).[6]

2 A simple example

To see how our system works, consider the example polity from the introduction with one-half Republican- and Democratic-leaning precincts each. There are $2N$ equal-sized legislative districts to be defined, plus one at-large district that also sends one representative and ensures an odd number of representatives in the legislature.

Consider the following redistricting system. Democrats start and assign each precinct to a district (such that each district consists of the same number of precincts). After the Democrats are done, it's the Republicans' turn to assign each precinct to a district. Observe that each precinct is now in two districts, and votes on both local races.[7]

While a full characterization of the subgame-perfect equilibrium of this game is cumbersome, it is straightforward to show that each party has a strategy that can

[4] See Osborne and Rubinstein (1994) for a textbook treatment.

[5] Our results also mirror a well-known Theorem by Zermelo (1913) on the game of chess. According to Zermelo's theorem, either *White* has a strategy that guarantees a victory, or *Black* has a strategy that guarantees a victory, or both have a strategy that guarantees a draw. While Zermelo, of course, cannot characterize these strategies for chess, we do not just show that there exist, for both parties, strategies that guarantee winning the election (conditional on winning the popular vote), but we can also describe them.

[6] Other papers in this line of work include Friedman and Holden (2008), who study optimal partisan gerrymandering with noisy signals about voters' party preferences, and Gul and Pesendorfer (2010), who analyze partisan gerrymandering when each party controls some territory (as in U.S. House redistricting).

[7] In principle, we allow for the parties to assign a particular precinct to the same district; if that happens, the votes from voters in that precinct would simply count twice in that district election, relative to voters from precincts that are assigned to two different districts.

guarantee itself a majority in the legislature whenever they win the popular vote (i.e., the Democrats if $\omega < 0.5$, and the Republicans when $\omega > 0.5$).

Consider first the Republicans who move second. If the Democrats allocated shares of Democratic-leaning precincts and of Republican-leaning precincts to some district k, the Republicans can just flip this. For example, if Democrats assigned 60 percent Democratic-leaning and 40 percent Republican-leaning precincts to district k, Republicans can produce a perfectly balanced district by assigning, in their move, 60 percent Republican-leaning and 40 percent Democratic-leaning precincts.

It is clearly feasible for Republicans to play this balancing strategy for each district, and this would result in each district going to the winner of the popular vote. Observe, though, that the balancing strategy just described is not necessarily optimal for Republicans. This depends on how Democrats distributed the precincts, and on the Republican party's objective. For these reasons, a full characterization of best responses or of the subgame-perfect equilibrium would be more involved.

Consider now the Democrats, who are the first movers. Suppose that Democrats assign only Democratic-leaning precincts to the first N districts, and only Republican-leaning precincts to districts $N + 1$ to $2N$. Clearly, this is feasible as it uses up all Democratic and Republican-leaning precincts. Furthermore, no matter what the Republicans do in their move, the first N districts will have at least a 50 percent share of Democratic-leaning precincts, so will be won by a Democrat whenever $\omega < 0.5$. Since Democrats also win the at-large district whenever $\omega < 0.5$, they are guaranteed a majority in the legislature whenever $\omega < 0.5$.

In the following analysis, we will show how to generalize this example to the case that the number of Democratic- and Republican-leaning precincts is not the same, and that the average partisan lean of these two types of districts is not the same. For example, throughout the United States, Democrats are often very strongly concentrated in urban areas, and Chen and Rodden (2013) suggest that this geographic fact alone provides a significant advantage for Republicans in a traditional redistricting process. We will show that we can maintain a fair system, although we do need, in general, have to make the redistricting system somewhat more complicated by having the parties distribute precincts over several rounds.

To understand why matters become more complicated, consider an example in which 2/3 of precincts are Democratic-leaning, with Republican vote share $0.3 + 0.2\omega$, and 1/3 of precincts are Republican-leaning, with Republican vote share $0.6 + 0.2\omega$. Note that it is still the case that the Democrats win the popular vote whenever $\omega < 0.5$ and that the Republicans win the popular vote when $\omega > 0.5$. As the share of Democratic-leaning precincts is greater than one-half, it is not possible for Democrats to block all of them together in one-half of the districts. That is, the type of move that guaranteed Democrats a victory whenever they won the popular vote in the previous example is no longer feasible. Blocking them in 2/3 of districts is feasible, but this strategy does not work in the sense of ensuring a majority whenever $\omega > 0.5$.[8] Thus, it becomes more difficult for the Democratic party to make sure that it wins a majority of districts whenever it wins the popular vote. Things are easier for the Republican party: By assigning a percentage share of 2/3 of Republican leaning districts to half of the districts, it can ensure that these districts are won whenever $\omega > 0.5$. This disadvantage for the Democrats can, however, be overcome when voters are assigned over more than just two rounds.

3 The Model

There are $2N$ local districts, indexed by $k \in \{1, 2, \ldots, 2N\}$, and one at-large district. The two parties are labeled R and D. There are two types of "voters," $t \in \{t_1, t_2\}$, that we interpret either as individuals, or as the smallest unit that can be assigned to a district, such as a census block or precinct.

The set of aggregate states of the world is denoted by $\Omega \subset \mathbb{R}$, with generic element ω, taken to be the realization of a real-valued random variable. Let $v(t, \omega)$ denote the probability that a type t person votes for party R in state ω. We adopt a law of large numbers convention and also interpret $v(t, \omega)$ as the share of type t voters voting for party R in state ω. The function v is taken to be strictly increasing in both arguments; i.e., in any given state ω, type 2 is more likely to vote R than type 1, and higher ω

[8] To see this, suppose that Democrats create 2/3 of districts that are composed only of Democratic-leaning precincts, and 1/3 of districts that are exclusively Republican-leaning. Then, Republicans can add only Democratic-leaning precincts to the latter, and block their Republican-leaning precincts in another third of districts, while the remaining third is composed only of Democratic-leaning precincts. Thus, in the two-thirds of districts that consist of an equal share of Democratic- and Republican-leaning precincts, the Republican vote share is

$$\frac{1}{2}[0.3 + 0.2\omega] + \frac{1}{2}[0.6 + 0.2\omega] = 0.45 + 0.2\omega,$$

which is greater than 0.5 whenever $\omega > 1/4$. Thus, if $\omega \in [0.25, 0.5)$, Republicans win the majority while losing the popular vote.

increases the share of R voters among both types. The mass of type t_j voters is given by

$$b_j = 2N\,\beta_j\ , \quad \text{where} \quad \beta_1 + \beta_2 = 1 \quad \text{and} \quad \beta_1 \leq \frac{1}{2}\ .$$

The popular vote. Let $\hat{\omega} \in \Omega$ denote the state that yields a popular vote tie, i.e.,[9]

$$\beta_1\, v(t_1, \hat{\omega}) + \beta_2\, v(t_2, \hat{\omega}) = \frac{1}{2}\ . \tag{1}$$

Party R wins the popular vote if $\omega > \hat{\omega}$, while party D wins the popular vote if $\omega < \hat{\omega}$. Conditional on state $\hat{\omega}$, type 1 voters are more likely to vote for party D and type 2 voters are more likely to vote for party R,

$$v(t_1, \hat{\omega}) \quad < \quad \frac{1}{2} \quad < \quad v(t_2, \hat{\omega})\ .$$

We also assume that type 1 voters are weakly more partisan than type 2 voters in the sense that, in the critical state $\hat{\omega}$, type 1 votes D with a probability that is at least as high as the probability that type 2 votes R,

$$1 - v(t_1, \hat{\omega}) \quad \geq \quad v(t_2, \hat{\omega})\ .$$

Interpretation. One special case of this setup has $v(t_1, \hat{\omega}) = 0$ and $v(t_2, \hat{\omega}) = 1$ and $\beta_1 = \beta_2$. In this case a "voter" is really an individual whose vote, conditional on the state, the parties can perfectly predict. In state $\hat{\omega}$, type 1 (2) votes for party D (R). For states $\omega > \hat{\omega}$, some type 1 voters – formally, a fraction that is increasing in ω – vote R. Likewise, for $\omega < \hat{\omega}$, some type 2 voters vote D.

By contrast, when $v(t_1, \hat{\omega}) \in \left(0, \frac{1}{2}\right)$ or $v(t_2, \hat{\omega}) \in \left(\frac{1}{2}, 1\right)$ a "voter" can be interpreted as a precinct, a street, or a census block that needs to be treated as an indivisible unit for the purposes of redistricting. Any such unit contains a mix of individuals who vote for each party; $v(t_1, \hat{\omega}) \in \left(0, \frac{1}{2}\right)$ can then be interpreted as the R-vote share in a unit that predominantly votes for party D. Likewise, $v(t_2, \hat{\omega})$ is the R-vote share in a more R-leaning unit.

When $1 - v(t_1, \hat{\omega}) \quad > \quad v(t_2, \hat{\omega})$, i.e., type t_1 is strictly more partisan than type 2, then $\beta_2 > \frac{1}{2}$. Hence, while there are equal numbers of D and R voters at the aggregate level in state $\hat{\omega}$, D voters are more concentrated: Fewer units mostly vote for D, $\beta_1 < \beta_2$, but in those units, D's vote share is higher than R's vote share in those R-leaning blocks. As ω increases above $\hat{\omega}$, R's vote share increases in both types of blocks, and vice versa.

[9] To assume the existence of $\hat{\omega}$ is without loss of generality because it may have zero probability.

District outcomes. As we describe in more detail below, voters are allocated to districts over several rounds. In this process, each party assigns every voter to one of the districts. Thus, any one voter is assigned twice, once by D and once by R. If a voter is assigned to district k by party D and to some other district $k' \neq k$ by party R, he simply casts one vote in each district election. If $k' = k$ (i.e., both parties assign the voter to the same district), then his vote is counted twice in that district.

Ultimately, every district k contains some mix of type t_1 and type t_2 voters. More formally, a voter assignment by party $P \in \{D, R\}$ is a collection $\sigma_P = (\sigma_{Pk})_{k=1}^{2N}$, where

$$\sigma_{Pk} = (\sigma_{Pk}^1, \sigma_{Pk}^2) \quad \text{with} \quad \sigma_{Pk}^1 + \sigma_{Pk}^2 = 1 \ ,$$

is the assignment of voters to district k by party P. Party R wins district k in state ω if

$$(\sigma_{Dk}^1 + \sigma_{Rk}^1)\ v(t_1, \omega) + (\sigma_{Dk}^2 + \sigma_{Rk}^2)\ v(t_2, \omega) \quad > \quad \frac{1}{2} \ , \tag{2}$$

and vice versa.

The sequence of moves. In each of L rounds, each Party P assigns a mass of $\frac{1}{L}$ voters to any one district k. Formally, in each round l, any party P specifies $\sigma_{Pl} = (\sigma_{kPl}^1, \sigma_{kPl}^2)_{k=1}^{2N}$ so that

$$\sigma_{kPl}^1 + \sigma_{kPl}^2 = \frac{1}{L} \ .$$

The percentage shares of type t_1 and type t_2 voters are then, respectively, given by

$$\beta_{kPl}^1 \ := \ L\, \sigma_{kPl}^1 \quad \text{and} \quad \beta_{kPl}^2 \ := \ L\, \sigma_{kPl}^2 \ .$$

For concreteness, we assume that, for l odd, R moves first and D second. For l even, D moves first and R second. Thus, the second-mover advantage, if any, alternates between D (in odd rounds)and R (in even rounds).

Denote the total mass of type t_1 partisans assigned by party P to district k over the L rounds by $\sigma_{kP}^1 := \sum_{l=1}^{L} \sigma_{kPl}^1$. Analogously, let $\sigma_{kP}^2 := \sum_{l=1}^{L} \sigma_{kPl}^2$. To be consistent with the overall distribution of voters, $(\sigma_{kP})_{k=1}^{2N}$ must satisfy

$$\frac{1}{2N} \sum_{k=1}^{2N} \sigma_{kP}^1 = \beta_1 \quad \text{and} \quad \frac{1}{2N} \sum_{k=1}^{2N} \sigma_{kP}^2 = \beta_2 \ .$$

Winning a majority of seats. Recall that there are $2N$ districts and an at-large-district. Thus, the party that wins at least $N+1$ seats wins a majority in the legislature. Given a pair of voter assignments (σ_D, σ_R), we denote the probability that party R wins a majority of seats, conditional on it winning the popular vote, by $\Pi_R(\sigma_D, \sigma_R \mid \omega > \hat{\omega})$. We define $\Pi_D(\sigma_D, \sigma_R \mid \omega < \hat{\omega})$ analogously.

4 The main result

Theorem 1 below shows that, when the number of rounds L is sufficiently large, each party has a strategy that guarantees winning a majority of seats whenever it wins the popular vote. By "guarantee," we mean that, whatever the opponent does over the various rounds, a party can make sure that it wins a legislative majority if it wins the popular vote.

The Theorem encapsulates our definition of a fair system. We do not claim that both parties will necessarily play in a way that maximizes their probability of winning a majority, as they may also have other objectives in redistricting, for example incumbent protection or representation of various groups. But these are issues that each party has to deal with internally, we cannot expect that a redistricting system solves them for the parties.

Theorem 1 *Let $N \geq 3$. For every $\varepsilon > 0$, there is $\hat{L}$, so that, for $L \geq \hat{L}$: There is a strategy σ_R so that*

$$\Pi_R\left(\sigma_D, \sigma_R \mid \omega > \hat{\omega}\right) = 1 \ , \quad \textit{for every} \quad \sigma_D \ ,$$

and there is a strategy σ_D so that

$$\Pi_D\left(\sigma_D, \sigma_R \mid \omega < \hat{\omega}\right) = 1 \ , \quad \textit{for every} \quad \sigma_R \ .$$

Theorem 1 follows from Propositions 1 and 2 below.

Proposition 1 $\exists\ \sigma_D$ *so that* $\Pi_D(\sigma_D, \sigma_R \mid \omega < \hat{\omega}) = 1$, $\forall\ \sigma_R$ *and* $\forall L \geq 1$.

Proposition 1 generalizes an observation from the analysis of the example in Section 2 (a formal proof can be found in the Appendix): Suppose that D assigns, over the course of the whole procedure, a mass of $2\beta_1$ type 1 voters to half of the districts, say, to any district k with $k \geq N+1$. Then, whatever, the strategy of party R, the percentage share of type 1 voters in those district is bounded from below by β_1, which is the share necessary to win a district whenever $\omega < \hat{\omega}$. Furthermore, as Democrats also win the at-large district whenever $\omega < \hat{\omega}$, this guarantees a majority of seats for party D. Note that this strategy can be implemented for any $L \geq 1$.

Observe that R has a variety of moves to react to D's strategy that leave its probability of winning a majority constant, but have very different implications for the seat distribution. On the one extreme, R can basically choose to focus their supporters on the same N districts and thereby turn every district into a replica of the electorate at large; in this case, all districts are won by the same party, for most states of the world. On the other extreme, the second mover can double down on the first mover's choice and spread all Democratic-leaning precincts in the same N districts that D built up; in this case, each party is essentially guaranteed N seats, with the legislative majority being decided by the at-large district. Of course, combinations of these strategies – for example, building $N/2$ districts that are safe for D and R, respectively, and N that are replicas – are also feasible.

When $\beta_2 = \beta_1 = \frac{1}{2}$ the strategy of blocking ones own supporters into one half of the districts, respectively, is available to both parties, and hence both can use it to ensure wining a majority of seats in those states that favor them, respectively. In the following, we will therefore assume that $\beta_2 > \frac{1}{2}$. A strategy that assigns $2\beta_2$ voters to half of the districts is then infeasible. However, by Proposition 2, party R can overcome this difficulty when L is large, by using a different strategy.

Proposition 2 *Let $N \geq 3$. There is $\hat{L}$ so that, for all $L \geq \hat{L}$, there exists a strategy σ_R so that $\Pi_R(\sigma_D, \sigma_R \mid \omega > \hat{\omega}) = 1$, for all σ_D.*

4.1 Proof of Proposition 2

For Party R to win a majority of seats whenever $\omega > \hat{\omega}$, there need to be at least N districts with a percentage share of type 1 voters that is weakly below β_1. We now show that, for L large, party R indeed has a strategy available that ensures this outcome.

A zero-sum game. Our analysis builds on a specific zero-sum game. To reiterate, we do not assume that the parties are actually playing that game. However, an understanding of the equilibrium strategies of that game facilitates the proof of Proposition 2. The sequence of moves is as outlined in Section 3. To this game form we add the following payoff specifications: Party R gets a payoff of $\pi_R = 1$ when there are at least N districts with a type t_2 voter share of at least β_2. Otherwise, party R gets a payoff of $\pi_R = 0$. The payoff of party D is given by $\pi_D = 1 - \pi_R$.

In the following, we will assume that districts are ordered according to their share of type 1 voters, so that district 1 has the (weakly) lowest, and district $2N$ has the (weakly) highest share of type 1 voters. Lemma 2 in the Appendix shows that this is without loss of generality throughout all rounds. Strategies that involve a reordering of districts (say, adding voters in a way such that District 5 after the move has strictly more type 1 voters than District 6) can be shown to be weakly dominated by order-preserving strategies. Thus, whenever a party is called upon to move, it will only consider assignments that preserve the ranking of districts.

Lemma 1 *Let $N \geq 3$. There is $\hat{L}$ so that $L > \hat{L}$ implies $\pi_R = 1$ in equilibrium.*

A formal proof of Lemma 1 is in the Appendix.

Before we turn to an illustration of the main argument in the proof, we explain the significance of Lemma 1 for the proof of Proposition 2. Party R's equilibrium strategy in the zero-sum game allows it to hold the share of type t_1-voters in half of the districts (weakly) below β_1, for any strategy of party D. Consequently, if R plays the same strategy in the original redistricting game, it wins all of these districts whenever $\omega > \hat{\omega}$.

Could party D prevent this outcome by deviating from its equilibrium strategy in the zero-sum game? The answer is negative because the game is zero sum. Any equilibrium strategy of party R solves a maximin-problem, i.e., it maximizes R's payoff under the assumption that D's strategy is chosen to minimize the maximum attained by R; see e.g. Osborne and Rubinstein (1994). Thus, if D does not behave this way, the payoff R gets cannot go down. We therefore obtain the following Corollary to Lemma 1. This completes the proof of Proposition 2.

Corollary 1 *Let $N \geq 3$. There is $\hat{L}$ so that $L \geq \hat{L}$ implies the existence of a strategy σ_R so that $\Pi_R(\sigma_D, \sigma_R \mid \omega > \hat{\omega}) = 1$, for all σ_D.*

On the proof of Lemma 1. Remember that we can fix a district ranking, and focus on the zero-sum game being played in such a way that districts with lower numbers have a (weakly) lower share of type t_1-voters. For party R to secure a majority whenever $\omega > \hat{\omega}$, it needs to ensure that there are at least N districts so that, after L rounds of play, the percentage share of type t_1-voters is not higher than β^1. Thus, R's objective is to minimize, and D's objective is to maximize, the share of D partisans in district N.

Since party D seeks to maximize this share, it will not waste type t_1-voters in lower ranked districts. Thus, party D concentrates type t_1-voters in the $N + 1$ top-ranked districts. More specifically, whenever it is called upon to play in some round l, and plans to assign a certain mass of t_1-voters, the following pecking order is optimal: Assign t_1-voters to the district with rank N until its mass of t_1-voters is equal to the one in the district with rank $N+1$. From that point on, keep these two districts at a joint level and add further t_1-voters until this joint level equals the one in the district with rank $N+2$. From then on, the districts with ranks N, $N+1$ and $N+2$ are raised to the level of district $N+3$ and so on, until no further t_1-voters are left, see Figures 1 and 2 for an illustration, where $N=5$.

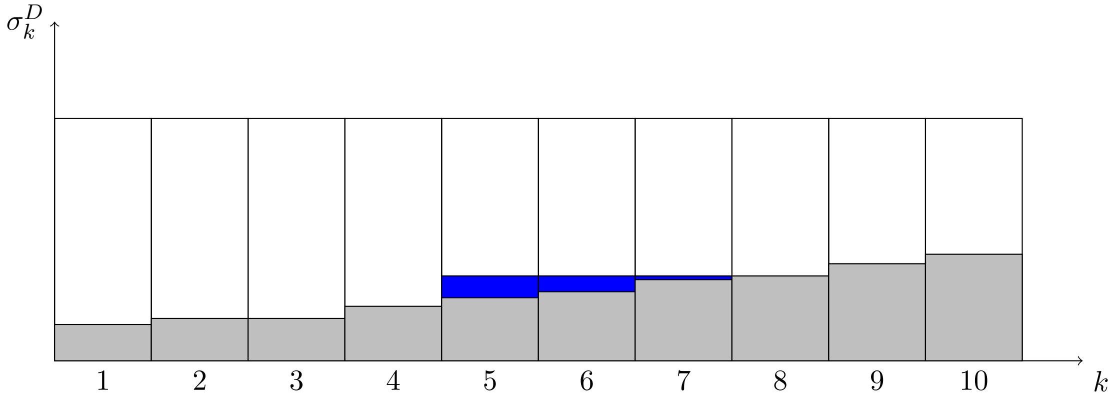

Figure 1: 10 Districts. In round l, D inherits, for every district, a stock of t_1-voters, illustrated in gray. It then adds further t_1-voters in round l, illustrated in blue. This figure is drawn under the assumption that D assigns only few t_1-voters in round l, so that, when assigning them optimally, its budget allows to raise the level of t_1-voters only in districts 5,6, and 7.

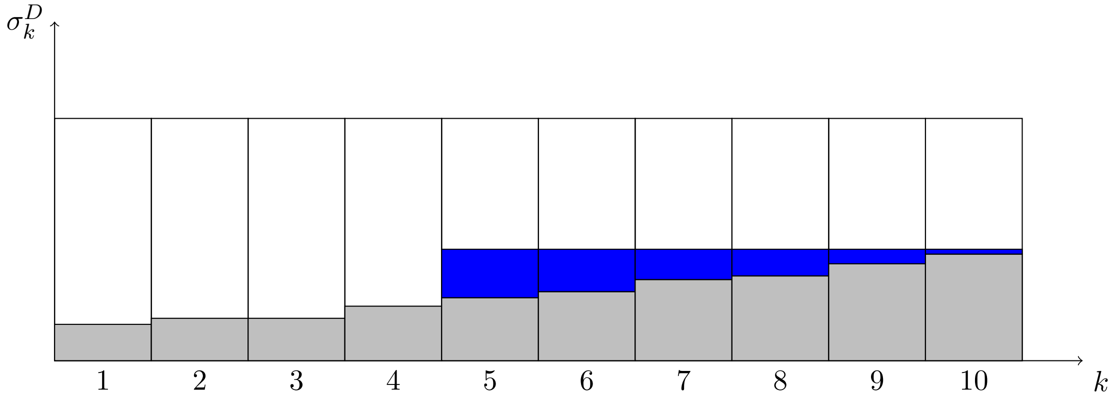

Figure 2: 10 Districts. In round l, D inherits, for every district, a stock of t_1-voters, illustrated in gray. It then adds further t_1-voters in round l, illustrated in blue. This figure is drawn under the assumption that D assigns many t_1-voters in round l, so that, when assigning them optimally, its budget allows to raise the level of t_1-voters in all districts with a rank weakly larger than 5.

What is an optimal response for party R? Its problem is to dispose of a total mass of $2N\beta^1$ t_1-voters in such a way that they contribute as little as possible to the mass of t_1-voters in district N. What is clearly harmless is to add t_1-voters to districts with ranks up to $N-1$, provided they are not yet at an equal level with the district that has rank N. Thus, when party R plans to assign some mass of t_1-voters in some round, it will first fill the bottom $N-1$ districts up to the point where a common level of t_1-voters is reached in the bottom N districts. This ensures a minimal level of t_1-voters in all districts; see Figure 3 for an illustration under the assumption that the mass of t_1-voters assigned in round l does not suffice to bring the bottom 4 districts to the level of district 5. Figure 4 is based on the alternative assumption that the mass exceeds what would be needed for that purpose.

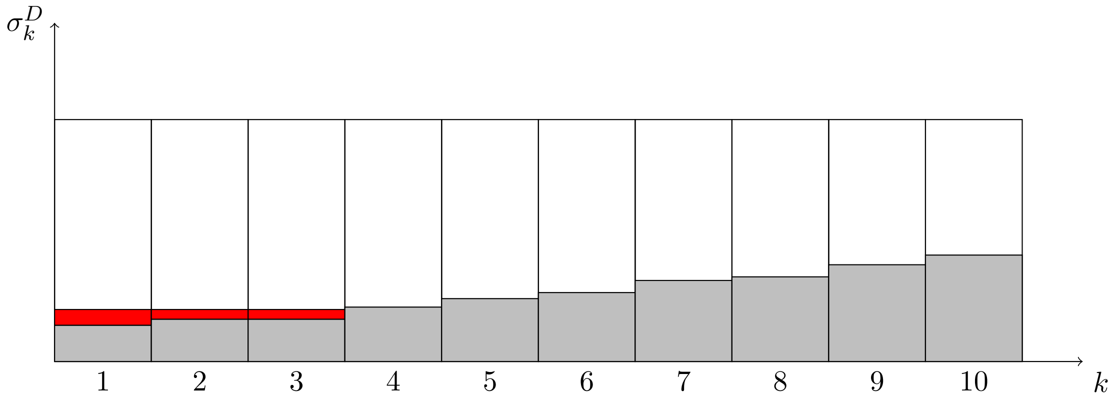

Figure 3: 10 Districts. In round l, R inherits, for every district, a stock of t_1-voters, illustrated in gray. It then adds further t_1-voters in round l, illustrated in red. This figure is drawn under the assumption that R assigns few t_1-voters in round l, so that, when assigning them optimally, the level in the districts with a rank below 5 cannot be raised to the level in the district with rank 5.

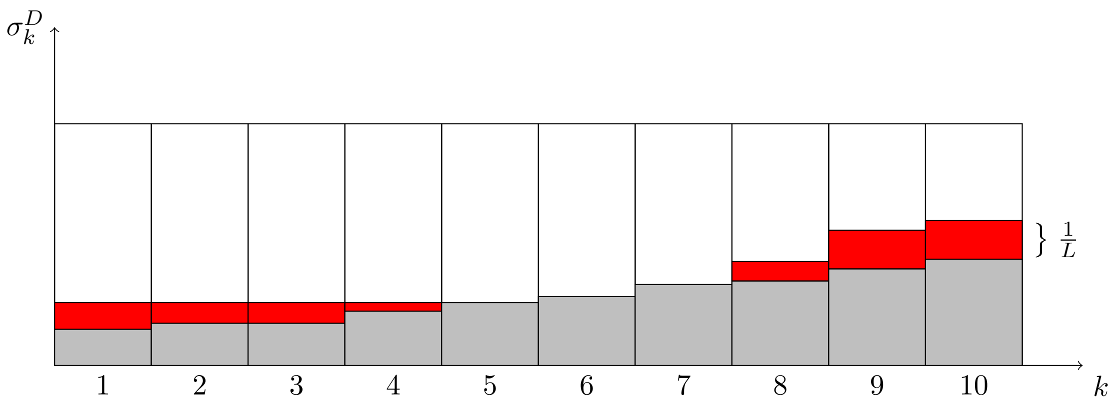

Figure 4: 10 Districts. In round l, R inherits, for every district, a stock of t_1-voters, illustrated in gray. It then adds further t_1-voters in round l, illustrated in red. This figure is drawn under the assumption that R assigns many t_1-voters in round l, so that, when assigning them optimally, the level in the districts with a rank below 5 is raised to the level in the district with rank 5. Additional t_1-voters are then assigned to the top-ranked districts.

Figure 4 illustrates the following logic: When additional t_1-voters need to be assigned after a common level in the bottom N districts has been achieved, party R continues with districts in the upper half. Here, it is optimal to start with the top-ranked district. If the capacity constraint of $\frac{1}{L}$ for that district in that round is reached, party R starts to fill the district with the second highest rank, and so on. Thus, party R concentrates on the top-ranked districts when assigning t_1-voters.

Party R discards the extra t_1-voters in very few districts in order to make it as difficult as possible for party D to "use" these t_1-voters in an attempt to raise the t_1-share in the pivotal district with rank N. To see intuitively why the distribution over non-pivotal districts matters at all, suppose instead that party R distributes the t_1-voters uniformly over districts $N+2$ to $2N$. That makes it easier for party D to raise the t_1-content of district $N+1$ in the next round: Remember that, when district $N+1$ reaches the level of district $N+2$, party D needs to allocate t_1-voters to both of these districts in order to avoid a district rank reversal. By allocating t_1-voters to the highest-ranked districts, R can insure that this no-rank-reversal constraint for party D kicks in as early as possible.

For a complete characterization of equilibrium strategies we would also need to describe how many t_1-voters are assigned by whom and when, i.e., we would need to characterize, for any party P and any round l the equilibrium value of β^1_{Pl}, defined as the percentage share of t_1-voters in the total mass of $\frac{2N}{L}$ voters assigned by party P in round l. We do not provide such a complete characterization, but show that R can choose the sequence $\{\beta^D_{Pl}\}^L_{l=1}$ so that the share of t_1-voters in district N remains below β_1. To this end, assume that R chooses $\beta^D_{R1} = 0$, and for any $l \geq 2$, $\beta^D_{Rl} = \beta^D_{Dl-1}$. Thus, R waits until D starts to assign t_1-voters and then assigns in, any round, as many t_1-voters as D assigned in the round before.

This implies that, after any of R's moves, the bottom $2N-2$ districts have the same level of t_1-voters, while there are some further t_1-voters in the top ranked district, and, possibly, also in the district with the second highest rank. To see this, suppose for concreteness, that D chooses $\beta^1_{D1} > 0$. Then, it will spread a mass of $\beta^1_{D1}\ \frac{2N}{L}$ t_1-voters evenly over $N+1$ districts. In round 2, R will use the mass of voters previously assigned to $N-1$ of those districts to have an equal level in the bottom half. The remaining mass of t_1-voters is then assigned to at most two top districts. See Figure 5 for an illustration. This pattern is now repeated over various rounds, with the implication that, after any move of R there is a joint level of t_1-voters in the bottom $2N-2$ districts.

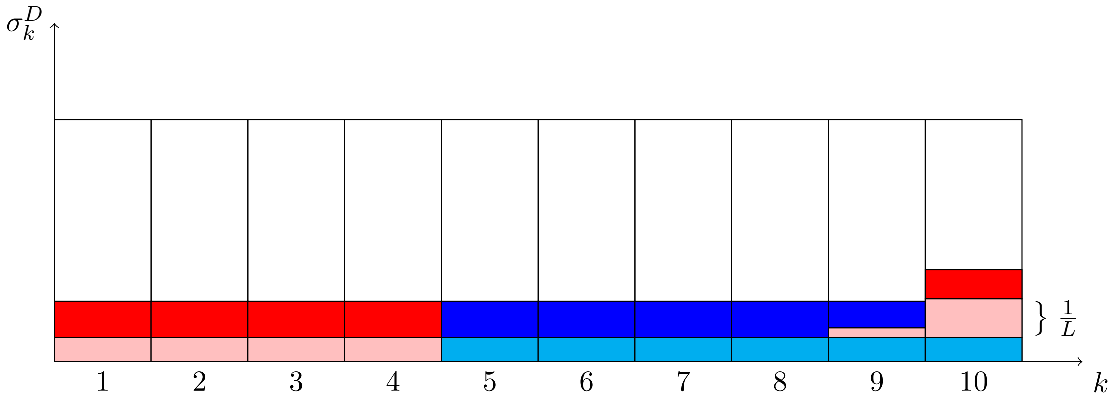

Figure 5: 10 Districts. R assigns as many t_1-voters as D did in the previous round. In light blue is the first round in which D assigns a positive mass of t_1-voters. R's response is in light red. In blue is the second round in which D assigns a positive mass of t_1-voters, and R's response is in red. As a consequence, there is a common level in the bottom eight districts, both after R's first and second response.

It is now easy to see that the share of t_1-voters in the pivotal district N cannot be strictly above β_1. This would imply a percentage share above β_1 in all districts and this is incompatible with the fact that the share of t_1-voters in the electorate at large is β_1.

Also note that there is a common level of t_1-voters in all districts, with the possible exception of the two top ranked ones. Thus, R's equilibrium strategy implies winning a majority whenever $\omega \in \Omega_R$, and moreover, implies that there are at most two districts that are "safe" for D. If the number of districts N is large, the fraction of districts where the outcome deviates from the popular vote is small.

5 Discussion

Our model shows that we can specify a dynamic game in which parties take turns in assigning voters to districts such that each party has a strategy that guarantees winning a majority in the legislature whenever it wins the popular vote. We now discuss some extensions.

Geographic constraints. As is standard in the gerrymandering literature, we do not impose geographic restrictions on the players. However, in contrast to the positive literature that makes this assumption, we also have a justification beyond tractability: Under the current redistricting system, the requirement that districts are contiguous can be interpreted as a second-best constraint to the gerrymanderer's power that (slightly) limits his ability to implement a map that distorts the popular vote outcome. Our mechanism directly gets rid of that power to distort election outcomes, so that indirect constraints for that purpose are unnecessary.

This said, suppose that parties, in addition to caring about their chance of winning in future elections, also prefer districts where voters live geographically close to each other, for example, in order for representatives to organize constituency services more conveniently.[10]

In this case, one could easily consider the outcome of the redistricting game only as a default endowment. If both parties agree, then reassignments of precincts (for example, to generate more compact districts) can certainly be permitted. A similar argument applies if there is a desire to bundle certain groups such as racial minorities for representation purposes. Note that dispensing with a contiguity requirement may actually makes it easier to form some minority districts.

Opposition representation. The analysis above shows that the proposed system of redistricting may yield an outcome so that most districts are replicas of the at-large districts. In this case, the majority-preferred party in an election wins a very large percentage of seats, with few or none going to the minority party.

Even though the minority party has limited influence on which policies are enacted even if it is represented in the legislature, this representation my have beneficial effects. For one, the minority can at least participate in the discussion of legislative proposals and provide additional information in this context, and, to the extent that they can persuade the majority party, they can have (possibly Pareto-improving) influence on policy. A strong opposition within the legislature may also be useful for providing information about legislative proposals to the public.

Finally, if legislative experience matters for performance, then the voters' opportunity to replace the current majority (if either voters' political preferences shift, or if the current majority party "misbehaves" and needs to be replaced for incentive reasons) is better if the opposition party contains at least some experienced legislators who do not have to learn from scratch how a legislature works.

So, how could we adjust our system if we wanted to guarantee a substantial oppo-

[10] Observe that the requirement of district contiguity in the current system does not guarantee that voters in a district live close to each other. See, for example, district TX-35 during the 2012-2020 time period.

sition representation in the legislature? One simple possibility is to turn each single-member district into a multi-member district.

For example, suppose that each district is represented by 3 legislators. Within each district, there is proportional representation (or some transferable vote system), so that the party that gets more votes in the district receives 2 representatives, and the other party the remaining seat if its vote share is above a threshold. The percentage of votes that is required to win one seat in a district of three representatives depends on the specific rules that map the votes obtained by the parties in the district to a seat allocation. For example, with both the Hare-Niemeyer procedure and the Webster/Sainte-Lague procedure (the methods used in German federal elections from 1987 to 2005, and after 2005, respectively), obtaining more than 1/6 of the vote entitles the weaker party in a district with three representatives to one seat.[11]

In this case, the redistricting game between the parties remains exactly the same as in the basic model, and the losing party is essentially guaranteed a representation of one-third in the legislature. In contrast to the current system with one representative per district, this system would also guarantee that each voter is represented, in the legislature, by (at least) one representative from his district and from his favorite party.

6 Concluding remarks

Our model provides one of the first normative analyses of gerrymandering, in the setting that is the "canonical" one in the large positive literature. We show that it is possible to neutralize the distortions due to partisan gerrymandering by having both parties participate in the redistricting process, in the sense that each party has a strategy that guarantees winning a majority in the legislature whenever it wins the popular vote.

While this possibility result is based on a particular sequential game, the protocol does not have be taken literally as a specific proposal for how redistricting should be done in practice. It is of theoretical value in that it provides an upper bound for what is achievable when the rules governing the redistricting process are well designed. Presumably, there are other protocols that also implement the popular vote, or at least something closer than the current system. Any such protocol must, however, have the property that the parties can keep each other in check. As the literature on partisan

[11]The methods would differ in the vote share that is required to guarantee the stronger party two seats if there are three or more parties.

gerrymandering has shown, when there is no possibility for the other party to interfere, there is also no hope to implement the popular vote.

References

Chen, Jowei and Jonathan Rodden, "Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures," *Quarterly Journal of Political Science*, June 2013, *8* (3), 239–269.

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

Friedman, John N. and Richard T. Holden, "Optimal Gerrymandering: Sometimes Pack, but Never Crack," *The American Economic Review*, 2008, *98* (1), 113–144.

Groseclose, Tim and James M. Snyder, "Buying Supermajorities," *The American Political Science Review*, 1996, *90* (2), 303–315.

Gul, Faruk and Wolfgang Pesendorfer, "Strategic Redistricting," *The American Economic Review*, 2010, *100* (4), 1616–1641.

Kolotilin, Anton and Alexander Wolitzky, "The Economics of Partisan Gerrymandering," Working Paper 2020.

Konrad, Kai A, *Strategy and dynamics in contests*, Oxford University Press, 2009.

Kovenock, Dan and Brian Roberson, "Generalizations of the General Lotto and Colonel Blotto games," *Economic Theory*, 2020, *7* (1), 5 – 22.

Laslier, Jean-François and Nathalie Picard, "Distributive Politics and Electoral Competition," *Journal of Economic Theory*, 2002, *103* (1), 106 – 130.

Lizzeri, Alessandro and Nicola Persico, "The provision of public goods under alternative electoral incentives," *American Economic Review*, 2001, *91* (1), 225–239.

_ **and** _ , "A drawback of electoral competition," *Journal of the European Economic Association*, 2005, *3* (6), 1318–1348.

Myerson, Roger, "Incentives to Cultivate Favored Minorities Under Alternative Electoral Systems," *American Political Science Review*, 1993, *87* (4), 856–869.

Osborne, Martin and Ariel Rubinstein, *A course in Game Theory*, MIT Press, Cambridge, MA., 1994.

Owen, Guillermo and Bernard Grofman, "Optimal partisan gerrymandering," *Political Geography Quarterly*, 1988, *7* (1), 5 – 22.

Vickrey, William, "On the prevention of gerrymandering," *Political Science Quarterly*, 1961, *76* (1), 105–110.

von Neumann, John, "Zur Theorie der Gesellschaftsspiele," *Mathematische Annalen*, 1928, *100* (1), 295—-320.

Zermelo, Ernst, "Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels," in "Proceedings of the fifth international congress of mathematicians," Vol. 2 Cambridge University Press 1913, pp. 501–504.

A Appendix

A.1 Proof of Proposition 1

Consider the following strategy for party D: In all rounds l, choose $\sigma^1_{kDl} = 0$, for $k \leq N$ and $\sigma^1_{kDl} = \frac{2\beta_1}{L}$, for all $k > N$. We seek to show that, with this strategy, for all σ_R and for all districts with an index $k > N$,

$$(\sigma^1_{kD} + \sigma^1_{kR})\ v(t_1, \omega) + (\sigma^2_{kD} + \sigma^2_{kR})\ v(t_2, \omega) \quad < \quad \frac{1}{2}\ , \tag{3}$$

whenever $\omega < \hat{\omega}$. Since the left-hand side of equation (3) decreases in ω, it suffices to show that

$$(\sigma^1_{kD} + \sigma^1_{kR})\ v(t_1, \hat{\omega}) + (\sigma^2_{kD} + \sigma^2_{kR})\ v(t_2, \hat{\omega}) \quad \leq \quad \frac{1}{2}\ , \tag{4}$$

or, equivalently, that

$$\sigma^1_{kD} + \sigma^1_{kR} \quad \geq \quad \frac{v(t_2, \hat{\omega}) - \frac{1}{2}}{v(t_2, \hat{\omega}) - v(t_1, \hat{\omega})} \quad = \quad \beta_1\ , \tag{5}$$

where the inequality in the left part of (5) follows from (4) upon using that $\sigma^2_{kD} = 1 - \sigma^1_{kD}$ and $\sigma^2_{kR} = 1 - \sigma^1_{kR}$. The equality in the right part of (5) then follows from (1).

After L rounds, the total mass of voters assigned by the two parties to any one district k equals 2. Under party D's strategy the share of type 1 voters is in any district with an index $k > N$ is bounded from below by β_1. To see this note that

$$\sigma^1_{kD} + \sigma^1_{kR} \quad \geq \quad \frac{L\ \frac{2\beta_1}{L}}{2} \quad = \quad \beta_1\ .$$

□

A.2 On the ranking of districts

Ordering districts. If the game were to end after round l, party R would win district k in state ω when

$$\sum_{j=1}^{2} v(t_j, \omega)\ \frac{L}{2l}\ \left(s^l_{Dk}(t_j) + s^l_{Rk}(t_j)\right) \quad > \quad \frac{1}{2}\ , \tag{6}$$

where $s^l_{Dk}(t_j) := \sum_{l'=1}^{l} \sigma_{Dkl'}(t_j)$ and $s^l_{Rk}(t_j) := \sum_{l'=1}^{l} \sigma_{Rkl'}(t_j)$ are the stocks of type t_j voters who have been assigned by parties D and R, respectively, over the first l rounds of play. To interpret this inequality, note that $\frac{L}{2l}\ (s^l_{Dk}(t_j) + s^l_{Rk}(t_j))$ is the share of type t_j voters among those voters who have been assigned to district k in the first l periods.

Thus, if ω is such that the above inequality holds, then party R has majority support in district k after round l.

Let $s_k^l(t_j) := s_{Dk}^l(t_j) + s_{Rk}^l(t_j)$. We define a rank order of districts according to their republican vote share after l rounds of play. Thus, the rank of district k is higher than the rank of district k' if, for some ω,

$$\sum_{j=1}^{2} v(t_j,\omega)\,\frac{L}{2l}\,s_k^l(t_j) \quad \geq \quad \sum_{j=1}^{2} v(t_j,\omega)\,\frac{L}{2l}\,s_{k'}^l(t_j)\ . \tag{7}$$

Using that, in any district the shares of type 1 and type 2 voters add up to 1, inequality (7) can equivalently be written as

$$\begin{aligned} &v(t_1,\omega) + \tfrac{L}{2l}\,s_k^l(t_2)\Big(v(t_2,\omega) - v(t_1,\omega)\Big) \\ &\geq v(t_1,\omega) + \tfrac{L}{2l}\,s_{k'}^l(t_2)\Big(v(t_2,\omega) - v(t_1,\omega)\Big)\ . \end{aligned} \tag{8}$$

or, more simply, as

$$s_k^l(t_2) \geq s_{k'}^l(t_2)\ .$$

Thus, ordering districts according to their republican vote share is equivalent to ordering them according to the share of type 2 voters. Also, the Republican vote share in any district k is, for every state ω, a monotonic function of the mass of type 2 voters.

Order preserving assignments. Assume without loss of generality that after l rounds of play district 1 has a weakly lower republican vote share than district 2, that district 2 has a weakly lower republican vote share than district 3 and so on. District $2N$ is then among those with a maximal republican vote share. Now consider round $l+1$. Suppose that party R moves first in round $l+1$. It then assigns a mass of $\frac{1}{L}$ voters to any district k. Thus, for any district k,

$$\sum_{j=1}^{2} \sigma_{kRl+1}(t_j) \quad = \quad \frac{1}{L}\ .$$

This move of R induces a new order of districts according to

$$s_k^l(t_2) + \sigma_{kRl+1}(t_2)\ .$$

Let $r_\sigma(k) \in \{1,\dots,2N\}$ be the new rank of the district with initial rank k.

Lemma 2 *Given a move $\sigma_{Rl+1} = (\sigma_{kRl+1})_{k=1}^{2N}$ of party R in round $l+1$ with a resulting ranking $k \mapsto r_\sigma(k)$ according to the republican vote share, there is an alternative move $\sigma'_{Rl+1} = (\sigma'_{kRl+1})_{k=1}^{2N}$ of party R with the following properties:*

i) The alternative move uses the same voter types: For every j,

$$\sum_{k=1}^{2N} \sigma_{kRl+1}(t_j) \quad = \quad \sum_{k=1}^{2N} \sigma'_{kRl+1}(t_j) \; .$$

ii) The alternative move preserves the old ranking; formally, it induces a new ranking $k \mapsto r_{\sigma'}(k)$ so that $r_{\sigma'}(k) = k$, for every k.

iii) The republican vote share in the district with rank $N+1$ under the alternative move σ'_{Rl+1} is at least as high as in the district with rank $N+1$ under the initial move.

Proof of Lemma 2. Suppose there is some district with initial rank k' that has rank k in the ranking induced by $\sigma_{Rl+1} = (\sigma_{kRl+1})_{k=1}^{2N}$; i.e. $k' = r^{-1}(k)$. The mass of type t_2 voters after R's move under $\sigma_{Rl+1} = (\sigma_{kRl+1})_{k=1}^{2N}$ is given by

$$s^l_{k'}(t_2) + \sigma_{k'Rl+1}(t_2) \; .$$

We now choose $\sigma'_{kRl+1}(t_2)$ so that

$$\sigma'_{kRl+1}(t_2) \quad = \quad \max\left\{0, s^l_{k'}(t_2) + \sigma_{k'Rl+1}(t_2) - s^l_k(t_2)\right\} \; .$$

Proceeding in the same way for all k implies that

$$s^l_k(t_2) + \sigma'_{kRl+1}(t_2) \quad \geq \quad s^l_{r^{-1}(k)}(t_2) + \sigma_{r^{-1}(k)Rl+1}(t_2) \; .$$

The mass of type t_2 voters used by $\sigma'_{Rl+1} = (\sigma'_{kRl+1})_{k=1}^{2N}$ across all districts is such that

$$\sum_{k=1}^{2N} \sigma'_{kRl+1}(t_2) \quad = \quad \sum_{k=1}^{2N} \max\left\{0, s^l_{r^{-1}(k)}(t_2) + \sigma_{r^{-1}(k)Rl+1}(t_2) - s^l_k(t_2)\right\}$$

An upper bound is obtained under the assumption that

$$s^l_{r^{-1}(k)}(t_2) + \sigma_{r^{-1}(k)Rl+1}(t_2) - s^l_k(t_2) > 0 \; ,$$

for all k, i.e. so that type t_2 voters have to be assigned to all districts. Therefore,

$$\begin{aligned}
\textstyle\sum_{k=1}^{2N} \sigma'_{kRl+1}(t) &\leq \textstyle\sum_{k=1}^{2N} s^l_{r^{-1}(k)}(t) + \sigma_{r^{-1}(k)Rl+1}(t) - s^l_k(t) \\
&= \textstyle\sum_{k=1}^{2N} s^l_{r^{-1}(k)}(t) - \sum_{k=1}^{2N} s^l_k(t) + \sum_{k=1}^{2N} \sigma_{r^{-1}(k)Rl+1}(t) \\
&= \textstyle\sum_{k=1}^{2N} \sigma_{r^{-1}(k)Rl+1}(t) \, .
\end{aligned}$$

Thus, σ'_{Rl+1} does not use more type t_2 voters than σ_{Rl+1}, and it yields, in any district, at least as type 2 voters in total. If there is a strict inequality, i.e. if σ'_{Rl+1} use strictly less type t_2 voters than σ_{Rl+1}, then those voters can be assigned to the districts in such a way that the initial ranking is preserved. □

A.3 Proof of Lemma 1

A strategy for party R. In any round l, given a – for now exogenous – budget of $\beta^1_{Rl} \frac{2N}{L}$ type t_1-voters to be assigned, proceed sequentially in the following way – until the budget of type t_1-voters for that round is exhausted:

i) Add type t_1-voters to the lowest ranked district until the mass of t_1-voters equals the mass in the district with the second lowest rank. From then on, keep the mass in these two districts equal.

ii) Add type t_1-voters to the two lowest ranked districts until the mass of t_1-voters equals the mass in the district with the third lowest rank. From then on, keep the mass in these two districts equal.

iii) Proceed analogously for all districts with a rank smaller or equal $N-2$. From then on, keep the mass in all these districts equal. Add t_1-voters to the $N-1$ lowest ranked districts until the mass of t_1-voters equals the mass in the district with rank N. From then on, don't add further t_1-voters to one of the bottom N districts.

iv) Add t_1-voters to the top ranked district.

v) If there are still t_1-voters left in the budget after a mass of $\frac{1}{L}$ t_1-voters has been assigned to the top ranked district, add t_1-voters to the district with the second highest rank, etc, then move to the district with the third highest rank, etc.

vi) Stop when no further t_1-voters are left.

Note that, as an implication, R's play in any round leaves the ranking of districts unchanged.

A best response for party D. Consider a – for now exogenous – sequence of budgets for party D's play $\{\beta^1_{Dl}\}_{l=1}^{L}$. Note that since party R never affects the ranking of districts, the ranking of districts in any round is entirely due to party D. As argued above it entails no loss of generality to assume that party D's moves do neither affect the ranking of districts. This also implies that it is never optimal to have a budget of partisan D voters in some round that makes it necessary to assign D voters to strictly more than $N+1$ districts. Thus, we may assume that, for any round l,

$$\beta^1_{Dl} \frac{2N}{L} \quad \leq \quad \frac{N+1}{L},$$

or, equivalently,

$$\beta^D_{Dl} \quad \leq \quad \frac{1}{2} + \frac{1}{2N}.$$

Given some budget for moves in round l, the optimal strategy for party D is now as follows:

i) Add type t_1-voters to the district with rank N until the mass of t_1-voters equals the mass in the district with the rank $N+1$. From then on, keep the mass in these two districts equal.

ii) Add type t_1-voters to the two districts with ranks N and $N+1$ until the mass of t_1-voters equals the mass in the district with rank $N+2$. From then on, keep the mass in these three districts equal.

iii) Proceed analogously for all districts with a rank larger or equal $N+2$, until the budget of D voters is exhausted.

Party R's sequence of budgets. We now specify a particular sequence of budgets for party R: As the first mover in the initial round, it does not assign any type t_1-voters, $\beta^1_{R1} = 0$. In any round $l \geq 2$, and as long os this is feasible, party R assigns as many t_1-voters as party D did in the previous round

$$\beta^1_{Rl+1} \quad = \quad \beta^1_{Dl} \; .$$

This is clearly feasible in early rounds. If, however, party D keeps some type t_1-voters for the last round so that $\beta^1_{DL} > 0$, then party R will have to assign an additional mass of $\beta^1_{DL} \leq \frac{2N}{L}$ type t_1-voters somewhen in the game. Otherwise party R would violate its overall budget constraint. Note that this quantity vanishes for $L \rightarrow infty$.

Thus, there is a subset of rounds L' so that

$$\sum_{l' \in L'} \beta^1_{Dl'} < \sum_{l' \in L'} \beta^1_{Rl'+1} \quad \leq \quad \sum_{l' \in L'} \beta^1_{Dl'} + \frac{2N}{L} \; .$$

and for l not in L' we let

$$\beta^1_{Rl+1} \quad = \quad \beta^1_{Dl} \; .$$

Party R's strategy has the following implication: Whenever party R moves, it brings the mass of t_1-voters in the bottom $N-1$ districts to the level that party D has generated for the district with rank N in the previous round. Moreover, party R adds t_1-voters at most to the two top-ranked districts, and does not assign any D voters to districts with the ranks $N, N+1, \ldots, 2N-2$.

To see this, first consider rounds 1 and 2:

- In round 1, party D assigns an equal mass of D voters to $N+1$ districts.
- In round 2, party R fills the bottom $N-1$ districts. It then has additional t_1-voters left. According to party R's strategy, as many t_1-voters as possible are assigned to the district with the top rank $2N$. If additional t_1-voters are left, they go to the district with rank $2N-1$ and then, possibly, to the district with rank $2N-2$.

Now consider rounds 3 and 4:

- In round 3, party D's best response stipulates to assign an equal mass of t_1-voters to the districts with ranks $N, N+1, \ldots, 2N-2$. Those are $N-1$ districts. Possibly, it also assigns t_1-voters to the three top ranked districts.
- In round 4, party R fills the bottom $N-1$ districts. It can do so by adding to the districts in the bottom $N-1$ exactly the amount of D voters that party D has added to the districts with ranks $N, N+1, \ldots, 2N-2$ in round 3.
- If party D has previously added t_1-voters to the two top ranked districts, then party R has additional t_1-voters left after the bottom $2N-2$ districts have been leveled. Again, by party R's strategy, of these voters as many as possible are assigned to the district with the top rank $2N$. If additional t_1-voters are left, they go to the district with rank $2N-1$.

Completing the argument. Suppose first that, for all l,

$$\beta^D_{Rl+1} \quad = \quad \beta^D_{Dl} \, .$$

The strategies of parties R and D described above then imply that after the last move in round L, there is an equal mass of type t_1-voters for all districts with a rank smaller or equal to $2N-2$. The mass of these voters is (weakly) larger in the two top ranked districts. Now suppose that the percentage share of t_1-voters in the district with rank N is strictly larger than β_1. Equivalently, the total mass of t_1-voters in that district exceeds 2 β_1. Then, the mass of t_1-voters exceeds 2 β_1 in all districts. Hence, the total mass of assigned t_1-voters is strictly larger than $4N\ \beta_1$. But this is infeasible as the two parties' total endowments with partisan t_1-voters only sum to $4N\ \beta_1$. Thus, the assumption that party D can generate $N+1$ districts with a percentage share of type t_1-voters strictly larger than β_D leads to a contradiction, and must be false.

Now suppose, there needs to be a subset of rounds L' so that

$$\sum_{l'\in L'} \beta^1_{Dl'} < \sum_{l'\in L'} \beta^1_{Rl'+1} \quad \leq \quad \sum_{l'\in L'} \beta^1_{Dl'} + \frac{2N}{L} \, .$$

For L sufficiently large, we can chose the number of such rounds equal to $2N$, i.e. $\#L' = 2N$. Party R can then satisfy its overall budget constraint by assigning, for every round $l' \in L'$, an additional mass of t_1-voters that is bounded from above by $\frac{1}{L}$.

Then, party R's moves in rounds $l' \in L'$ may require to add type t_1-voters to the three highest ranked districts, with the mass going to the district with rank $2N-2$ being bounded from above by $\frac{1}{L}$. The strategies of parties R and D described above then imply that after the last move in round L, there is an equal mass of type t_1-voters for all districts with a rank smaller or equal to $2N-3$. The mass of these voters is (weakly) larger in the three top ranked districts. Again, the assumption that the percentage share of t_1-voters in the district with rank N is strictly larger than β_1 leads to a contradiction, and must be false. $\square$