ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 034

Competitive Gerrymandering and the Popular Vote

Felix Bierbrauer Matthias Polborn

October 2020 www.econtribute.de

UNIVERSITÄT BONN Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866 is gratefully acknowledged.

Competitive Gerrymandering and the Popular Vote*

Felix J. Bierbrauer† Mattias Polborn‡

October 16, 2020

Abstract

Gerrymandering undermines representative democracy by creating many uncompetitive legislative districts, and generating the very real possibility that a party that wins a clear majority of the popular vote does not win a majority of districts. We present a new approach to the determination of electoral districts, taking a design perspective. Specifically, we develop a redistricting game between two parties who both seek an advantage in upcoming elections, and show that we can achieve two desirable properties: First, the overall election outcome corresponds to the popular vote. Second, most districts are competitive.

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

*We benefitted from conversations with Peter Cramton, 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

State legislatures in the United States are generally elected from single-member districts. The process by which the state is divided up into districts differs from state to state. However, in most states, the main power lies with the state legislatures themselves that pass by majority vote, after each decennial census, a new map of districts.

In this process, the state legislature (or, more precisely, the majority party that controls the legislature) faces some legal constraints. For example, the population in each district cannot deviate too much from the average district size; the districts have to be contiguous; too-overtly racial gerrymandering may be banned by the Voting Rights Act. However, even with those constraints, being able to decide on the allocation of voters to districts is hugely advantageous to the party in control of the redistricting process.

For example, Republicans took over the Pennsylvania state legislature in the 2010 Republican wave election and used the opportunity to create a district map that is very favorable to them. 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 control 110 out of 203 seats in the Pennsylvania House of Representatives. Similarly, 63 out of 99 representatives in the 2019/2020 Wisconsin State Assembly (the lower chamber of the legislature) are Republicans, in spite of Democratic candidates winning the corresponding state-wide aggregated vote 53% to 45%.

While, in the current decade, the gerrymandering advantage is mostly with Republicans because they had a very strong showing in the 2010 election that gave them control over the last redistricting process, in the 1990s and 2000s, Democrats managed, through gerrymandering, to hold on to majorities in Southern state legislatures at a time when these states were solidly Republican in all presidential elections.[1]

The resulting disconnect between the political preference of the majority of the electorate and the election outcome is problematic for two reasons. First, and most obviously, there is a representation problem: A voting system where a minority of voters consistently gets to decide what the majority should do lacks democratic legitimacy.

Second, if most representatives are elected in districts that are not competitive in the general election, they will have insufficient incentives for good behavior, both in terms of valence provision and in terms of positioning.[2] Regarding valence provision, think of constituency service; an incumbent who is more or less guaranteed reelection is less likely to feel compelled to provide adequate service for his constituents than one who feels that, if enough voters are unhappy, his job is in danger.

Regarding positioning, the problem in predominantly "blue" or "red" districts is that incumbents are more concerned with the primary election than with the general election and are therefore more likely to cater to the median primary voter in their party than to

[1] At the federal level, McCarty et al. (2009) show that gerrymandering has increased the Republican seat share in the House of Representatives.

[2] See, e.g., Callander (2005), Van Weelden (2015) and Krasa and Polborn (2018).

the median voter in the district. And, if (what seems reasonable) voters have preferences that depend both on local candidates positions and on those of the national or state party, then primary voters of the party whose national position is sufficiently favored by the district median voter do not have to nominate a candidate whose position appeals to this district median voter in order to win — rather, a candidate who mostly appeals to the party faithful is essentially unbeatable in the general election (Krasa and Polborn, 2018).

Because of these problems, there is substantial backlash against existing gerrymanders and also the institutions that allow for it to happen. Existing gerrymanders (i.e., the 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.

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."[3]

While one might suspect that justices' partisan preferences add to their unwillingness to interfere with the specific gerrymandering cases brought before them, it is genuinely difficult to define general abstract rules that can be used to determine which district maps should still be acceptable and which ones should be illegal. Measures of electoral district compactness such as the Polsby–Popper 1991 Test or measures of "wasted votes" such as the Efficiency Gap define "ideal" fair situations and provide some measure of how far away a particular redistricting map is from that ideal, but are clearly conceptually unsatisfying. For example, the plaintiffs in *Gill v. Whitford* argued that an Efficiency Gap of more than 7% indicates illegal gerrymandering; but why not 5% or 10% as limit instead?

In this article, we argue that, rather than coming up with an ideal measure of gerrymandering and a necessarily somewhat arbitrary boundary line between "still legal" and "sufficiently outrageous to be illegal," it is more fruitful to design a system of rules for a redistricting process that leads to, in equilibrium, a map that has certain desirable features.

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. This process in which two self-interested agents participate is more likely to lead to a fair outcome than the alternative of devising general rules and constraints under which only one child chooses both their own and the other child's piece.

Specifically, we provide the rules for a redistricting game in which both parties participate and want to maximize their respective probability of winning a majority of seats in

[3]See `https://en.wikipedia.org/wiki/Vieth_v._Jubelirer`, and also the US Supreme Court *Gill v. Whiford* and *Benisek v. Lamone* decisions in 2018, upholding *Vieth*.

the legislature in a future election. Our primary objective is that the district map generated in equilibrium has the property that the party that wins the support of a majority of the electorate in an election always also wins a majority of seats in the legislature.

In our setting, there are Democratic and Republican partisans (i.e., voters who can be relied on voting for their party for sure), as well as independents. Some independents vote Republican, and some Democrat, and the exact split depends on a random state of the world. The redistricting game yields an order of districts according to the net advantage of, say, the Democrats. Whatever the state of the world, the party that wins the median district according to that order wins a majority of districts. Thus, in the redistricting game, each party seeks to maximize the probability of winning the median district.

We analyze a sequential districting system in which parties assign voters to $2N$ districts over L rounds of play. Each party starts with a budget set that contains all voters in the state (i.e., each voter is both an initial element of the Democratic and the Republican budget set), and has to be completely distributed at the end. Thus, in each round, each party assigns a fraction of $1/(2NL)$ of all voters to each district, and each district eventually ends up with a fraction of $1/2N$ of all voters in the state. In addition to the regular districts, there is also one at-large district, so that there are in total $2N+1$ representatives in the legislature.

Our main result is that, if both parties play this assignment game, then the equilibrium district map has the property that the party that is favored by a majority of voters statewide always (i.e., in every state of the world) wins the majority of legislative districts.[4]

While the game is too complex to derive the exact equilibrium play, we prove this result by constructing, for each party, a strategy that allows it to win, against any opponent strategy, a majority of districts whenever it wins the majority of the statewide votes.

The intuition for this result is as follows. Observe first that, for a party to win a majority in the legislature more often than it wins the statewide popular vote, it is necessary to introduce an asymmetry between the median and the average district.[5] For example, if Republicans want to win even in some cases in which they lose the popular vote, they need a district map in which the median district has more Republican partisans or fewer Democratic partisans than the statewide average.

In a system in which one party can unilaterally decide on district allocations (such as the one currently in place in most states), such an asymmetry is easy to achieve. In contrast, the key to the fairness of our proposed sequential system is that, when Republicans attempt to generate such an asymmetry in a round (essentially, by allocating disproportionately more Republican voters to slightly more than half of the districts, and more

[4]This result holds in the limit as the number of rounds L is sufficiently large.

[5]When we talk about the "median" district here, we think of districts ordered with respect to their net partisan advantage.

Democratic voters to the remainder of districts), Democrats can react by balancing, i.e., making the Republican-targeted districts more Democratic, and the remainder districts more Republican. Thereby, we show, they can offset the asymmetry that Republicans attempt to create.

The remainder of the paper is organized as follows. In Section 2, we comment on related literature. A simple example is presented in Section 3. Section 4 introduces our general framework, with the main results of the formal analysis in Section 4.2. Sections 5 and 6 contain discussions of extensions, robustness and conclusions. Proofs are collected in the Appendix.

2 Related Literature

This paper uses ideas from mechanism and market design and applies them to gerrymandering. We share with the literature on mechanism design that we seek to implement a particular outcome –namely, an assignment of voters to districts such that the party that wins the popular vote always wins a majority of seats in the legislature,– and ask whether we can find a game that generates this result as an equilibrium outcome.[6] There is a trivial "solution" of this problem, which is to have proportional representation of parties based on a single national district. However, such a "solution" would eliminate the connection between local constituencies and their representatives, whose legitimacy is based on majority support in that district. So, we take as given that there have to be many districts and that all voters have to be assigned to a new district map from time to time. Working with these predetermined institutional constraints – rather than proceeding axiomatically on the basis of a game-theoretic solution concept – is a similarity to many papers on market design.[7]

The question how to organize competition for a market, for example in natural monopolies, is a central theme in market design, see e.g. Iossa et al. (2019). In this context, the concern is to benefit from scale economies while limiting the monopolists's abuse of market power. Limiting power abuse in the design of electoral districts is a different application and requires very different modeling choices. Still, the overarching question is the same: How to organize competition for electoral districts at an ex ante stage, so as to ensure that competition, when it takes place, yields desirable outcomes ex post.

The theory of mechanism design and implementation theory is multifaceted.[8] We in-

[6]William Vickrey emphasized the need for "an automatic and impersonal procedure for carrying out a redistricting" already in the early in the 1960s, see Vickrey (1961). He did not approach this challenge from a mechanism design perspective, though. Moreover, the class of mechanisms that he helped to develop, the well-known Vickrey-Clarke-Groves (VCG) mechanisms, align private interests with an objective of surplus-maximization. The mechanism developed in this paper focuses on a different alignment, namely, the alignment of political party incentives with the legitimacy of an election outcome.

[7]See Roth (2002) for an outline of the market design agenda.

[8]For textbooks treatments of the mechanism design problem, see e.g. Mas-Colell et al. (1995) or

sist on finding a game so that *every* equilibrium implements the popular vote, as opposed to the more modest objective of showing that there is some game with some equilibrium that implements the popular vote, see Jackson (2001) for a discussion. We also consider a game with symmetric rather than asymmetric information and use the solution concept of a subgame perfect equilibrium, see e.g. Moore and Repullo (1988) for a seminal reference.

Our paper develops a notion of implementation by means of 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.

Previous theoretical literature on gerrymandering has analyzed the decision problem of the party that happens to be in power when redistricting is due. We depart from this perspective and, instead, propose a system in which district boundaries are determined as the outcome of a game that is played by two competing parties. This game has some similarities with the divide-the-dollar game, but an important difference is that our system does not give rise to Condorcet cycles and mixed strategy equilibria. Instead, it involves a sequential mechanism that implements the popular vote as a pure strategy equilibrium.[9] In the following, we discuss the connection to the previous literature on gerrymandering in more detail. Further remarks on the divide-the-dollar game can be found in Section 5.

Coate and Knight (2007) define a district map that is, in a utilitarian sense, optimal, and provide conditions under which the optimum can be reached as an outcome of gerrymandering. The main differences to our approach is that we focus on a different objective (implementing a majority for the popular vote winner), and that we construct a game that achieves this objective as an equilibrium outcome. Most of the technical aspects of our settings are similar: There is an exogenous number of districts; the only constraint imposed on gerrymanders is that all districts have to contain the same number of voters; and the electorate is composed of partisan and independent voters, with uncertainty about election outcomes resulting from aggregate shocks to the behavior of independents.

Friedman and Holden (2008) and to Gul and Pesendorfer (2010) both focus on (selfishly) optimal gerrymandering strategies in a setting in which one party assigns voters to districts in order to maximize its advantage in future elections. In Friedman and Holden

Börgers et al. (2015).

[9]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. The design of sequential mechanisms is also a theme in auction theory, see e.g. Benoît and Krishna (2001).

(2008), the incumbent party observes noisy signals of the voters' party preferences and thus can order voters according to the probability that they will support it. The paper investigates whether 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). [10] While there is a difference in focus, a similarity between Friedman and Holden (2008) and our work is that both provide a fully microfounded analysis of the assignment of voters to districts.

Gul and Pesendorfer (2010) consider a strategic game of gerrymandering, but assume that each party has an own territory that it controls. Thus, in equilibrium, the districts drawn in the territory of party 1 are a best response to the districts drawn in the territory of party 2 and vice versa. Again, a key question is to what extent there is cracking and packing in equilibrium. In our setting, parties also have incentives for cracking and packing. However, in equilibrium, parties neutralize each other in their attempts to create districts in which they are favored to win, and thus, each party wins if and only if it wins the popular vote.

3 A simple example

Before we proceed to the general model, it is helpful to present a simplified example that encompasses the main ideas. Our inspiration comes, to some extent, from mechanisms for the fair division of a private good. For the cake-cutting problem in the introduction, a fair outcome can be decentralized by means of a mechanism so that one party can take advantage of any bias created by the other party. A mechanism for the determination of districts, however, has to deal with two additional complications: First, only one party can win a majority, i.e. there is no perfect divisibility. Second, the winning party should be the one with more support in the electorate at large.

A related problem arises in competitive chess when the winner of a tournament needs to be found via a decisive "Armageddon" game that cannot end in a draw. In one version, *White* has to win (i.e., a draw is equivalent to a win for *Black*). To compensate for this big advantage, both players submit a "bid" for the right to play as *Black*: While *White* has a fixed time budget for the game (e.g., 30 minutes), each player submits a bid indicating the lowest time budget that they would accept to play *Black*, and the lower bidder gets to play *Black*. Thus, as in the districting problem, ultimately a winning party needs to be found, and it is desirable that the stronger party wins. Moreover, the competing parties jointly determine the rules of the game. The following example illustrates how we bring these ideas to gerrymandering.

Suppose there are two parties, D and R, and three districts. In the electorate at large, there are 600 independent voters, 300 voters who always vote for party R (R partisans) and 300 voters who always vote for party D (D partisans). There is a state of the world

[10] This terminology goes back to Owen and Grofman (1988).

$\omega \in [-1, 1]$ that determines the Republican margin of victory in the pool of independent voters. For instance, $\omega = 1$ indicates that all independents vote for party R, $\omega = \frac{1}{2}$ indicates that 75 percent of the independents vote for R and 25 percent vote for D, and so on. Consequently, party R wins the popular vote when $\omega > 0$ and party D wins the popular vote when $\omega < 0$. We take ω to be the realization of a random variable.

We now look at the following game of district determination: In each district, there are 200 independent voters, and there is nothing the parties can do about this. (This is relaxed in the more general analysis below.) The 600 partisan voters, by contrast, need to be allocated, and moreover, districts have to be equal-sized. Thus, every district receives 200 partisan voters and what needs to be determined is the mix between R partisans and D partisans. After districts are determined, an election takes place, and whoever wins at least two districts wins the election.

An ideal gerrymander for a party is the one that maximizes its probability of winning the election. For instance, the ideal gerrymander for party R, assigns 200 D partisans to one district. The two remaining districts are each assigned 50 D partisans and 150 R partisans. Consequently, R wins the election whenever $\omega > -\frac{1}{2}$. Thus, whenever ω lies between $-\frac{1}{2}$ and 0, party D wins the popular vote, but party R wins a majority of districts.

The following mechanism alleviates this problem: Each party comes up with an own assignment of partisan voters to districts. This is done sequentially: First, party D presents its assignment. Party R observes it and then proposes its assignment. Thus, each voter is assigned twice, once by each party. In total, each district now has 400 independent voters, and is assigned a total of 400 partisans by the two parties.[11] Suppose party D assigns 100 D partisans and 100 R partisans to each district (one can show that Party D cannot do better). The best response of party R then is, again, to have 200 D partisans in one district, and two further districts with 50 D partisans and 150 R partisans each. In this case, party R wins the election whenever $\omega > -\frac{1}{4}$. Compared to the gerrymander in the previous paragraph, there are now more states in which the party that wins the popular vote is also the party that wins the election. However, there are still states of the world where this is not the case.

As we now argue, playing this game over many rounds gets us closer to an implementation of the popular vote. Specifically, suppose that, in every round, each party assigns two voters to every district, with D moving first in every round. Thus, in total, six voters are assigned by each party in every round. We claim that, in equilibrium, party D wins the election whenever $\omega < 0$ and party R wins the election whenever $\omega > 0$. The proof of this claim is constructive: We first argue that party R has a strategy that ensures winning the election whenever $\omega > 0$, i.e. in these states party R wins whatever the strategy

[11] For now, we stick to the assumption that independent voters are exogenously assigned and distributed evenly over districts. With one assignment per party, there are 400 independent voters in each district, 200 per party.

played by party D. Analogously, we show that party D can ensure to win almost surely whenever $\omega < 0$, provided that the number of rounds is sufficiently large.

Consider the following strategy for party R: In each round, simply "neutralize" party D's choice. For each district, assign an R partisan for every D partisan assigned by party D in the previous round, and assign a D partisan for every R partisan. Clearly, this strategy is feasible and ensures that parties D and R have the same number of partisan supporters in each district. Consequently, when R plays this strategy, R can ensure to win whenever $\omega > 0$.

Party D has to move first in every round and therefore cannot follow a strategy that achieves an exact neutralization of R's moves. However, when there are many rounds, the discrepancy becomes small. To see this, consider the following strategy for party D: In the first round, assign one D partisan and one R partisan to every district. From the second round on, neutralize R's move in the previous round, i.e. match every R partisan assigned by party R with a D partisan and every D partisan with an R partisan. While neutralization is not feasible in the last round (because D does not get to respond to R's last move), up to that point, districts remain balanced. With the last move, R may be able to gain an advantage, but it is easy to see that the size of this advantage in the median district is at most 1. Thus, when the number of rounds is large, party D gets close to an exact neutralization, and hence can ensure to win the election whenever $\omega < 0$.

Note that we do not claim that these strategies are optimal for the parties in every subgame, only that they are feasible and thus provide a lower payoff bound. In fact, it is quite clear that there are subgames in which R's strategy is not optimal. For example, if, after D's last move in the, say, 100th round, the net R-advantage by district is $(-2, 0, 0)$, then, rather than to equalize the forces in the first district, it is better for party R to use the final move to create an advantage in battles 2 and 3.

The general lessons from this stylized example are as follows: If one party gets to allocate all voters, it enjoys a large advantage in terms of winning probability. This is true, albeit to a lesser extent, if we impose additional constraints on what allocations can be chosen. Additional constraints (e.g., districts have to be "contiguous," i.e., only certain voter combinations are feasible assignments) may limit the advantage of the party that gets to decide on the voter allocation, but will generally not eliminate it completely. Designing a game in which both players participate in drawing up the allocation of voters to districts has the potential to create a completely fair outcome. Additional constraints are not needed to obtain the desirable outcome. They may actually be unhelpful, as they not only constrain the party presenting a first draw, but also constrain the other party' response.

4 A General Framework

We now introduce our general framework. The previous section provides a special case in which a biased election outcome due to strategic gerrymandering can be avoided. The purpose of the more general framework introduced here is to see whether these insights extend to settings so that the two parties may have different numbers of partisan supporters, an endogenous assignment of independent voters and an arbitrary number of districts.

There are $2N$ districts, indexed by $k \in \{1, 2, \ldots, 2N\}$ and an at-large district. The electorate consists of voters who always vote Republican (R partisans), voters who always vote Democrat (D partisans) and independent voters. The mass of Republican partisans, Democrat partisans and independent voters in the electorate at large is, respectively, given by

$$b_R = 2N\,\beta_R\ , \quad b_D = 2N\,\beta_D\ , \quad b_I = 2N\,\beta_I\ , \quad \text{where} \quad \beta_R + \beta_D + \beta_I = 1\ .$$

We assume, without loss of generality, that $\beta_R \leq \beta_D$. We also assume $\beta_D \leq \frac{1}{2}$. Absent this assumption, there would be no uncertainty about which party wins the popular vote.

Let p_R be the probability that an independent votes for the Republicans and p_D the probability that she votes for the Democrats. We denote the difference of these probabilities by $\omega = p_R - p_D$. Thus, $\omega \in \Omega = [-1, 1]$. With an appeal to a law of large numbers for large economies, we interpret p_R and p_D also as the fraction of independents voting, respectively, for Republicans and Democrats. Consequently, ω is the Republican's margin of victory in the pool of independent voters. In the following, ω is taken to be the realization of a random variable with *cdf* denoted by F. We assume F to be continuous.

Popular vote. We denote the set of states in which party D or party R wins the popular vote, respectively, by

$$\Omega_D := \left\{\omega : \omega < \frac{\beta_D - \beta_R}{\beta_I}\right\} \quad \text{and} \quad \Omega_R := \left\{\omega : \omega > \frac{\beta_D - \beta_R}{\beta_I}\right\}\ .$$

The probability that party D wins the popular vote is denoted by

$$\pi_D^* = \mathrm{pr}(\omega \in \Omega^D) = F\left(\frac{\beta_D - \beta_R}{\beta_I}\right)\ .$$

District outcomes. As we describe in more detail below, we consider games so that voters are allocated to districts over various rounds. In such a process, any one party $P \in \{D, R\}$ sends voters to any one district k. There is then a particular mix of Republican partisans, Democratic partisans and independent voters in that mass of voters. More formally, a strategy for party D is a collection $\sigma_D = (\sigma_{Dk})_{k=1}^{2N}$, and a strategy for party R is a collection $\sigma_R = (\sigma_{Rk})_{k=1}^{2N}$. In this collection,

$$\sigma_{Dk} = (\sigma_{Dk}^D, \sigma_{DK}^R, \sigma_{Dk}^I) \quad \text{with} \quad \sigma_{Dk}^D + \sigma_{Dk}^R + \sigma_{Dk}^I = 1\ ,$$

is the proposal of party D for district k. Thus, party D assigns a unit mass of voters to any one district k, and the shares of D partisans, R partisans and independent voters are, respectively, denoted by σ^D_{Dk}, σ^R_{Dk} and σ^I_{Dk}. We use analogous notation for party R. Party D wins district k if

$$\sigma^D_{Dk} + \sigma^D_{Rk} \geq \sigma^R_{Dk} + \sigma^R_{Rk} + \omega \left(\sigma^I_{Dk} + \sigma^I_{Rk}\right) \ , \tag{1}$$

or, equivalently, if

$$\frac{\sigma^D_{Dk} + \sigma^D_{Rk} - \left(\sigma^R_{Dk} + \sigma^R_{Rk}\right)}{\sigma^I_{Dk} + \sigma^I_{Rk}} \geq \omega \ . \tag{2}$$

Hence, if F is the *cdf* of ω, the probability that party D wins district k is

$$\pi_{Dk} = F\left(\frac{\sigma^D_{Dk} + \sigma^D_{Rk} - \left(\sigma^R_{Dk} + \sigma^R_{Rk}\right)}{\sigma^I_{Dk} + \sigma^I_{Rk}}\right) \ .$$

and the probability that party R wins district k is $\pi_{Rk} = 1 - \pi_{Dk}$. In the following, when we seek to emphasize the dependence of winning probabilities on the parties' strategies, we write $\pi_{Dk}(\sigma_D, \sigma_R)$.

We say that district k is a replica of the at-large district if $\pi_{Dk} = \pi^*_D$.[12] If all districts are replicas of the at-large-district, then the popular vote determines outcomes both at the local district level and at the aggregate state or national level. This avoids constellations so that one party wins the popular vote and the other party wins a majority of seats. It also creates healthy competition at the district-level so that no representative can take his reelection for granted. In the following, Theorem 2 will show that, for N large we can come very close to the ideal of all districts replicating the at-large-district.[13]

Winning the election. A party wins the election if it wins a majority of the seats. Recall that there $2N$ districts and an at-large-district. Thus, there are $2N + 1$ seats in total and winning a majority requires to win at least $N + 1$ of them. We denote by π^V_D,

[12]Note that "replica" here does not require that all three type proportions are the same in district k as in the electorate as a whole (though that would certainly be sufficient), but rather that the critical state

$$\hat{\omega}_k = \frac{\sigma^D_{Dk} + \sigma^D_{Rk} - \left(\sigma^R_{Dk} + \sigma^R_{Rk}\right)}{\sigma^I_{Dk} + \sigma^I_{Rk}}$$

is the same in district k and in the electorate as a whole. For example, if the type distribution in the electorate is $(0.4, 0.4, 0.2)$ (for D-partisans, R-partisans, independents), then the critical state is the same in a district where the type distribution is $(0.3, 0.3, 0.4)$.

[13]Katz et al. (2020) define various measures of fairness in redistricting, using the seat-vote curve $S(V)$ that maps the Democrats' overall vote share V into the proportion of seats that they win in the legislature. For example, symmetry (i.e., $S(V) = 1 - S(1 - V)$) requires that the Democrats' and the Republicans' seat share as a function of their vote share is the same.

In their language, the objective that the popular vote winner receives a majoritycan be written as $S(V) < 0.5$ if $V < 0.5$ and $S(V) > 0.5$ if $V > 0.5$.

the probability of such a Victory for party D. We define π_R^V analogously. Our focus is on whether we can ensure an alignment of the party that wins a majority of seats with the party that wins the popular vote.

For a formal treatment of this question the following notation proves helpful. Given a pair of strategies (σ_D, σ_R), we denote the probability of a victory for party R, conditional on party R winning the popular vote, by $\pi_R^V(\sigma_D, \sigma_R \mid \omega \in \Omega_R)$. A system that guarantees the "correct" outcome at the aggregate level has

$$\pi_R^V(\sigma_D, \sigma_R \mid \omega \in \Omega_R) = 1 \quad \text{and} \quad \pi_D^V(\sigma_D, \sigma_R \mid \omega \in \Omega_D) = 1 \ ,$$

for every pair of equilibrium strategies (σ_D, σ_R).

Our main result in Theorem 1 will show that we can indeed achieve this outcome through a sequential mechanism in which parties assign voters to districts over many rounds, alternating which party moves first and which one moves second. We now turn to formal description of this sequential protocol.

4.1 The protocol

Each party assigns every voter to one of the districts. As a consequence, 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 election.

Sequence of moves. Voters are assigned to districts over L rounds. In each round l, any party P specifies $\sigma_{Pl} = (\sigma_{kPl}^D, \sigma_{kPl}^R, \sigma_{kPl}^I)_{k=1}^{2N}$ so that

$$\sigma_{kPl}^D + \sigma_{kPl}^R + \sigma_{kPl}^I = \frac{1}{L} \ .$$

In words: Party P assigns a mass of $\frac{1}{L}$ voters to any one district k. The percentage shares of D partisans, R partisans and independents in that mass of voters are then, respectively, given by

$$\beta_{kPl}^D \ := \ L\,\sigma_{kPl}^D \ , \quad \beta_{kPl}^R \ := \ L\,\sigma_{kPl}^R \quad \text{and} \quad \beta_{kPl}^I \ := \ L\,\sigma_{kPl}^I \ .$$

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. D has this advantage in odd rounds and R has it in even rounds.

Feasibility. Let the total number of D partisans assigned by party P to district k over L rounds be denoted

$$\sigma_{Pk}^D := \sum_{l=1}^{L} \sigma_{Pkl}^D.$$

Analogously, let

$$\sigma_{Pk}^{R} := \sum_{l=1}^{L} \sigma_{Pkl}^{R} \quad \text{and} \quad \sigma_{Pk}^{I} := \sum_{l=1}^{L} \sigma_{Pkl}^{I} \ .$$

For any party P, $(\sigma_{Pk})_{k=1}^{2N}$ has to be consistent with the distribution of voters in the electorate at large, i.e.,

$$\frac{1}{2N} \sum_{k=1}^{2N} \sigma_{Pk}^{D} = \beta_D \ , \quad \frac{1}{2N} \sum_{k=1}^{2N} \sigma_{Pk}^{R} = \beta_R \ , \quad \text{and} \quad \frac{1}{2N} \sum_{k=1}^{2N} \sigma_{Pk}^{I} = \beta_I \ .$$

Winning probabilities. Winning probabilities for specific districts or for a majority of seats depend on the number of rounds L. We use superscript L to indicate this dependence. For instance, we write π_{Dk}^{L} for the probability that party D wins district k when there are L rounds of play, or π_{R}^{VL} for the probability that party R wins a majority of seats when there are L rounds of play.

4.2 The main result

Our main result, Theorem 1, shows that, with a sufficiently large number of rounds, every equilibrium is such that the "correct" party wins, namely the one that wins the popular vote.

Theorem 1 *For all $\varepsilon > 0$, there is $\hat{L}$, so that, for $L \geq \hat{L}$, in every equilibrium (σ_D, σ_R),*

$$\pi_{R}^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_R) \geq 1 - \varepsilon \quad \textit{and} \quad \pi_{D}^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_D) \geq 1 - \varepsilon \ .$$

Theorem 2 complements this finding: It shows that it is possible to achieve this outcome with only small distortions at the district level. With many districts, i.e. for $N \to \infty$, there is an equilibrium, so that almost every district is a replica of the at-large-district.

Theorem 2 *For all $\varepsilon > 0$ and all $\delta > 0$, there is $\hat{L}$ so that for $L \geq \hat{L}$, there exists a pair of strategies (σ_D, σ_R), so that*

$$\pi_{R}^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_R) \geq 1 - \varepsilon \quad \textit{and} \quad \pi_{D}^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_D) \geq 1 - \varepsilon \ ,$$

and

$$\# \left\{ k : \ \left| \frac{\sigma_{Dk}^{D} + \sigma_{Rk}^{D} - (\sigma_{Dk}^{R} + \sigma_{Rk}^{R})}{\sigma_{Dk}^{I} + \sigma_{Rk}^{I}} - \frac{\beta_D - \beta_R}{\beta_I} \right| \geq \delta \right\} \frac{1}{2N} \quad \leq \quad \frac{2}{N} \ .$$

Formal proofs of Theorems 1 and 2 are in the Appendix. Note that Theorems 1 is a statement about *all* equilibrium strategies, whereas Theorem 2 is a statement about *one* pair of strategies that approximates a particular equilibrium for L large.

The key insights that carry these proofs are Propositions 1, 2 and 3 below. Proposition 1 deals with the symmetric case of parties that have equal numbers of partisan supporters, $\beta_D = \beta_R$. As we will explain, in this case, each party can neutralize any attempt of its competitor to create districts that are more favorable than the average district. Propositions 2 and 3 then deal with the asymmetric case, $\beta_D > \beta_R$. Proposition 2 shows that party D can ensure to win whenever $\omega \in \Omega_D$ simply by spreading its partisans supporters evenly over at least fifty percent of the districts, i.e. by following a strategy that is referred to as *cracking* in the literature. Proposition 3 deals with the challenge for party R to counter this strategy of party D in such a way that it wins a majority whenever $\omega \in \Omega^R$. We will show that party R can achieve this outcome by a strategy of spreading D partisans over less than fifty per cent of the districts. Such a strategy is referred to as *packing*.

4.3 Symmetry

The following Proposition 1 deals with the case that both parties have the same number of committed supporters. In this case, both players can achieve (in the limit of $L \to \infty$) that each district looks like the electorate at large and therefore is won by D if $\omega \in \Omega_D$ and by R if $\omega \in \Omega_R$.

Proposition 1 *Suppose that $\beta_D = \beta_R$. For every $\varepsilon > 0$, there is $\hat{L}$ so that for $L \geq \hat{L}$,*

a) there is σ_R so that, for every σ_D, $\left| \frac{\sigma^D_{Dk}+\sigma^D_{Rk}-(\sigma^R_{Dk}+\sigma^R_{Rk})}{\sigma^I_{Dk}+\sigma^I_{Rk}} - \frac{\beta_D-\beta_R}{\beta_I} \right| \leq \varepsilon$ in every district k.

b) there is σ_D so that, for every σ_R, $\left| \frac{\sigma^D_{Dk}+\sigma^D_{Rk}-(\sigma^R_{Dk}+\sigma^R_{Rk})}{\sigma^I_{Dk}+\sigma^I_{Rk}} - \frac{\beta_D-\beta_R}{\beta_I} \right| \leq \varepsilon$ in every district k.

A proof of the Proposition can be found in the Appendix. The main idea is to specify a strategy, say, for party R that neutralizes any attempt of party D to engage in cracking or packing. To achieve this, in any round l, and for any district k, R assigns a number of R partisans that matches the number of D partisans assigned by party D in the previous round. Likewise, R assigns a number of D partisans equal to the number of R partisans assigned previously by party D. Consequently, in any district k, $\sigma^D_{Dk}+\sigma^D_{Rk}-\left(\sigma^R_{Dk}+\sigma^R_{Rk}\right)$ is zero, and hence party D wins the district with probability

$$F\left(\frac{\sigma^D_{Dk}+\sigma^D_{Rk}-\left(\sigma^R_{Dk}+\sigma^R_{Rk}\right)}{\sigma^I_{Dk}+\sigma^I_{Rk}}\right) = F(0) = F\left(\frac{\beta^D-\beta^R}{\beta^I}\right) = \pi^*_D .$$

The formal proof consists in showing that, for L large, such moves – or close approximations of them – are feasible. If party R follows such a strategy, then, for L sufficiently large, party R can turn every district into a copy of the at-large-district, and therefore win them whenever it wins the majority of the popular vote. It is worth emphasizing that this neutralizing strategy of party R works robustly, i.e. it does not depend on the behavior of party D.

In turn, party D also has such a neutralizing strategy at its disposal, i.e., can also turn each district into an approximate copy of the at-large district, and therefore also can ensure that it wins the districts whenever it wins the at-large district.

Proposition 1 implies, in particular, that for sufficiently large L, party D has a strategy so that it wins a majority of districts whenever $\omega \in \Omega_D$. (In fact, it wins all districts in these states.) By the same logic, party R has a strategy so that it wins a majority of districts whenever $\omega \in \Omega_R$. Therefore, in every equilibrium, it has to be the case that party D wins whenever $\omega \in \Omega_D$ and that party R wins whenever $\omega \in \Omega_R$: If party D won with a probability strictly larger than π_D^*, party R could profitably deviate to the strategy in part a) of Proposition 1. If party D won with a probability strictly below π_D^*, party D could profitably deviate to the strategy in part b) of Proposition 1.[14]

An illustration. Proposition 1 uses the properties of neutralizing strategies to bound equilibrium payoffs, but does not provide an explicit characterization of equilibrium strategies. In the limit case, as $L \to \infty$, the neutralizing strategy yields equilibrium payoffs and can therefore be regarded as an "equilibrium in the limit". For finite L, neutralizing moves and best responses do not usually coincide. We illustrate this by means of the following example: Suppose that $\beta_D = \beta_R = \frac{1}{3}$ and that there are 10 districts, so that a legislative majority requires winning (at least) six districts.

First observe that, if party R was in charge of assigning voters to districts, without being contested by party D, it would spreads its own partisan supporters evenly over six districts and pack the Democratic partisans in a complementary set of four districts, as illustrated in Figure 1.

Figure 2 below illustrates the neutralizing response of party D, and Figure 3 illustrates a best response, under the assumption that $L = 1$, so that the game ends after party D's move. The neutralizing strategy has every D partisan matched with an R partisan and every R partisan matched with a D partisan. Under symmetry, there is also a D partisan for every R partisan in the electorate at large. As a consequence, the parties' winning probabilities at the district level are equal to their probabilities of winning the popular vote.

The best response, in the subgame generated by R's move depicted in Figure 1, differs from the neutralizing response in that it takes advantage of the fact that D partisans have already been packed in four districts. Party D then tries to secure two further districts and, once they have each been assigned a unit mass of D partisans (the maximum possible amount), uses the remaining D partisans to increase its winning probability in the four districts that already lean towards it. Party D finally disposes of the R partisans in its budget set so that they do not interfere with the winning probability in those six D

[14]Strictly speaking, the preceding argument shows that, in any equilibrium, D's winning probability must be $\Pr(\Omega_D)$. The conclusion that D must win in *all* states $\omega \in \Omega_D$, and lose whenever $\omega \in \Omega_R$ follows from the fact that the votes for either party, are, in each district, monotone in ω.

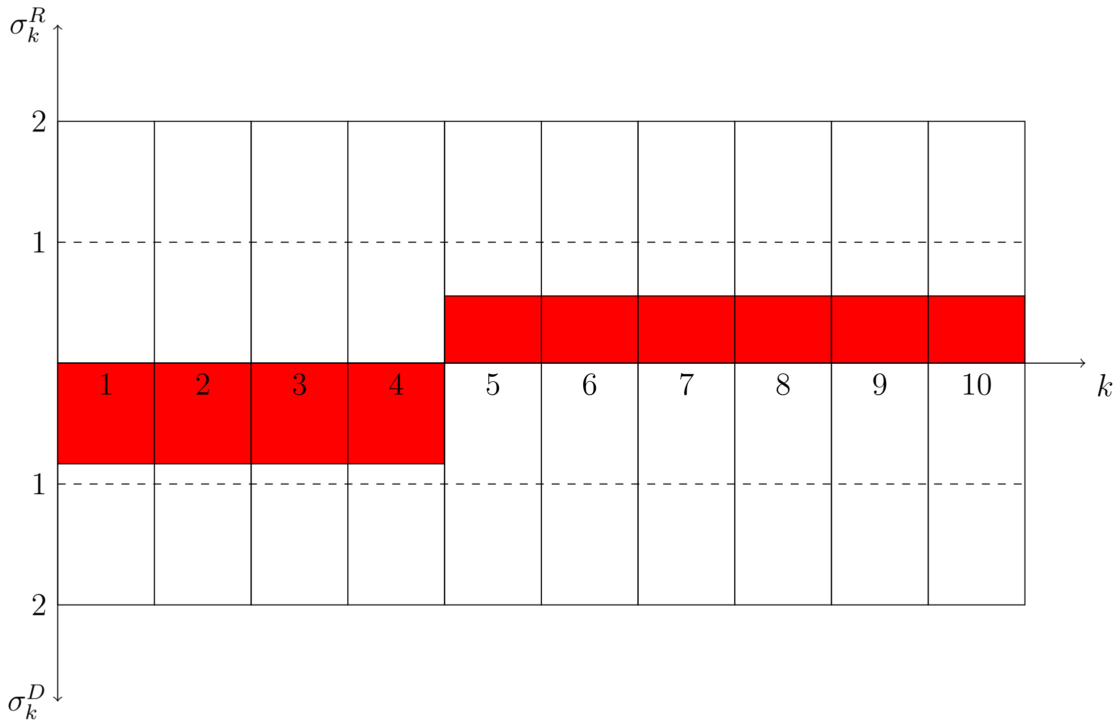

Figure 1: 10 Districts, $\beta_D = \beta_R = \frac{1}{3}$. Party R's optimal assignment of a unit mass of voters when it seeks to maximize the probability of winning a majority of 6 districts – and when party R's assignment is not contested by an assignment of party D. Party R then wins a majority of districts with probability $F(1.25) > F(0)$, its probability of winning the popular vote.

leaning districts; i.e. R partisans are assigned where party D's chances of winning are lowest. Consequently, party R wins a majority of six seats only with probability $F(-1)$.

4.4 Asymmetry

Proposition 1 cannot be adapted to the asymmetric case $\beta_R < \beta_D$. For concreteness, suppose that $\beta_D > 0$ and $\beta_R = 0$: party D has a significant number of loyal voters, whereas party R gets votes only from independent voters, if at all. Neutralizing strategies are not feasible in this case. For instance, Party D can no longer match the number of democratic partisans assigned to some district k by party R with an equal number of republican partisans.

In the following, we will first focus on the case $\beta_D > 0$ and $\beta_R = 0$, and subsequently extend the analysis to constellations with $0 < \beta_R < \beta_D$. We show that either party can ensure to win a majority of districts with a probability at least as large as its probability of winning the popular vote. In this sense, at the aggregate level, the "correct" outcome prevails. However, we can no longer ensure that all districts are turned into close replicas of the electorate at large. There may be some safe districts for party D. That said, for N and L large, we will show that there is an equilibrium in which the share of such districts

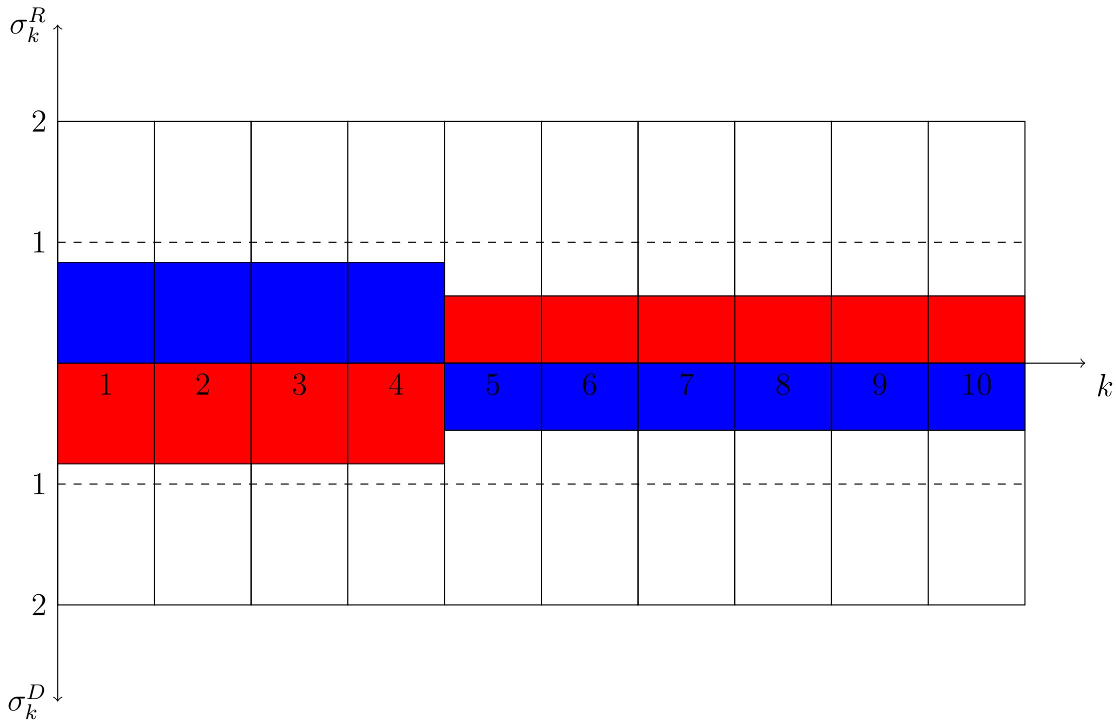

Figure 2: 10 Districts, $\beta_D = \beta_R = \frac{1}{3}$. Party R's neutralizing response to party D's assignment in Figure 1.

is small.

Recall that, due to the at-large-district, if party D wins N districts and the popular vote it wins a strict majority of seats. The following Proposition shows that party D can ensure to win N districts in all states of the world ω in which it wins the popular vote. Put differently, party D can ensure to win the overall election in all states of the world where it "should" win.

Proposition 2 *Suppose that $\beta_R = 0$ and $\beta_D > 0$. For every L, there is σ_D so that, for every σ_R, $\pi_D^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_D) = 1$.*

The Proposition follows from a simple argument. Suppose that party D assigns, over the course of the whole game, a mass of $2\beta_D$ partisan D voters to half of the districts, say, to any district with an index k larger or equal to $N + 1$.[15] Then, whatever, the strategy of party R, the percentage share of partisan D voters in those district is bounded from below by β_D. Equivalently, all such districts are won whenever the state ω is such that $\beta_D \geq \omega\, \beta_I$. Hence, $\omega \in \Omega^D$ implies that party D wins at least fifty percent of all districts. The at-large district then ensures a majority of seats for party D. Also note that this conclusion does not depend on the assumption that the number of rounds L is large.

For party R it is more challenging to ensure a victory whenever it wins the popular vote. When party R was alone in determining districts, it could simply pack all partisan

[15] In the literature such a strategy is also refereed to as "cracking".

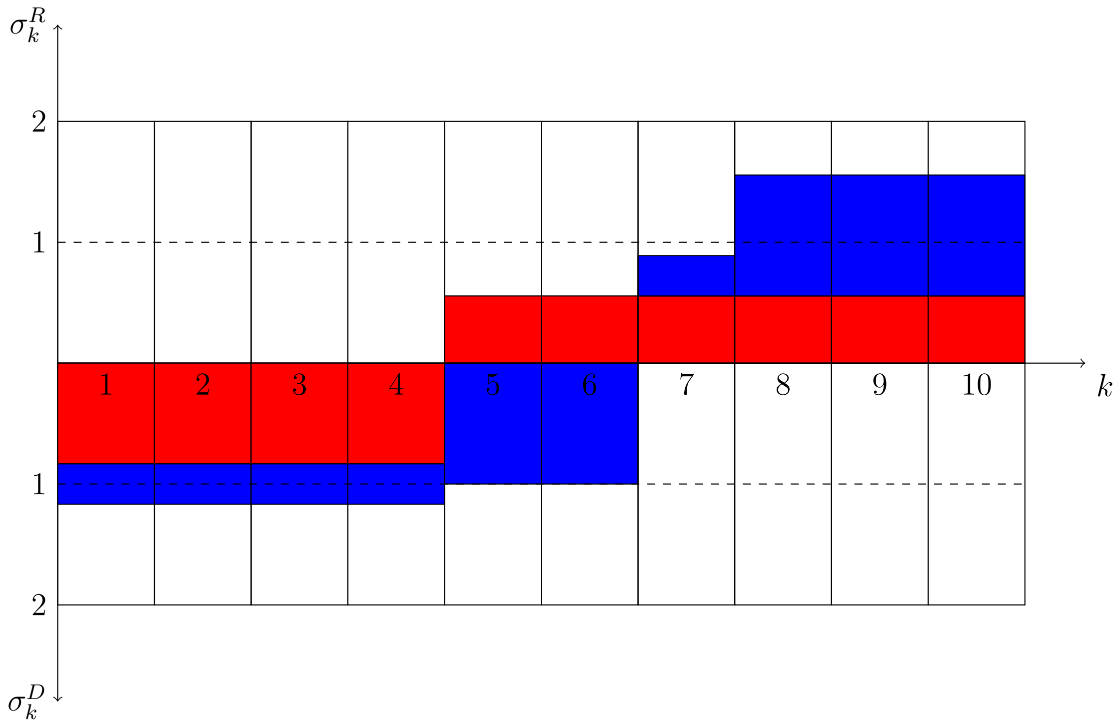

Figure 3: 10 Districts, $\beta_D = \beta_R = \frac{1}{3}$. Party D's best response to party R's assignment in Figure 1.

D voters in a small subset of districts, and, as a consequence, have a large number of districts with only independent voters. It would then win all of the latter whenever independents voters lean towards party R, i.e. whenever $\omega > 0$. Thus, party R would win too often, i.e. it would win in states ω in which it does not win the popular vote. The presence of party D upsets this strategic logic. Party D also has an incentive to concentrate partisan D voters, but would spread them over a larger number of districts. By Proposition 2, when they are spread over N districts, party D wins a majority of seats whenever it wins the popular vote. Thus, if party R engaged in a packing of partisan D voters in a small number of districts, party D would simply say "thank you" and add further districts that are dominated by D partisans to those that have been generated by party R.

The following Proposition shows that, for L large, party R can overcome these difficulties. More precisely, it can ensure to win a majority of districts whenever it wins the popular vote.

Proposition 3 *Suppose that $\beta_R = 0$ and $\beta_D > 0$. For every $\varepsilon > 0$, there is $\hat{L}$ so that $L \geq \hat{L}$ implies the existence of a strategy σ_R so that, for all σ_D,*

$$\pi_R^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_R) \geq 1 - \varepsilon \ .$$

A detailed proof can be found in the Appendix. Its logic is as follows: When party R wants to secure a majority whenever $\omega \in \Omega_R$, it needs to ensure that there are at least N

districts so that, after L rounds of play, the percentage share of partisan D voters is below β^D. Thus, the objective of party R is to minimize the share of D partisans in the district with rank N, where the ranking is such that lower ranked districts have a lower share of D partisans. To show that party R can indeed enforce an outcome where this share of D partisans remains below β_D, we assume that party D's objective is to maximize the share of D partisans in the district with rank N. When R's strategy succeeds under this assumption, then it succeeds against any strategy of party D.[16]

Since party D seeks to maximize the share of D partisans in that district, it will not waste partisan D voters in lower ranked ones. Thus, party D concentrates partisan D voters in the $N+1$ top-ranked districts, i.e. it engages in *cracking*. More specifically, whenever it is called upon to play in some round l, and plans to assign a certain mass of D partisans, the following pecking order is optimal: Assign D voters to the district with rank N until its mass of D partisans 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 D partisans 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 D partisans are left, see Figures 4 and 5 for an illustration.

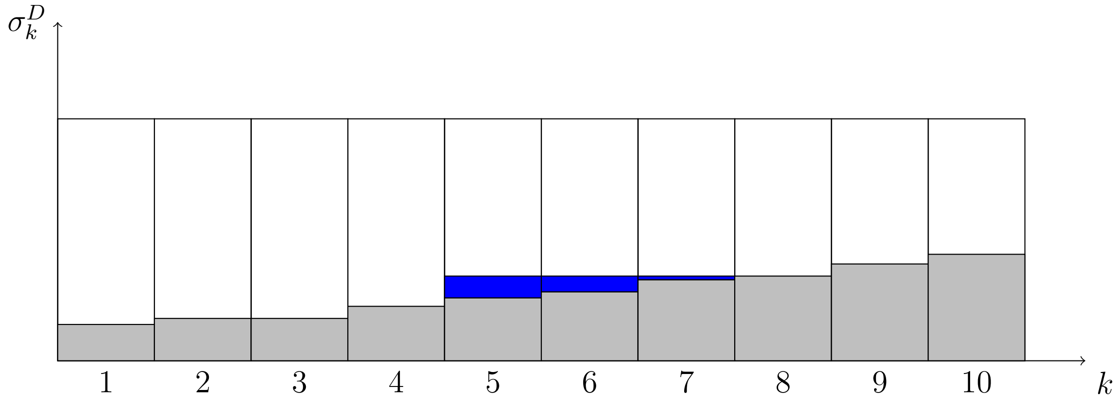

Figure 4: 10 Districts, $0 = \beta_R < \beta_D$. In round l, party D inherits, for every district, a stock of D partisans, illustrated in gray. It then adds further D partisans in round l, illustrated in blue. This figure is drawn under the assumption that party D assigns only few partisan D voters in round l, so that, when assigning them optimally, its budget allows to raise the level of partisan D voters only in districts 5,6, and 7.

What is an optimal response of party R? Its problem is to get rid of a total mass of $2N\beta^D$ partisan D voters in such a way that they contribute as little as possible to the mass of partisan D voters in the district with rank N. What is clearly harmless is to add partisan D voters to districts with ranks below $N-1$, provided they are not yet an

[16]We show in the Appendix, that we can assume that the ranking of districts remains unchanged over the whole course of the game. Any outcome that a party can achieve in some round l with a rank reversing voter assignment can also be achieved without a rank reversal. Thus, there is no ambiguity when we simply refer to the district with rank N.

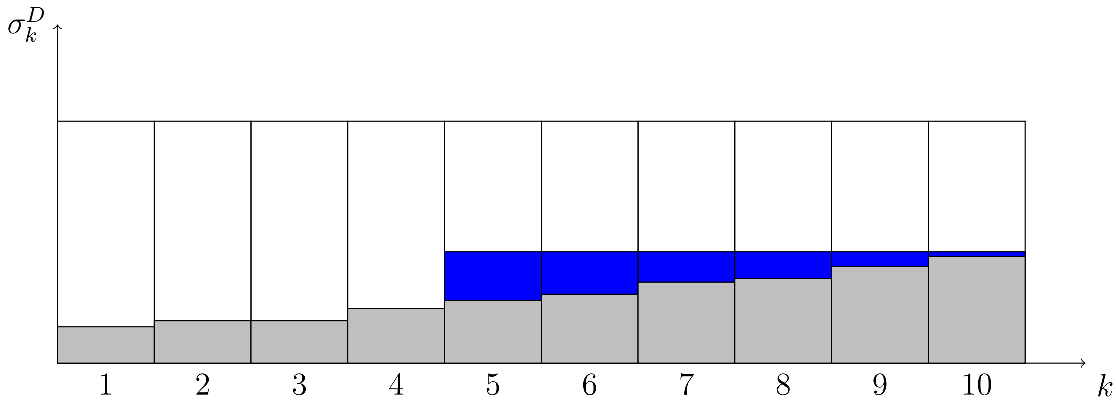

Figure 5: 10 Districts, $0 = \beta_R < \beta_D$. In round l, party D inherits, for every district, a stock of D partisans, illustrated in gray. It then adds further D partisans in round l, illustrated in blue. This figure is drawn under the assumption that party D assigns many partisan D voters in round l, so that, when assigning them optimally, its budget allows to raise the level of partisan D voters in all districts with a rank weakly larger than 5.

equal level with the district that has rank N.[17] Thus, when party R plans to assign some mass of partisan D voters in some round, it will first fill the bottom $N-1$ districts up to the point where a joint level of partisan D voters is reached in the bottom N districts. This ensures a minimal level of partisan D voters in all districts. See Figure 6 for an illustration under the assumption that the mass of partisan D voters assigned in round l does not suffice to bring the bottom 4 districts to the level of district 5. Figure 7 is based on the alternative assumption that the mass exceeds what would be needed for that purpose.

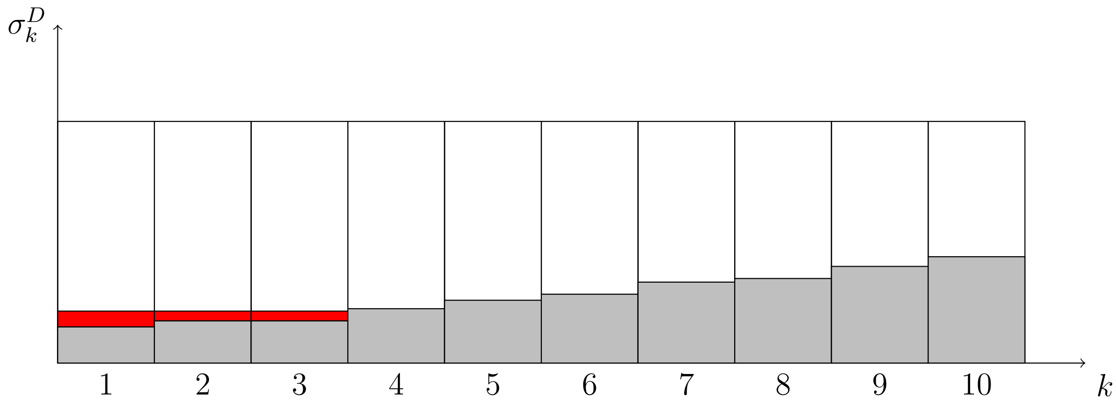

Figure 6: 10 Districts, $0 = \beta_R < \beta_D$. In round l, party R inherits, for every district, a stock of D partisans, illustrated in gray. It then adds further D partisans in round l, illustrated in red. This figure is drawn under the assumption that party R assigns few partisan D voters in round 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.

[17]In the literature this is refereed to as “packing”.

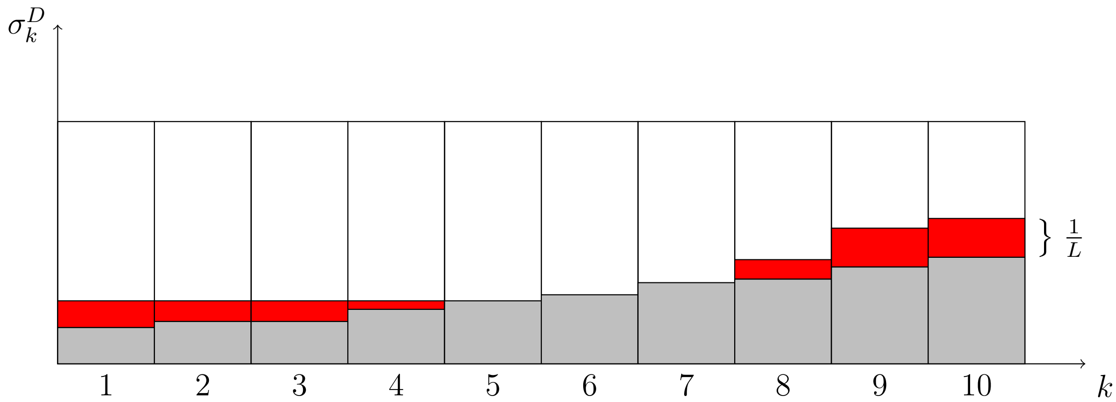

Figure 7: 10 Districts, $0 = \beta_R < \beta_D$. In round l, party R inherits, for every district, a stock of D partisans, illustrated in gray. It then adds further D partisans in round l, illustrated in red. This figure is drawn under the assumption that party R assigns many partisan D voters in round 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 D partisans are then assigned to the top-ranked districts.

Figure 7 illustrates the following logic: When further partisan D voters need to be assigned after a common level in the bottom N districts has been achieved, party R needs to continue with districts in the upper half. Here the logic is to assign partisan D voters where they are least likely to help party D in its attempts to raise the level in the district with rank N. Thus, party R begins with the top ranked district and there assigns as many partisan D voters as possible. If the capacity constraint of $\frac{1}{L}$ for that district and that round is reached, party R will start to fill the district with the second highest rank, and so on. Thus, there is a concentration on the top-ranked districts when party R assigns D partisans to districts in the upper half of the rank distribution.

Using a more colorful language, we also refer to R's strategy for the bottom half as a *water-level-strategy*. Given a sufficient volume of partisan D voters, the level in the basin consisting of the bottom $N-1$ districts is raised to the level prevailing in the district with rank N. We will refer to R's strategy for the upper half as *building-towers-strategy*. A tower is a district in the upper half with a level of partisan D voters that sticks out. If additional D partisans need to be assigned, party R will assign them with priority to the district that sticks out most, i.e. it will make the highest tower even higher. It will then move to the second highest tower, and so on.

For a complete equilibrium characterization, we would also need to describe how many D partisans 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_{Pl}^D, defined as the percentage share of D partisans in the total mass of $\frac{2N}{L}$ voters assigned by party P in round l. We do not provide such a complete equilibrium characterization, but show that party R can choose the sequence $\{\beta_{Pl}^D\}_{l=1}^L$ so that the share of partisan D voters in the district with rank N

remains below β_D. To this end, assume that party R chooses $\beta^D_{R1} = 0$, and for any $l \geq 2$, $\beta^D_{Rl} = \beta^D_{Dl-1}$. Thus, party R waits until party D starts to assign D partisans and then assigns in, any round, as many D partisans as party D assigned in the round before.

Given the partial characterization of equilibrium behavior above, this implies that, after any move of party R, the bottom $2N-2$ districts have the same level of partisan D voters, while there are some further partisan D voters in the top ranked district, and, possibly, also in the district with the second highest rank. To see this, suppose for concreteness, that party D chooses $\beta^D_{D1} > 0$. Then, it will spread a mass of $\beta^D_{D1}\ \frac{2N}{L}$ partisan D voters over evenly over $N+1$ districts. In round 2, party R will use the mass of voters previously assigned to $N-1$ of those districts to have an equal water-level in the bottom half. The remaining mass of partisan D voters is then assigned to at most two further districts. See Figure 8 for an illustration. This pattern is now repeated over various rounds, with the implication that, after any move of party R there is a joint level of partisan D voters in the bottom $2N-2$ districts.

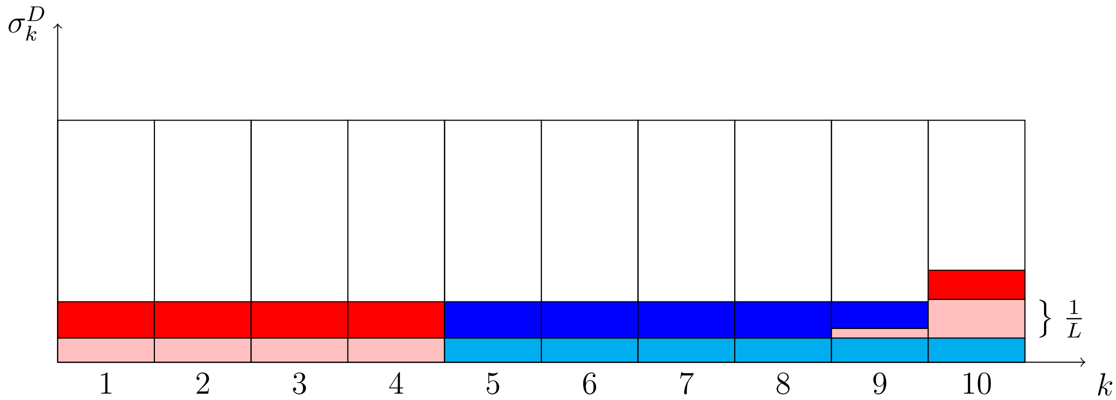

Figure 8: 10 Districts, $0 = \beta_R < \beta_D$. Party R assigns as many D partisans as party D did in the previous round. In light blue is the first round in which party D assigns a positive mass of partisan D voters. The response of party R is in light red. In blue is the second round in which party D assigns a positive mass of partisan D voters, and party R's response is in red. As a consequence, there is a common water-level in the bottom eight districts, both after R's first response and after R's second response.

It is now easy to see that the percentage share of partisan D voters in the pivotal district with rank N cannot be strictly above β_D. This would imply a percentage share above β_D in all districts and this is incompatible with the fact that the share if partisan D voters in the electorate at large equals β_D. Also note that there is a common level of partisan D voters in all districts, with exception of the two top ranked ones. Thus, party R's has a strategy that ensures winning a majority whenever $\omega \in \Omega_R$, and moreover, implies that there are at most two districts that are "safe" for party D. For $N \to \infty$ the fraction of districts where the outcome deviates from the popular vote becomes negligible.

By Proposition 3, and for L large, party R can ensure to win a majority of seats whenever $\omega \in \Omega_R$ and, by Proposition 2 party D can ensure to win a majority of seats whenever $\omega \in \Omega_D$. As an implication, for N and L large, a constellation where party D concentrates its partisan D supporters in fifty percent of the districts and party R concentrates them in the other half approximates the equilibrium. Also the share of districts in which outcomes are not approximately equal to the popular vote then becomes negligible.

Relaxing the assumption that $\beta^R = 0$. Proposition 4 below is a generalization of Proposition 3 that allows for the possibility that there are both R partisans and D partisans, but maintains the assumption that there are more of the latter.

Proposition 4 *Suppose that $\beta_R < \beta_D$.*

a) For every $\varepsilon > 0$, there is $\hat{L}$ so that $L \geq \hat{L}$ implies the existence of a strategy σ_R so that, for all σ_D,

$$\pi_R^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_R) \geq 1 - \varepsilon \ .$$

b) For every $\varepsilon > 0$, there is $\hat{L}$ so that $L \geq \hat{L}$ implies the existence of a strategy σ_D so that, for all σ_R,

$$\pi_D^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_D) \geq 1 - \varepsilon \ .$$

The key for the proof of Proposition 4 is the insight is that either party has an incentive to use its own partisan supporters so that they are spread evenly over fifty percent of the districts. The other party then has an incentive to respond to this attempt using a water-level and building-towers-strategy. When this logic is squared with the assumption that either party assigns in a round l as many rival partisans as the rival party used in the previous round – so that $\beta_{Rl}^D = \beta_{Dl-1}^D$ and $\beta_{Dl-1}^R = \beta_{Rl-2}^R$ – then there at most 2 safe districts for party D and at most two safe districts for party R. Theorem 2 follows from this last observation.

5 Discussion

In this section, we discuss issues of robustness, both in terms of substantive concerns (such as additional requirements to impose on the system), as well as theoretical extensions of the model.

Our system achieves two objectives, namely that the party that wins the support of a majority of voters wins the majority in the legislature, and that all, or at least a large share of, districts are competitive. This said, there may be other objectives that are desirable for a district map and the implied legislature, and so it is interesting to analyze whether it is possible to tweak our system so that it also satisfies these additional objectives.

5.1 Opposition representation in the legislature

We first discuss the objective of having the opposition party represented in the legislature. Because most or all districts are replicas of the electorate at-large, in most states of the world, the majority-preferred party wins a very large percentage of seats, with few or none going to the minority party.

Even though the minority party has very 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 opposition representation in the legislature? One simple possibility s to turn each 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.[18]

In this case, the redistricting game between the parties remains exactly the same as in the basic model, while 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.

[18]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.

5.2 Representation of demographic minorities

Another conceivable objective is that there is a certain subset of districts whose majority has to be composed of a certain demographic type such as African Americans or Hispanics ("majority-minority districts"). Generally, there is a tension between imposing this constraint and an implementation of the popular vote: If demographic minorities are extremely likely to vote for Democrats, then generating a subset of districts in which minority voters are a majority necessarily implies that the remaining districts have a below-average share of Democratic partisans. Thus the objectives of ensuring a fair election outcome in terms of a correspondence between the outcome of the popular vote, and creating a large set of "majority-minority" districts, may be logically incompatible.

This said, with the system of competitive gerrymandering described in the previous section, if a party wants to generate districts that overrepresent certain demographic groups, it can plausibly do so that without negatively impacting its winning probability. For example, suppose that the Democratic party has its core supporters among Blacks and certain urban Whites, while they are weaker among other groups of voters, e.g., rural Whites. How the Democrats mix these voter types into legislative districts is their choice – in particular, it seems well feasible to create some districts in which the Democratic partisans allocated to these districts are predominantly Black, so that they would have a strong influence on the outcome of the Democratic primary.

5.3 Geographic constraints

In most states, electoral districts are currently required to be contiguous. Our system allows parties to allocate voters to districts without any geographic constraints, and the districts generated as an equilibrium outcome are unlikely to be contiguous. Is that problematic? We now argue that the answer is "no".

There are three justifications for imposing a contiguity requirement for districts, neither of which is particularly compelling from a logical point of view.

First, most (though not all) polities throughout history have contiguous maps, which is useful for the provision of public goods, say, the protection against outside invaders. As public goods are not really provided at the legislative district level, and there is, in particular, no danger of invasion that concerns single districts and could be easier countered if districts were contiguous, this is not a convincing argument.

Second, and relatedly, geographic closeness may be a proxy for preference homogeneity, and it could be argued that it may be more efficient, or at least easier, for representatives to represent more homogeneous districts. While this argument clearly applies to cities and other polities that determine autonomously which public goods are and are not provided for their inhabitants, this argument is less compelling for legislative districts. State legislators generally set policy as a team, and it applies to the state as a whole. From this perspective, it is not clear what advantage more homogeneous legislative districts would convey.

Third, it could be argued, that contiguity makes it easier for representatives to provide constituency services. Note that what plausibly influences the costs of providing constituency services is the average geographic distance of constituents, not contiguity of the district by itself. For example, Texas' 28th district (in Figure 9) is contiguous, but clearly not designed to minimize travel time to the representative's district office.

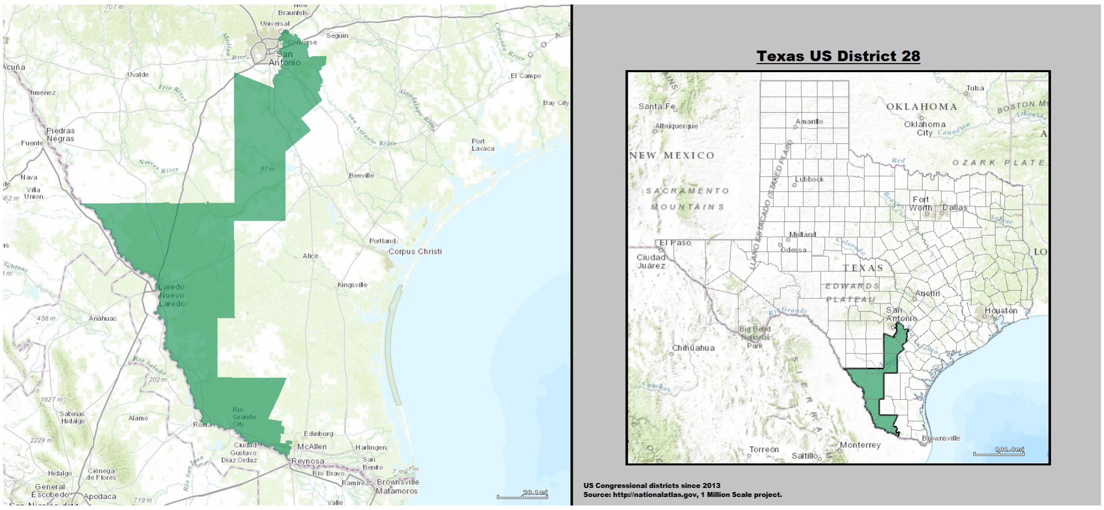

Figure 9: TX-28: Contiguous, but not compact

To the extent that minimizing travel is important to lower the costs for representatives' constituency service, both parties should prefer geographically close districts, and we would therefore expect both parties in our system to coordinate their choices in a way that voters in any one district are geographically clustered (albeit probably not in a contiguous way), rather than coming from all corners of the state.[19]

The fourth and final argument for a contiguity requirement is that, in a time of severe gerrymandering, contiguity serves as a de-facto constraint that limits the party in power from even more egregious gerrymandering. While this may be a valid second-best type argument within the current redistricting system, our system shows that it is possible to reach a better outcome without this constraint.

5.4 National elections and statewise gerrymandering

Our model directly applies to the choice of state legislative districts, i.e., the choice of the district map is made at the level of the polity that is voting in later elections.

[19] Formally, suppose that parties are lexicographically-first interested in their winning probability, and lexicographically-second in the average distance of voters within districts (or some other measure of geographic closeness). Our results imply that, in any equilibrium, the parties' winning probabilities are equal to their respective probability of winning the popular vote. Both parties have the same ranking over equilibrium maps in terms of their geographic closeness, and therefore can be expected to coordinate on the mutually best equilibrium.

For federal elections, the power to draw district maps lies at the state level, while the relevant outcome is clearly which party wins a majority in the House of Representatives. This mismatch (which one can interpret as another geographic constraint) would be problematic in our system.

Specifically, suppose there is a set S of states, with the electorate of state s captured by the population shares $\beta_s = (\beta_{Ds}, \beta_{Rs}, \beta_{Is})$. A distribution function F_s describes the random behavior of independent voters in state s.

Clearly, if the states are very similar in terms of their preference distributions (in the sense that $\frac{\beta_{Ds}-\beta_{Rs}}{\beta_{Is}}$ is equal for all s, and the same state ω obtains in all states, which implies that the outcome of the popular vote is the same in all states), then the restriction that certain districts have to be formed from voters that live in the same state is no restriction for the parties, and the equilibrium is substantively the same as in our basic model.

In contrast, suppose that states differ substantially in their political preference distribution. In this case, a party is not really interested in maximizing the probability of winning the majority of seats *in a particular state*, but rather in winning a majority of seats in the federal legislature.

Consider, for example, the game in a state that is substantially more conservative than the country at-large (say, Tennessee). If all districts in Tennessee are replicas of the state, then the most likely outcome would be that all districts go to the Republicans. In contrast, if Democrats are able to win some seats in the state by concentrating their supporters, these additional seats may well be pivotal for who wins a majority in Congress. Thus, it appears that, in Tennessee, Democrats would have an incentive to concentrate their supporters, while Republicans would benefit from generating similar districts. A complete analysis of the equilibrium in our system, but with the geographical constraint that, in each district, all voters have to come from the same state, is beyond the scope of this paper and is left for future research.

Overall, the discussion here bears some resemblance to the classical decentralization theorem of Oates (1972) - except that the policy recommendations are turned upside down: In Oates (1972), public goods should be provided locally if there is pronounced heterogeneity in local preferences. In this case, respecting local preferences is more important than realizing scale economies or internalizing spillovers. When the public good in question is the legitimacy of a national election, considerable heterogeneity between different local polities calls for gerrymandering at the national level, and homogeneity makes gerrymandering at the local level tolerable.

5.5 Relationship to the divide-the-dollar / Colonel Blotto game

Our formal analysis in the previous section is based on strictly competitive game in which parties assign partisan and independent voters to districts to maximize their probability of winning an election. This can be regarded as a variation of the well known Colonel Blotto or divide-the-dollar game. In the Colonel Blotto game, troops are allocated over various battlefields in an attempt to win a war.[20] In the divide-the-dollar game, two parties compete for the support of voters. Each party is endowed with one dollar. A party proposal specifies which voter receives how much of that dollar.

Our game differs from the canonical divide-the-dollar game in two important points. Most importantly, we look at sequential rather than simultaneous moves. As is well-known, see e.g. the discussion in Laslier and Picard (2002), the simultaneous move game gives rise to Condorcet cycles and admits only equilibria in mixed strategies. The same would apply in our model if we made the parties choose their allocation of voters in all districts simultaneously. In particular, note that, in a simultaneous move setup, there is no pure strategy equilibrium in which all districts are replicas of the at-large district: Hypothesize an equilibrium in which both parties distribute D partisans, R partisans and independent voters uniformly, with the implication that every district is a copy of the at-large-district. Then, each party has an incentive to deviate so as to make a majority of districts more favorable to itself, by taking out some of its partisans from the complementary set of districts and reassigning them.

A further difference to the divide-the-dollar game comes from the consideration of asymmetric endowments. In the divide-the-dollar-game, each party has the same endowment, one dollar. A comprehensive treatment of gerrymandering needs to allow for the possibility that one party has more partisan supporters than the other. Moreover, allowing for independent voters yields randomness in voting outcomes, whereas payoffs are deterministic in the divide-the-dollar game. In the divide the dollar game, a voter supports a party with probability 1 if it offers more than the rival party, flips a coin if the two offers are equal, and votes for the rival otherwise. Here, with the behavior of independent voters governed by a continuous probability distribution, there is no such discontinuity.

These differences raise two questions: First, is it true that the divide-the-dollar game and our game of gerrymandering can be viewed as different formalizations of broadly similar strategic situations? Second, to what extent is our possibility result for an implementation of the popular vote sensitive to alternative modeling choices?

We provide detailed answers in part B of the Appendix. To answer the first question, we show that a simplified version of our model, combined with simultaneous rather than sequential moves, yields the divide-the-dollar game analyzed in Myerson (1993). Our answer to the second question is that modeling choices can be altered without upsetting our basic insight; i.e. that it is possible to implement the popular vote with alternating moves in a sequential game of gerrymandering. Specifically, we provide an illustration based on a simple model with the following features: Districts are filled not over various rounds,

[20] The classical reference is Gross and Wagner (1950), see Kovenock and Roberson (2020) for a recent treatment.

but one after the other. Moreover, there are no independent voters, with the consequence that payoff functions are deterministic and discontinuous, i.e. payoff functions are as in the divide-the-dollar game.

5.6 Maximizing the expected number of districts

To what extent does our analysis rest on the assumption that parties seek to maximize the probability of winning a majority of seats? An alternative objective is the maximization of the expected number of seats. We argue in the following that this alters the parties' incentives. Even though, by Theorem 1, each party can make sure that it wins a majority of seats whenever it wins the popular vote, it may now prefer to follow a different strategy.

The fraction of districts won by, say, party D is a random variable with a Poisson binomial distribution. Its expected value is given by

$$\begin{aligned} \Pi_D(\sigma_D, \sigma_R) \;&:=\; \tfrac{1}{2N} \textstyle\sum_{k=1}^{2N} \pi_{Dk}(\sigma_D, \sigma_R) \\ &=\; \tfrac{1}{2N} \textstyle\sum_{k=1}^{2N} \; F\left(\frac{\sigma_{Dk}^D + \sigma_{Rk}^D - \left(\sigma_{Dk}^R + \sigma_{Rk}^R\right)}{\sigma_{Dk}^I + \sigma_{Rk}^I} \right) \;. \end{aligned}$$

The parties' incentives are now shaped by the curvature of F. To see this, consider the following assumptions which imply that F is symmetric and that a party faces decreasing returns from assigning own partisans when being ahead in a district, and increasing returns when lagging behind.

i) **Symmetry.** For any x, $F(x) = 1 - F(-x)$.

ii) **Curvature.** $f'(x) < 0$, for $x > 0$,

iii) **Inada.** $\lim_{x \to 0} f(x) = \infty$ and $\lim_{x \to \infty} f(x) = 0$.

In this case, the logic that own partisan supporters should be concentrated on at most $N+1$ districts no longer applies. A party with more partisan supporters than its competitor will try to spread its advantage evenly over many districts. It mostly operates on the concave part of F and therefore seeks to equate the marginal returns from the assignment of own partisan supporters. Analogously, a party that has fewer partisan supporters might now have an incentive to concentrate them on a narrow set of districts, i.e. to engage in packing: With the number of seats as the objective, securing a small number with a large probability can be preferable to winning a majority with a small probability. More formally, the party with fewer partisan supporters mostly operates on the convex part of F and therefore prefers a convex combination with own partisans concentrated in some districts and none in others over an even distribution across districts.

6 Concluding Remarks

In a democracy, voters elect their representatives. However, when representatives are elected by plurality rule in districts, gerrymandering has the potential of reversing this relationship into one where parties select their voters.

A majority party that can unilaterally select the district map in which future elections are held can usually guarantee that it maintains majority status even if a majority of voters should prefer the opposition party in future elections. This subverts a fundamental principle of democracy.

Furthermore, the US Supreme Court, while recognizing that gerrymandering is very problematic, has generally refused to intervene against gerrymanders because there is no operational procedure to separate an "excessive" gerrymander from an admissible one. Thus, while several state courts have taken a more aggressive stance against gerrymandering, it is unlikely that there are purely judicial solutions to the gerrymandering problem.

In this paper, we have developed a system for redistricting legislative districts that has several attractive features. In particular, we show that the equilibrium district map satisfies the property that the election (almost surely) leads to the party winning the popular vote receiving a majority of seats in the legislature. Furthermore, a robust representation of the minority party in the legislature can be guaranteed by using the modified system discussed in Section 5.1.

While it is unlikely that majority parties will simply resign their redistricting powers and adopt the system proposed in this paper, in many states, voters have the power to institute a new redistricting process through a referendum process.

References

Benoît, Jean-Pierre and Vijay Krishna, "Multiple-Object Auctions with Budget Constrained Bidders," *The Review of Economic Studies*, 2001, *68* (1), 155–179.

Börgers, Tilman, Daniel Krähmer, and Roland Strausz, *An introduction to the theory of mechanism design*, Oxford University Press, USA, 2015.

Callander, Steven, "Electoral Competition in Heterogeneous Districts," *Journal of Political Economy*, 2005, *113* (5), 1116–1145.

Coate, Stephen and Brian Knight, "Socially Optimal Districting: A Theoretical and Empirical Exploration," *The Quarterly Journal of Economics*, 2007, *122* (4), 1409–1471.

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.

Gross, Oliver Alfred and R. A. Wagner, "A Continuous Colonel Blotto Game," Santa Monica, CA: RAND Corporation 1950.

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

Iossa, Elisabetta, Patrick Rey, and Michael Waterson, "Organizing Competition for the Market," CEPR Discussion Paper, No. 13461, 2019.

Jackson, Matthew, "A crash course in implementation theory," *Social Choice and Welfare*, 2001, *18* (4), 655—708.

Katz, Jonathan N., Gary King, and Elizabeth Rosenblatt, "Theoretical Foundations and Empirical Evaluations of Partisan Fairness in District-Based Democracies," *American Political Science Review*, 2020, *114* (1), 164–178.

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.

Krasa, Stefan and Mattias K. Polborn, "Political competition in legislative elections," *American Political Science Review*, 2018, *112* (4), 809–825.

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.

Mas-Colell, A., M. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995.

McCarty, Nolan, Keith T Poole, and Howard Rosenthal, "Does gerrymandering cause polarization?," *American Journal of Political Science*, 2009, *53* (3), 666–680.

Moore, John and Rafael Repullo, "Subgame Perfect Implementation," *Econometrica*, 1988, *56* (5), 1191–1220.

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

Oates, Wallace E., *Fiscal Federalism*, Edward Elgar Publishing, 1972.

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.

Polsby, Daniel D. and Robert D. Popper, "The Third Criterion: Compactness as a procedural safeguard against partisan gerrymandering," *Yale Law and Policy Review*, 1991, *9* (2), 301–353.

Roth, Alvin E., "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics," *Econometrica*, 2002, *70* (4), 1341–1378.

Van Weelden, Richard, "The welfare implications of electoral polarization," *Social Choice and Welfare*, 2015, *45* (4), 653–686.

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

A Proofs

A.1 Proof of Proposition 1

We first prove statement a) in Proposition 1. We describe a strategy for party R. Subsequently, we verify that, upon playing this strategy, party R ensures that all district outcomes approximate the popular vote.

Neutralizing moves. In round 1, party R moves first. In this case, party R chooses, for any district k,

$$\sigma^D_{kR1} = \frac{1}{L}\,\beta_D\ , \quad \sigma^R_{kR1} = \frac{1}{L}\,\beta_R \quad \text{and} \quad \sigma^I_{kR1} = \frac{1}{L}\,\beta_I\ ,$$

i.e. the mix of the different types of voters is as in the electorate at-large. R moves second in round 2, and then, immediately after, R moves first in round three. When party R makes these moves it can condition on the choices of D in rounds 1 and 2. We say that R's move in round 2 *neutralizes* D's move in round 1 if, for every district k,

$$\sigma^D_{kR2} = \sigma^R_{kD1} \quad \text{and} \quad \sigma^R_{kR2} = \sigma^D_{kD1}\ .$$

Analogously, we say that R's move in round 3 *neutralizes* D's move in round 2 if

$$\sigma^D_{kR3} = \sigma^R_{kD2} \quad \text{and} \quad \sigma^R_{kR3} = \sigma^D_{kD2}\ .$$

Thus, a neutralizing move of R in round l is such that R assigns as many own supporters to district k as D did in the previous round, and also assigns as many opponents to district k as D did before. Consequently, after R has moved in round 3, any district k is a replica of the at-large-district: For any k,

$$\frac{\sum_{l=1}^{2}(\sigma^D_{kDl} - \sigma^R_{iDl}) + \sum_{l=1}^{3}(\sigma^D_{kRl} - \sigma^R_{iRl})}{\sum_{l=1}^{2}\sigma^I_{kDl} + \sum_{l=1}^{3}\sigma^I_{kRl}} \quad = \quad \frac{\beta_D - \beta_R}{\beta_I} \quad = \quad 0\ ,$$

where the last equality follows from the assumption of symmetry, $\beta_D = \beta_R$.

Feasibility. At this stage, a discussion of feasibility is warranted: Are such neutralizing moves feasible? For L small, they are not, at least not, in general. For instance, if $L = 3$, then the move in round 3 is already the last move of party R and it will have to assign voters subject to a binding budget constraint that may prevent a neutralization in all districts. We now argue that, for large L, however, there is a number of initial rounds where such a neutralization is possible without violating the feasibility constraint.

After l rounds of play, denote the budgets available to R for the last $L - l$ rounds by

$$2N\ \beta^R_{Rl}\ , \quad 2N\ \beta^D_{Rl}\ , \quad \text{and} \quad 2N\ \beta^I_{Rl}\ .$$

Feasibility of a neutralizing move of R in round l requires that,

$$\beta_{Rl}^{R} \quad \geq \quad \frac{1}{2N} \sum_{k=1}^{2N} \sigma_{kRl}^{R} \quad = \quad \frac{1}{2N} \sum_{k=1}^{2N} \sigma_{kDl-1}^{D} \,,$$

and

$$\beta_{Rl}^{D} \quad \geq \quad \frac{1}{2N} \sum_{k=1}^{2N} \sigma_{kRl}^{D} \quad = \quad \frac{1}{2N} \sum_{k=1}^{2N} \sigma_{kDl-1}^{R} \,.$$

Now recall that, in any round l, the pair $(\sigma_{iRl}^{D}, \sigma_{iRl}^{R})$ has to lie in an $\frac{1}{L}$-simplex. Thus, feasibility is guaranteed provided that

$$2N\,\beta_{Rl}^{R} \quad \geq \quad \frac{1}{L} \qquad \text{and} \qquad N\beta_{Rl}^{D} \quad \geq \quad \frac{1}{L}\,.$$

In this case, for any district k, all points in the $\frac{1}{L}$-simplex are available to R.

Finally, note that, for l small, and L large β_{Rl}^{R} is close to β_R and β_{Rl}^{D} is close to β_D. Thus, for L is sufficiently large, these constraints are automatically satisfied for small l, implying that neutralizing moves of R are feasible in early rounds.

A binding feasibility constraint. Suppose that there is some round l' where neutralizing moves are no longer feasible for R. We will now argue that, in such a case, for L large, R can still get close to neutralization. To make this point, we relate the budget available to R in round l' to the budget available to D in round $l'-1$, on the assumption that the neutralizing strategy was feasible in all previous rounds. Note that, for any l,

$$\begin{aligned}
2N\,\beta_{Rl}^{R} &= 2N\,\beta_R - \textstyle\sum_{k=1}^{2N}\sum_{j=1}^{l-1}\sigma_{kRj}^{R} \\
&= 2N\,\beta_R - 2N\,\tfrac{\beta_R}{L} - \textstyle\sum_{k=1}^{2N}\sum_{j=2}^{l-1}\sigma_{kRj}^{R} \\
&= 2N\,\beta_R - 2N\,\tfrac{\beta_R}{L} - \textstyle\sum_{k=1}^{2N}\sum_{j=1}^{l-2}\sigma_{kDj}^{D} \\
&= 2N\,\beta_R - 2N\tfrac{\beta_R}{L} - \left(2N\,\beta_D - 2N\,\beta_{Dl-1}^{D}\right) \\
&= 2N\,\beta_{Dl-1}^{D} - 2N\,\tfrac{\beta_R}{L}\,.
\end{aligned} \tag{3}$$

where the last line follows from the assumption of symmetry, $\beta_D = \beta_R$. Also note that, by the feasibility constraint for D,

$$\sum_{k=1}^{2N} \sigma_{kDl-1}^{D} \quad \leq \quad 2N\,\beta_{Dl-1}^{D}\,, \tag{4}$$

and that, using (3), the feasibility constraint for R in round l can be written as

$$\sum_{k=1}^{2N} \sigma_{kRl}^{R} \quad \leq \quad 2N\,\beta_{Dl-1}^{D} - 2N\,\frac{\beta_R}{L}\,.$$

Thus, for L large, R faces a constraint in round l that is close to the constraint for D in round $l-1$. Still, a problem of feasibility arises when the choices of D in round $l-1$ are such that R lacks the endowment that would be needed to neutralize this move. This constellation arises if

$$2N\,\beta^D_{Dl-1} - 2N\,\frac{\beta_R}{L} \quad < \quad \sum_{k=1}^{2N} \sigma^D_{kDl-1} \quad \leq \quad 2N\,\beta^D_{Dl-1}\,.$$

This chain of inequalities has to hold in round l', the round where the feasibility constraint starts to bind.

Approximating neutralizing moves. We can now complete the description of R's strategy, by specifying R's moves from round l' onwards.

Let

$$\Delta := \sum_{k=1}^{2N} \sigma^D_{kDl'-1} - \Big(2N\,\beta^D_{Dl'-1} - 2N\,\frac{\beta_R}{L}\Big)$$

and also note that, as an implication of the feasibility constraint for D,

$$\frac{\Delta}{2N} \quad \leq \quad \frac{\beta_R}{L}\,. \tag{5}$$

Then define the strategy for R in round l' so that, for every k,

$$\sigma^R_{kRl'} = \sigma^D_{kDl'-1} - \frac{\Delta}{2N}\,,$$

and note that this construction ensures that R's feasibility constraint holds as an equality.

As a consequence, in all rounds larger than l', there are no R partisans left for an assignment by party R. Also, there are at most $2N\,\frac{\beta_D}{L}$ D partisans left for an assignment by party D.[21] Therefore, for any k,

$$\begin{aligned}
\textstyle\sum_{l=1}^{L}\sigma^D_{kDl} - \sum_{l=1}^{L}\sigma^R_{kRl} \quad &= \quad \textstyle\sum_{l=1}^{l'-2}\sigma^D_{kDl} - \sum_{l=2}^{l'-1}\sigma^R_{kRl} - \frac{\beta_R}{L} \\
&\qquad \textstyle +\sigma^D_{kDl'-1} - \sigma^R_{kRl'} + \sum_{l=l'}^{L}\sigma^D_{kDl} \\
&= \quad \textstyle -\frac{\beta_R}{L} + \frac{\Delta}{2N} + \sum_{l=l'}^{L}\sigma^D_{kDl} \\
&\leq \quad 2N\,\frac{\beta_D}{L}\,,
\end{aligned} \tag{6}$$

where the last inequality holds because of (5). Thus, for L large, $\sum_{l=1}^{L}\sigma^D_{iDl} - \sum_{l=1}^{L}\sigma^R_{iRl}$ is bounded from above by a term that is close to 0.

[21] To see this, note that $2N\,\beta^D_{Dl'-1} - 2N\,\frac{\beta_R}{L} \quad < \quad \sum_{k=1}^{2N}\sigma^D_{kDl'-1}$ implies $2N\,\beta^D_{Dl'} < 2N\,\frac{\beta_D}{L}$.

A lower bound is found on the (counterfactual) assumption that the feasibility constraint never binds, which implies

$$\sum_{l=1}^{L} \sigma_{kDl}^{D} - \sum_{l=1}^{L} \sigma_{kRl}^{R} \quad \geq \quad -\tfrac{\beta_R}{L} \ . \tag{7}$$

Together, inequalities (6) and (7) imply that, for any district k,

$$\lim_{L\to\infty} \quad \sum_{l=1}^{L} \sigma_{kDl}^{D} - \sum_{l=1}^{L} \sigma_{kRl}^{R} \qquad = 0 \ . \tag{8}$$

We can proceed in the analogous way for the specification of σ_{kRl}^{D}. The critical value of l where the feasibility constraint starts to bind may be different for the assignment of D partisans by R. But this is inconsequential for the conclusion that, for any k,

$$\lim_{L\to\infty} \quad \sum_{l=1}^{L} \sigma_{kRl}^{D} - \sum_{l=1}^{L} \sigma_{kDl}^{R} = 0 \ . \tag{9}$$

Consequently, in any district k, for L large we have

$$\frac{\sigma_k^D - \sigma_k^R}{\sigma_k^I} \quad \simeq \quad \frac{\beta_D - \beta_R}{\beta_I} \quad \simeq \quad 0 \ ,$$

where

$$\sigma_k^D := \sigma_{Dk}^D + \sigma_{Rk}^D \ , \sigma_k^R := \sigma_{Dk}^R + \sigma_{Rk}^R \quad \text{and} \quad \sigma_k^I := \sigma_{Dk}^I + \sigma_{Rk}^I \ .$$

Thus, any district is close to a replica of the at-large district as far as the mix between republican, democrat and independent voters is concerned, and for $L \to \infty$ there is actually convergence to the mix of these voter types in the electorate at large. Since F is a continuous function, this also implies that

$$\pi_{Dk}^L = F\left(\frac{\sigma_k^D - \sigma_k^R}{\sigma_k^I}\right)$$

converges to

$$\pi_D^* = F\left(\frac{\beta_D - \beta_R}{\beta_I}\right)$$

as L grows without limit.

Statement b) in Proposition 1. The prove of statement a) in Proposition 1 constructs a strategy for party R and verifies that it has the property calimed in statement a). Party D has "the same" strategy available: For round 1, choose in every district a mix of voters that is as in the electorate at-large. For all other early rounds so that the feasibility constraint is not binding, use the move in round l to neutralize R's move in round $l-1$. For the round in which the feasibility becomes binding, assign voters so that the departure from full normalization is the same in all districts and the remaining budget is exhausted. Thus, the proof of statement b) is a straightforward adaptation of the proof of statement a) and therefore omitted.

A.2 Proof of Proposition 2

Party D wins the popular vote whenever $\omega < \frac{\beta_D}{\beta_I}$. Hence, to ensure winning a majority of seats whenever $\omega < \frac{\beta_D}{\beta_I}$, party D needs to ensure that there are N districts so that, after L rounds of play,

$$\frac{\sum_{l=1}^{L} \sigma_{kDl}^{D} + \sum_{l=1}^{L} \sigma_{kDl}^{R}}{2 - \sum_{l=1}^{L} \sigma_{kDl}^{D} - \sum_{l=1}^{L} \sigma_{kDl}^{R}} > \omega \ ,$$

whenever $\frac{\beta_D}{\beta_I} > \omega$. Since the distribution F of ω is continuous, the probability that

$$\frac{\sum_{l=1}^{L} \sigma_{kDl}^{D} + \sum_{l=1}^{L} \sigma_{kDl}^{R}}{2 - \sum_{l=1}^{L} \sigma_{kDl}^{D} - \sum_{l=1}^{L} \sigma_{kDl}^{R}} > \omega \ ,$$

is equal to the probability that

$$\frac{\sum_{l=1}^{L} \sigma_{kDl}^{D} + \sum_{l=1}^{L} \sigma_{kDl}^{R}}{2 - \sum_{l=1}^{L} \sigma_{kDl}^{D} - \sum_{l=1}^{L} \sigma_{kDl}^{R}} \geq \omega \ .$$

Therefore, $\pi_D^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_D) = 1$ holds when there are N districts so that, after L rounds of play,

$$\frac{\sum_{l=1}^{L} \sigma_{kDl}^{D} + \sum_{l=1}^{L} \sigma_{kDl}^{R}}{2 - \sum_{l=1}^{L} \sigma_{kDl}^{D} - \sum_{l=1}^{L} \sigma_{kDl}^{R}} \geq \frac{\beta_D}{\beta_I} \ .$$

Consider the following strategy for party D: In all rounds l, choose $\sigma_{kDl}^{D} = 0$, for $k \leq N$ and $\sigma_{kDl}^{D} = \frac{2\beta^D}{L}$, for all $k > N$. Consequently,

$$\begin{aligned}
\frac{\sum_{l=1}^{L} \sigma_{kDl}^{D} + \sum_{l=1}^{L} \sigma_{kDl}^{R}}{2 - \sum_{l=1}^{L} \sigma_{kDl}^{D} - \sum_{l=1}^{L} \sigma_{kDl}^{R}} &= \frac{2\beta_D + \sum_{l=1}^{L} \sigma_{kDl}^{R}}{1 - 2\beta_D + 1 - \sum_{l=1}^{L} \sigma_{kDl}^{R}} \\
&\geq \frac{2\beta_D}{2(1-\beta_D)} \\
&= \frac{\beta_D}{\beta_I}
\end{aligned}$$

Thus, whenever the state of the world ω is such that $\beta_D < \omega\ \beta_I$, implying that D wins the popular vote, then D also wins all districts with an index $k \in \{N+1, \dots, 2N\}$ with probability 1.

A.3 Proof of Proposition 3

Party R wins the popular vote whenever the state ω is such that $\omega > \frac{\beta_D}{\beta_I} = \frac{\beta_D}{1-\beta_D}$. To win a majority of seats in all such states, after L rounds of play, there need to be at least N districts with

$$\frac{\sum_{l=1}^{L} \sigma_{kDl}^{D} + \sum_{l=1}^{L} \sigma_{kDl}^{R}}{2 - \sum_{l=1}^{L} \sigma_{kDl}^{D} - \sum_{l=1}^{L} \sigma_{kDl}^{R}} < \frac{\beta_D}{1-\beta_D} \ .$$

Equivalently, there need to be N districts with a percentage share of partisan D voters below β_D. In the following, we show that, for L large, party R indeed has a strategy available that ensures this outcome. The proof will be indirect: We will show that there is no strategy for party D that so that the percentage share of partisan D voters is larger than β_D in at least $N+1$ districts.

District ranks. Parties assign voters to districts over various rounds. We denote by $s^D_{k,l}$ the percentage share of partisan D voters in district k after l rounds of play. We denote the corresponding mass of partisan D voters by $\mu^D_{k,l}$. We will often rank districts according to the share of D voters. Let $r_l(k) \in \{1, \ldots, 2N\}$ be the rank of district k after l rounds of play. We assign ranks so that $r_l(k) > r_l(k')$ implies $s^D_{k,l} > s^D_{k',l}$. Hence, the district with largest share of D voters has rank $2N$, the district with the second-largest share has rank $2N-1$ and so on. The mapping r_l is taken to be injective implying that every rank in $\{1, \ldots, 2N\}$ is assigned. Thus, if two districts have the same share of partisan D voters one is (arbitrarily) assigned a higher rank than the other. What matters for the analysis that follows is that, if some district k has rank r, this implies that there are $2N-r$ further districts with a share of at least $s^D_{k,l}$.

Party Objectives. As explained above, we seek to show that there is no strategy for party D that so that the percentage share of partisan D voters is larger than β^D in at least $N+1$ districts. We therefore assume that it is party D's objective to maximize the percentage share of partisan D voters in the district with rank N after L rounds of play. Specifically, we will show that party R has a strategy under which this percentage share will be (weakly) below β_D, on the assumption that D's objective is to maximize this share. As an implication, party R's strategy implies a percentage share (weakly) below β_D, for any strategy of party D.[22]

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

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

ii) Add D partisans to the two lowest ranked districts until the mass of D partisans equals the mass in the district with the third lowest rank. From then on, keep the

[22]This is an implication of the game being zero sum. Any equilibrium strategy of party R solves a maximin-problem, i.e. it maximizes party R's payoff under the assumption that party D' strategy is chosen with the objective to make the maximum attained by R as small as possible, see e.g. Osborne and Rubinstein (1994). Thus, if party D does not behave this way, the payoff realized by party R can only get larger.

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 D partisans to the $N-1$ lowest ranked districts until the mass of D voters equals the mass in the district with rank N. From then on, don't add further D partisans to one of the bottom N districts.

iv) Add D voters to the top ranked district.

v) If there are still D voters left in the budget after a mass of $\frac{1}{L}$ D voters has been assigned to the top ranked district, add D 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 D 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 $\{\beta^D_{Dl}\}_{l=1}^L$.

Note that since party R never affects the ranking of districts, the ranking of districts in any round is entirely due to party D. We now argue that it entails no loss of generality to assume that party D's moves do neither affect the ranking of districts.

To be specific, consider party D's move in a round $l+1$, where l is odd, implying that D moves first in round $l+1$.[23] Suppose that two districts k and k' are such that $\mu^D_{k,l} > \mu^D_{k',l}$. Also suppose that, after party D's move in round $l+1$, the ranking is reversed, so that $\mu^D_{k,l+1} < \mu^D_{k',l+1}$. We now argue that an equivalent outcome can be induced without a rank reversal, and with the same implications for the budget of partisan D voters.

- Note first that the the rank reversing move requires a mass of D partisans equal to

$$\sigma^D_{k,l+1} + \sigma^D_{k',l+1} = \left(\mu^D_{k,l+1} - \mu^D_{k,l}\right) + \left(\mu^D_{k',l+1} - \mu^D_{k',l}\right) \tag{10}$$

- Now consider an alternative strategy in round $l+1$, $\bar{\sigma}_{l+1} = (\bar{\sigma}_{k,l+1})_{k=1}^{2N}$ that is the same for all districts, except for the two districts k and k' with the rank reversal. Under this alternative strategy, the mass of partisan D voters in district k is raised to the high level equal to $\mu^D_{k',l+1}$ and the mass of partisan D voters in district k' is raised to the low level of $\mu^D_{k,l+1}$. Thus, this alternative strategy yields an equivalent outcome as the original strategy: Districts k and k' flip their ranks, but in any case there are $\mu^D_{k',l+1}$ D partisans in the higher ranked district and $\mu^D_{k,l+1}$ D partisans in the lower ranked district.

[23] The same logic applies to party D's moves in odd rounds. Writing this down formally would require some obvious adjustments of notation, taking account of the fact that party D moves second in those rounds.

- The mass of D partisans required under the alternative strategy is

$$\bar{\sigma}^D_{k,l+1} + \bar{\sigma}^D_{k',l+1} = \left(\mu^D_{k',l+1} - \mu^D_{k,l}\right) + \left(\mu^D_{k,l+1} - \mu^D_{k',l}\right), \tag{11}$$

and therefore equal to the mass required by the rank reversing strategy.

We can therefore assume without loss of generality that, from the initial round onward, party D assigns partisan D voters only to $N+1$ districts. The ranking of these districts can be assumed to remain unchanged throughout the whole game. From now on, we assume for notational ease, that the index k coincides with the ranking of district k, i.e. we let $r_l(k) = k$, for all k and l.

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 more than $N+1$ districts. Thus, we may assume that, for any round l,

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

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 partisan D voters to the district with rank N until the mass of D 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 partisan D voters to the two districts with ranks N and $N+1$ until the mass of D 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 partisan D voters, $\beta^D_{R1} = 0$. In any round $l \geq 2$, and as long os this is feasible, party R assigns as many partisan D voters as party D did in the previous round

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

This is clearly feasible in early rounds. If, however, party D keeps some partisan D voters for the last round so that $\beta^D_{DL} > 0$, then party R will have to assign an additional mass of $\beta^D_{DL}\frac{2N}{L}$ late in the game. Otherwise party R would violate its budget constraint. This amount is bounded from above by $\frac{2N}{L}$ and vanishes for L large. Thus, for $L \to \infty$, and given that F is a continuous *cdf*, this will not affect the parties' winning probabilities in any one district.

In the following, we will focus on the limit case $L \to \infty$. For expositional ease, and without loss of generality, we assume that $\beta^D_{DL} = 0$, and that

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

for all $l < L$.

Party R's strategy has the following implication: Whenever party R moves, it brings the mass of D 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 D 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 D voters left. But those fill at most two further districts. According to party R's strategy, as many as possible are assigned to the district with the top rank $2N$. If additional D voters are left, they go to the district with rank $2N-1$.

Now consider rounds 3 and 4:

- In round 3, party D's best response stipulates to assign an equal mass of D voters to the districts with ranks $N, N+1, \ldots, 2N-2$. Those are $N-1$ districts. Possibly, it also assigns D voters to the two 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 added voters to the two top ranked districts, then party R has additional D 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 D voters are left, they go to the district with rank $2N-1$.

Completing the argument. The strategies of parties R and D described above imply that after the last move in round L, there is an equal mass of partisan D 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 D

partisans in the district with rank N is strictly larger than β_D. Equivalently, the mass of D voters in that district exceeds 2 β_D. Then, the mass of D voters exceeds 2 β_D in all districts. Hence, the total mass of assigned D voters is strictly larger than $2N$ β_D. But this is infeasible as the two parties' total endowments with partisan D voters only sum to $4N$ β_D. Thus, the assumption that party D can generate $N+1$ districts with a percentage share of partisan D voters strictly larger than β_D leads to a contradiction, and must be false.

A.4 Proof of Proposition 4: Sketch

We now sketch how the proofs of Propositions 2 and 3 need to be adapted when there is a non-negligible fraction of partisan R voters.

On the pivotal district. When we seek to show that party D has a strategy that ensures a victory whenever it wins the popular vote, we need to show that party D can ensure to win all districts with a rank larger or equal to $N+1$ whenever $\omega \in \Omega^D$; where ranks, after some round l, now refer to the order of districts according to

$$\Delta^D(k,l) := \frac{\mu^D_{k,l} - \mu^R_{k,l}}{2\frac{l}{L} - (\mu^D_{k,l} + \mu^R_{k,l})} \ .$$

In this expression, $\mu^D_{k,l}$ denotes, as before, the total mass of partisan D voters assigned to district k over the first l rounds of play, and $\mu^R_{k,l}$ is the analogously defined mass of partisan R voters. To show that party D has such a strategy, we assume that party D seeks to maximize $\Delta^D(k,l)$ in the district with rank $N+1$, and that party R seeks to minimize this quantity. This strategy of R is the one that makes it most difficult for party D to achieve, in the district with rank $N+1$, a value of $\Delta^D(k,l)$ that exceeds $\frac{\beta_D - \beta_R}{\beta_I}$, and hence ensures winning a majority of seats whenever $\omega \in \Omega_D$. We thereby construct an equilibrium on the assumption that the pivotal district is the one with rank $N+1$.

When we seek to show that party R has a strategy that ensures a victory whenever it wins the popular vote, we provide an indirect proof. We show that there is a strategy for party R which prevents party D from winning all districts with a rank larger or equal to N whenever $\omega \in \Omega_R$. This analysis amounts to constructing an equilibrium on the assumption that the pivotal district is the one with rank N.

A.4.1 Proof of statement b) in Proposition 4

Party R's strategy. We adapt party R's strategy in the following way: The strategy for the assignment of partisan D voters is the same as in the proof of Proposition 3, except that there is an adjustment for the pivotal district which instead of being the district with rank N is now the district with rank $N+1$. In any round l, given a – for now exogenous – budget of $\beta^D_{R,l}\ \frac{2N}{L}$ partisan D voters to be assigned, proceed sequentially in the following way – until the budget of partisan D voters for that round is exhausted:

i) Add D partisans to the lowest ranked district until $\Delta^D(k,l)$ equals the value for the district with the second lowest rank. From then on, keep $\Delta^D(k,l)$ in these two districts equal.

ii) Add D partisans to the two lowest ranked districts until the joint level of $\Delta^D(k,l)$ equals the value in the district with the third lowest rank. From then on, keep $\Delta^D(k,l)$ in these three districts equal.

iii) Proceed analogously for all districts with a rank smaller or equal $N-1$. From then on, keep $\Delta^D(k,l)$ in all these districts equal. Add D partisans to the Nth lowest ranked districts until the value of $\Delta^D(k,l)$ equals the one for the district with rank $N+1$. From then on, don't add further D partisans to one of the bottom $N+1$ districts.

iv) Add D voters to the top ranked district.

v) If there are still D partisans left in the budget after a mass of $\frac{1}{L}$ D voters has been assigned to the top ranked district, add D 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 D partisans are left.

The assignment of partisan R voters is the mirror image of the assignment of D partisans by party D in the proof of Proposition 3. Party R will focus on bringing down $\Delta^D(k,l)$ in the bottom $N+1$ districts. This also implies that party R will always choose

$$\beta^R_{Rl} \leq \frac{1}{2} + \frac{1}{2N}$$

to avoid having to assign partisan R voters to a district with rank $N+2$ or larger. Now, given a budget of $\beta^R_{R,l}\ \frac{2N}{L}$ partisan R voters to be assigned, party R proceeds sequentially in the following way – until the budget of partisan R voters for that round is exhausted:

i) Add R partisans to the district with rank $N+1$ until $\Delta^D(k,l)$ falls to the value for the district with rank N. From then on, keep $\Delta^D(k,l)$ in these two districts equal.

ii) Add R partisans to the districts with ranks $N+1$ and N until their joint level of $\Delta^D(k,l)$ equals the value in the district with rank $N-1$. From then on, keep $\Delta^D(k,l)$ in these three districts equal.

iii) Proceed analogously for all districts with a rank smaller or equal $N-1$.

Party D's strategy. Party D can now respond to party R' strategy in the following way:

- Assign R partisans according to a water-level and building-towers-strategy as outlined in the proof of Proposition 3: Assign as many R partisans as party R did in the previous round. Bring the top $2N-2$ districts to a joint level of R partisans and possibly have additional R partisans in the two bottom districts.

- Assignment of D partisans: Over the L rounds of play, assign a mass of $2\beta^D$ voters to any district with a rank lager or equal to $N+1$.

Outcome in the pivotal district. Consequently, after L rounds of play, and for L sufficiently large, in any one of the top $2N-2$ districts, the mass of partisan R voters is bounded from above by $2\beta_R$. Moreover, the districts in the bottom half are filled with partisan D voters assigned by party R, and the districts in the upper half have are filled with the partisan D voters assigned by party D, i.e. $2\beta_D$ per district in the upper half. All this implies that, in the district with rank $N+1$ after L rounds of play,

$$\begin{aligned}
\Delta^D(N+1,L) &= \tfrac{\mu^D_{N+1,L}-\mu^R_{N+1,L}}{2-(\mu^D_{N+1,L}+\mu^R_{N+1,L})} \\
&= \tfrac{2\beta_D-\mu^R_{N+1,L}}{2-(2\beta_D+\mu^R_{N+1,L})} \\
&\geq \tfrac{2\beta_D-2\beta_R}{2-(2\beta_D+2\beta_R)} \\
&= \tfrac{\beta_D-\beta_R}{\beta_I} \; .
\end{aligned}$$

The inequality in the third line follows from the fact that $\frac{2\beta_D-x}{2-(2\beta_D+x)}$ is a decreasing function of x provided that $\beta_D \leq \frac{1}{2}$.

A.4.2 Proof of statement a) in Proposition 4

The reasoning parallels the one from part 1, except that we now seek to show that party R can respond to party D's optimal behavior in such a way that, after L rounds of play, it is ensured that, in the district with rank N,

$$\Delta^D(N,L) \leq \tfrac{\beta_D-\beta_R}{\beta_I} \; .$$

To achieve this outcome, party R can respond with a water-level and building-towers-strategy to party D's assignment of partisan D voters. As a consequence, there is a common level of $2N-2$ partisan D voters in the bottom $2N-2$ districts and possibly a higher level in the two top districts. Consequently, the mass of partisan D voters in any one district in the bottom $2N-2$ is bounded from above by $2\beta_D$. Moreover, R can, over the L rounds of play, assign a mass of $2\beta_R$ partisan R voters to any district with a rank smaller or equal to N. This implies that

$$
\begin{aligned}
\Delta^D(N,L) &= \tfrac{\mu^D_{N,L}-\mu^R_{N,L}}{2-(\mu^D_{N,L}+\mu^R_{N,L})} \\
&= \tfrac{\mu^D_{N,L}-2\beta_R}{2-(\mu^D_{N,L}+2\beta_R)} \\
&\leq \tfrac{2\beta_D-2\beta_R}{2-(2\beta_D+2\beta_R)} \\
&= \tfrac{\beta_D-\beta_R}{\beta_I}\ ,
\end{aligned}
$$

where the inequality in the third line follows from the fact that $\frac{x-2\beta_R}{2-(x+2\beta_R)}$ is an increasing function of x.

A.5 Proof of Theorems 1 and 2

Propositions 1 – 4 imply that for all (β_D, β_R) with $0 \leq \beta_R \leq \beta_D \leq \frac{1}{2}$, the following statements hold true:

a) For every $\varepsilon > 0$, there is $\hat{L}$ so that $L \geq \hat{L}$ implies the existence of a strategy σ_R so that, for all σ_D,

$$\pi_R^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_R) \geq 1 - \varepsilon\ .$$

b) For every $\varepsilon > 0$, there is $\hat{L}$ so that $L \geq \hat{L}$ implies the existence of a strategy σ_D so that, for all σ_R,

$$\pi_D^{VL}(\sigma_D, \sigma_R \mid \omega \in \Omega_D) \geq 1 - \varepsilon\ .$$

Thus, for $L \geq \hat{L}$, if party R plays the strategy in part a) it realizes a payoff of at least $1-\varepsilon$ conditional on $\omega \in \Omega_R$, whatever the strategy chosen by party D. Therefore, in any equilibrium party R's equilibrium payoff in these states is bounded from below $1-\varepsilon$. (It is also bounded from above by 1.) The same is true for party D. This proves Theorem 1.

For L large, the strategies constructed in the proof of Proposition 4 approximate equilibrium strategies for all (β_D, β_R) with $0 \leq \beta_R \leq \beta_D \leq \frac{1}{2}$: With these strategies party R ensures to win with probability arbitrarily close to 1 when $\omega \in \Omega_R$ and party D ensures to win with probability arbitrarily close to 1 whenever $\omega \in \Omega_D$. Theorem 2 then follows from the observation that, with these strategies, there are at most 4 districts, out of a total of $2N$ districts, that are not replicas of the electorate at large.

B Alternative modelling choices

In this part of the Appendix, we show formally that there is a close connection between the divide-the-dollar game and the competitive-districting game. In particular, we show that the competitive-districting game is equivalent to the divide-the-dollar under the following assumptions: (i) simultaneous moves, (ii) unbiased and uncorrelated behavior of independents, (iii) equal population shares of R and D partisans, and (iv) deterministic and discontinuous winning probabilities. We begin with a recap of the divide-the-dollar-game in Myerson (1993).

B.1 The divide-the-dollar game

The analysis in Myerson (1993) includes a divide-the-dollar-game with two parties and a continuum of voters. For any one voter i, party j draws an offer from a distribution with *cdf* G^j. The offers to different voters are taken to be *iid* draws from this distribution. Admissible distributions satisfy the resource constraint

$$\int_0^\infty x \, dG^j(x) = e \ ,$$

where e is the endowment that can be redistributed in the electorate. Myerson (1993) contains a proof of the following Claim: There is one and only one symmetric equilibrium. In this equilibrium, G^j is a uniform distribution with support $[0, 2e]$.

For later reference, we reproduce the proof that parties playing a uniform distribution with support $[0, 2e]$ is an equilibrium. Suppose that party 2 plays this hypothetical equilibrium strategy. We now verify that it is a best response for party 1 to also play this strategy: Suppose that party 1 makes an offer of x_i^1 to voter i. The probability that i votes for party 1 is the probability of the event that $x_i^2 \leq x_i^1$, where, by assumption, x_i^2 is a random variable with a uniform distribution on $[0, 2e]$. Thus the probability that i votes for party 1 is given by

$$prob(x_i^2 \leq x_i^1) = \begin{cases} 0, & \text{if} \quad x_i^1 < 0 \ , \\ \frac{x_i^1}{2e}, & \text{if} \quad x_i^1 \in [0, 2e] \ , \\ 1, & \text{if} \quad x_i^1 > 2e \ . \end{cases}$$

Note that offers larger than $2e$ are dominated for party 1. We may therefore assume that the support of G^1 is bounded from above by $2e$. Given G^1, the probability that voter i votes for party 1 is given by

$$\int_0^{2e} prob(x_i^2 \leq x_i^1) \, dG^1(x_i^1) \quad = \quad \tfrac{1}{2e} \int_0^{2e} x_i^1 \, dG^1(x_i^1) \quad = \quad \tfrac{1}{2} \ ,$$

where the second equality follows from the resource constraint. Thus, all distributions G^1 with support $[0, 2e]$ and mean e yield a payoff of $\frac{1}{2}$, and there is no strategy yielding a higher payoff. Playing a uniform distribution with support $[0, 2e]$ is therefore a best response for party 1.[24]

[24] Playing e with probability 1 would also be a best response for party 1; but then the strategies are no longer mutually best responses. If party 1 plays e with probability 1, then party 2 can offer a tiny fraction of voters less than e and anyone elso more than e and thereby generate a winning probability close to 1.

B.2 The districting game

We now show that a strictly competitive game of gerrymandering with simultaneous moves and unbiased independent voters is essentially equivalent to the divide-the-dollar-game introduced in the previous subsection.

B.2.1 No independents

Assume that there is a continuum of districts of mass one. Every party sends a continuum of voters with mass one to every district. We initially assume that there are no independents. This assumption will be relaxed below.

Let $\sigma_{jD} = (\sigma_{jD}^D, \sigma_{jD}^R)$ be a generic voter assignment of party D to district j. The share of democratic voters assigned to this district is denoted by σ_{jD}^D and the share of republican voters is $\sigma_{jD}^R = 1 - \sigma_{jD}^D$. Given a pair $(\sigma_{jD}, \sigma_{jR})$, district j is won by the democrats if

$$\sigma_{jD}^D + \sigma_{jD}^R > \sigma_{jR}^D + \sigma_{jR}^R \ ,$$

or, equivalently, if

$$\sigma_{jD}^D > \sigma_{jR}^R \ .$$

There is a tie if $\sigma_{jD}^D = \sigma_{jR}^R$. If $\sigma_{jD}^D < \sigma_{jR}^R$, the district is won by party R. Note the similarity to the divide-the-dollar-game: a district is won by the party that is offering more, in the sense of sending more of its supporters.

Suppose both parties have a popular vote share of $\frac{1}{2}$. Then we can formalize a strategy of party D as a distribution G^D so that

$$\int_0^1 \sigma_{jD}^D \, dG^D(\sigma_{jD}^D) = \frac{1}{2} \ .$$

Analogously, a strategy of party R is a distribution G^R so that

$$\int_0^1 \sigma_{jR}^R \, dG^R(\sigma_{jR}^R) = \frac{1}{2} \ .$$

It follows from Myerson's analysis that the unique symmetric equilibrium has both parties play a uniform distribution on $[0, 1]$.

B.2.2 Adding independents

Suppose that a fraction $1 - e$ of all voters is of the independent type. These voters vote for the democrats with probability $p_D = \frac{1}{2}$. By the law of large numbers, a fraction $\frac{1}{2}$ e votes for D with probability 1 and a fraction $\frac{1}{2}$ e votes R with probability 1. The parties' voter assignments are now triplets that also specify the fraction of independents assigned to any one district. Thus, for a generic district j, we write $\sigma_{jD} = (\sigma_{jD}^D, \sigma_{jD}^R, \sigma_{jD}^I)$ and $\sigma_{jR} = (\sigma_{jR}^D, \sigma_{jR}^R, \sigma_{jR}^I)$.

Under the above assumptions, the behavior of independents is not source of uncertainty. In district j the mass of independents voting for party D is given by

$$p_D(\sigma^I_{jD} + \sigma^R_{jD})$$

and the mass of independents voting for party R is

$$(1 - p_D)(\sigma^I_{jD} + \sigma^R_{jD}) \ .$$

With $p_D = \frac{1}{2}$ these expressions are equal so that independents play no role for which party wins district j.

Suppose that party D sends a fraction $1 - e$ of independent voters to every district. Also suppose that it draws the share of democratic voters from a uniform distribution on $[0, e]$. We now show that it is a best response for party R to follow the same strategy. Given the behavior of D, party R wins district j if

$$\sigma^D_{jD} + \sigma^D_{jR} < \sigma^R_{jD} + \sigma^R_{jR}$$

or, equivalently, if

$$\sigma^D_{jD} < \sigma^R_{jR} + \frac{1}{2}\left(\sigma^I_{jR} - (1 - e)\right) \ .$$

The probability of this event is given by

$$prob(R \text{ wins } j \mid \sigma_{jD}, \sigma_{jR}) = \begin{cases} 0, & \text{if } \ \sigma^R_{jR} + \frac{1}{2}\left(\sigma^I_{jR} - (1-e)\right) \leq 0 \ , \\ \frac{\sigma^R_{jR} + \frac{1}{2}\left(\sigma^I_{jR} - (1-e)\right)}{e}, & \text{if } \ \sigma^R_{jR} + \frac{1}{2}\left(\sigma^I_{jR} - (1-e)\right) \in [0, e] \ , \\ 1, & \text{if } \ \sigma^R_{jR} + \frac{1}{2}\left(\sigma^I_{jR} - (1-e)\right) \geq e \ . \end{cases}$$

Thus, strategies for R so that $\sigma^R_{jR} + \frac{1}{2}\left(\sigma^I_{jR} - (1 - e)\right) < 0$ or $\sigma^R_{jR} + \frac{1}{2}\left(\sigma^I_{jR} - (1 - e)\right) > e$ are dominated.

Now suppose that R draws $(\sigma^R_{jR}, \sigma^I_{jR})$ from a joint distribution G^R with marginal distributions denoted, respectively, by G^{RR} and G^{RI}. The marginal distributions satisfy the resource constraints

$$\int \sigma^R_{jR} \, dG^{RR}(\sigma^R_{jR}) = \frac{1}{2} \, e \ ,$$

and

$$\int \sigma^I_{jR} \, dG^{RI}(\sigma^I_{jR}) = 1 - e \ .$$

The probability that R wins district j is given by

$$\begin{aligned}
prob(R \text{ wins } j \mid \sigma_{jD}, \sigma_{jR}) &= \tfrac{1}{e} \int\int \sigma_{jR}^R + \tfrac{1}{2}\left(\sigma_{jR}^I - (1-e)\right)\; dG^R(\sigma_{jR}^R, \sigma_{jR}^I) \\
&= \tfrac{1}{e} \int\int \sigma_{jR}^R \, dG^R(\sigma_{jR}^R, \sigma_{jR}^I) \\
&\quad + \tfrac{1}{e}\tfrac{1}{2}\left(\int\int \sigma_{jR}^I \, dG^R(\sigma_{jR}^R, \sigma_{jR}^I) - (1-e)\right) \\
&= \tfrac{1}{e} \int \sigma_{jR}^R \, dG^{RR}(\sigma_{jR}^R) \\
&\quad + \tfrac{1}{e}\tfrac{1}{2}\left(\int \sigma_{jR}^I \, dG^{RI}(\sigma_{jR}^I) - (1-e)\right) \\
&= \tfrac{1}{2}\,.
\end{aligned}$$

Thus, party R cannot do better than mimicking the strategy of party D, thereby securing a winning probability of $\frac{1}{2}$.

B.3 Alternating moves in a divide-the-dollar game in which endowments may be unequal

To what extent are the specifics of the game laid out in Section 4 necessary for our main insight in Theorem 1? Would alternative modelling choices still give rise to the conclusion that it is possible to implement the popular vote. Here, we suggest an affirmative answer. We look at a simple model that differs from the model in the main text in various dimensions: Districts are filled not over various rounds, but one after the other. There are no independent voters and winning probabilities in districts are deterministic and discontinuous, as in the divide-the-dollar game. The party that wins the popular vote is the one that has more partisan supporters. We show that this party also wins a majority of districts.

Setup. There are 3 districts. Each district is assigned a unit mass of voters by each party. There are no independent voters. The mass of democratic voters in the electorate at large is $b_D = 3\ \beta_D$, and the mass of republican voters is $b_R = 3\ \beta_R = 3\ (1-\beta_D)$. A strategy for party D is a vector $\sigma^D = (\sigma_1^D, \sigma_2^D, \sigma_3^D)$ specifying the number of democratic voters sent to any one district. The understanding is that the number of republican voters sent to these districts is then given by $(1-\sigma_1^D, 1-\sigma_2^D, 1-\sigma_3^D)$. Analogously, a strategy for party R is a vector $\sigma^R = (\sigma_1^R, \sigma_2^R, \sigma_3^R)$.

The probability that party D wins district k is a non-decreasing function of

$$\left(\sigma_k^D + (1-\sigma_k^R)\right) - \left(\sigma_k^R + (1-\sigma_k^D)\right)\,,$$

i.e. of the difference between the numbers of democratic and republican voters in the district. Equivalently, we can view this probability simply as a function of $\sigma_k^D - \sigma_k^R$ so that

$$\pi_{Dk}(\sigma_k^D - \sigma_k^R) = \begin{cases} 1, & \text{if} \quad \sigma_k^D - \sigma_k^R > 0 \ , \\ \frac{1}{2}, & \text{if} \quad \sigma_k^D - \sigma_k^R = 0 \ , \\ 0, & \text{if} \quad \sigma_k^D - \sigma_k^R < 0 \ . \end{cases}$$

We consider a sequential game: Parties first assign voters to district 1, then to district 2 and finally to district 3. Who moves first and who moves second alternates. Specifically, party D moves first for districts 1 and 3, and second for district 2. Party R moves first in district 2, and second in districts 1 and 3. Thus, the sequence of events is as follows:

1. Party D makes a proposal for $k = 1$.
2. Party R makes a proposal for $k = 1$.
3. Party R makes a proposal for $k = 2$.
4. Party D makes a proposal for $k = 2$.
5. Party D makes a proposal for $k = 3$.
6. Party R makes a proposal for $k = 3$.

Once districts 1 and 2 are filled, the outcome in district 3 is pinned down, implying that stages 5 and 6 no longer involve strategic choices. Specifically, the mass of democratic and republican voters that the two parties can assign to the last district are given by

$$\sigma_3^D = b_D - \sigma_2^D - \sigma_1^D \quad \text{and} \quad \sigma_3^R = b_R - \sigma_2^R - \sigma_1^R \ .$$

Thus,

$$\pi_{D3}(\sigma_3^D - \sigma_3^R) \quad = \quad \begin{cases} 0, & \text{if} \quad \sigma_3^D - \sigma_3^R < 0 \ , \\ \frac{1}{2}, & \text{if} \quad \sigma_3^D - \sigma_3^R = 0 \ , \\ 1, & \text{if} \quad \sigma_3^D - \sigma_3^R > 0 \ . \end{cases} \tag{12}$$

We assume that parties maximize the number of districts that they win. In the given deterministic setup, this is equivalent to maximizing the probability of winning.

Equilibrium in the subgame that begins after district 1 has been filled. The following Proposition characterizes the equilibrium for the subgame that starts after the outcome for district 1 has been determined. We denote by b_2^D and by b_2^R the parties' endowments with own partisans for the last two districts, i.e. after district 1 has been filled. We denote by

$$\Pi_{D2} = \pi_{D2}(\sigma_2^D - \sigma_2^R) + \pi_{D3}(\sigma_3^D - \sigma_3^R)$$

party D's payoff in that subgame. We use analogous notation for party R.

Proposition 5 *Consider the subgame that begins after the outcome for district 1 has been determined. The equilibrium payoff of party D is*

$$\Pi_{D2}(b_2^D, b_2^R) \quad = \quad \begin{cases} 2, & \textit{if} \;\; b_2^D > b_2^R \;, \\ 1, & \textit{if} \;\; 2b_2^D > b_2^R \geq b_2^D \;, \\ \frac{1}{2}, & \textit{if} \;\; b_2^R = 2b_2^D \;, \\ 0, & \textit{if} \;\; b_2^R > 2b_2^D \;. \end{cases} \tag{13}$$

The payoff Π_{R2} *of party R for all districts* $k \geq 2$ *is given by*

$$\Pi_{R2}(b_2^D, b_2^R) = 2 - \Pi_{D2}(b_2^D, b_2^R) \;. \tag{14}$$

An equal split

$$\sigma_3^R = \sigma_2^R = \frac{1}{2}\, b_2^R$$

is equilibrium behavior for R, the first mover in district 2*. An equal split*

$$\sigma_3^D = \sigma_2^D = \frac{1}{2}\, b_2^D$$

is equilibrium behavior for D, the second mover in district 2*, if* $b_2^D \geq b_2^R$ *or if* $b_2^R > 2b_2^D$*.*

For a proof of Proposition 5, see Section B.4 below. Note that we can view Π_{D2} also as a function of $b_2^D - b_2^R$ so that

$$\Pi_{D2}(b_2^D - b_2^R) \quad = \quad \begin{cases} 2, & \text{if} \;\; b_2^D - b_2^R > 0 \;, \\ 1, & \text{if} \;\; 0 \geq b_2^D - b_2^R \geq -b_2^D \;, \\ \frac{1}{2}, & \text{if} \;\; b_2^D - b_2^R = -b_2^D \;, \\ 0, & \text{if} \;\; b_2^D - b_2^R < -b_2^D \;. \end{cases} \tag{15}$$

Equilibrium in the overall game. We now turn to the overall game. The following Proposition characterizes equilibrium payoffs.

Proposition 6 *Equilibrium payoffs in the three district game are*

$$\Pi_{R1} \quad = \quad \begin{cases} 3, & \textit{if} \;\; b^R > 2b^D \;, \\ 2, & \textit{if} \;\; 2b^D \geq b^R > b^D \;, \\ \frac{3}{2}, & \textit{if} \;\; b^D = b^R \;, \\ 1, & \textit{if} \;\; b^D > b^R \geq \frac{1}{2}b^D \\ 0, & \textit{if} \;\; \frac{1}{2}b^D > b^R \;. \end{cases} \tag{16}$$

and

$$\Pi_{D1} = 3 - \Pi_{R1}$$

For a proof of Proposition 6, see Section B.5 below. To facilitate the comparison to equation (15) we can also write

$$\Pi_{D1}(b^D - b^R) \quad = \quad \begin{cases} 3, & \text{if} \quad b^D - b^R > b^R \ , \\ 2, & \text{if} \quad b^R \geq b^D - b^R > 0 \ , \\ \frac{3}{2}, & \text{if} \quad b^D - b^R = 0 \ , \\ 1, & \text{if} \quad 0 > b^D - b^R \geq -b^D \ , \\ 0, & \text{if} \quad -b^D > b^D - b^R \ . \end{cases} \tag{17}$$

Note that this is perfectly symmetric. This implies, in particular, that there is no first or second mover advantage. A party's payoff is entirely determined by the margin of victory in the popular vote, i.e. by $b^D - b^R$. Also, $b^D - b^R > 0$ implies that party D wins a majority of districts and $b^D - b^R < 0$ implies that party R wins a majority of districts. Thus, the popular vote determines which party wins a majority of districts.

The proofs of Propositions 5 and 6 below also contain a characterization of equilibrium behavior. It implies that the three districts will not typically look like replicas of the electorate at large. This would require that both parties spread republican and democratic voters evenly over the three districts. Such an equal split is compatible with equilibrium behavior only in the symmetric case $\beta_D = \beta_R$.

B.4 Proof of Proposition 5

District 2, second move by D. Party D chooses σ_2^D and σ_3^D subject to

$$\sigma_2^D + \sigma_3^D \quad \leq \quad b_2^D \ ,$$

where the budget $b_2^D = b^D - \sigma_1^D$ is predetermined at this stage. Note that σ_3^R and σ_2^R are also known at this stage. Thus, given b_2^D and σ^R, party D chooses σ_2^D and σ_3^D to maximize

$$\Pi_{D2} \quad = \quad \pi_{D2}(\sigma_2^D - \sigma_2^R) + \pi_{D3}(\sigma_3^D - \sigma_3^R) \ .$$

Lemma 1 *Consider the second stage of the subgame at district 2. The payoff Π_2^D of party D for all districts $d \geq 2$ is given by*

$$\Pi_{D2} \quad = \quad \begin{cases} 2, & \textit{if} \quad b_2^D > b_2^R \ , \\ 1, & \textit{if} \quad b_2^R \geq b_2^D > \min\{\sigma_2^R, \sigma_3^R\} \ , \\ \frac{1}{2}, & \textit{if} \quad b_2^D = \min\{\sigma_2^R, \sigma_3^R\} \ , \\ 0, & \textit{if} \quad b_2^D < \min\{\sigma_2^R, \sigma_3^R\} \ . \end{cases} \tag{18}$$

The payoff Π_2^R of party R for all districts $d \geq 2$ is given by

$$\Pi_{R2} = 2 - \Pi_{D2} \ . \tag{19}$$

Proof.

1. Party D can win districts 2 and 3 if the budget suffices to pay more than σ_2^R for district 2 and more than σ_3^R for district 3. This is the case if $b_2^D > \sigma_2^R + \sigma_2^R$, or, equivalently, if $b_2^D > b_2^R$.

2. If $b_2^R \geq b_2^D > \min\{\sigma_2^R, \sigma_3^R\}$, then party D can win district $k = 2$ or district $k = 3$. For $b_2^R > b_2^D$, party D chooses either $\sigma_2^D > \sigma_2^R$ and $\sigma_3^D < \sigma_3^R$ or $\sigma_2^D < \sigma_2^R$ and $\sigma_3^D > \sigma_3^R$. For $b_2^R = b_2^D$, there is an additional possibility, which is to win both with prob $\frac{1}{2}$ by choosing $(\sigma_2^D, \sigma_3^D) = (\sigma_2^R, \sigma_3^R)$.

3. If $b_2^D = \min\{\sigma_2^R, \sigma_3^R\}$ then party D either wins district 2 or district 3 with probability $\frac{1}{2}$ and loses the other one for sure.

4. If $b_2^D < \min\{\sigma_2^R, \sigma_3^R\}$ party D loses both district 2 and district 3.

Note that for districts 2 and 3, there is a second mover advantage for party D that plays out in the following ways: (i) If D has a tiny budget advantage it can win both districts. (ii) Party D can have a much smaller budget than party R and still get one district. $\square$

District 2, first move by R. Party R chooses σ_2^R and σ_3^R to minimize Π_{D2} subject to $\sigma_2^R + \sigma_3^R = b_2^R$.

Lemma 2 $\Pi_{D2} = 0 \iff b_2^R > 2b_2^D$.

Proof.

$\Longrightarrow$: Let $\Pi_{D2} = 0$. Then it follows from (18) that $\sigma_3^R > \sigma_2^D$ and $\sigma_2^R > b_2^D$. Hence, $\sigma_3^R + \sigma_2^R > 2b_2^D$. Equivalently, $b_2^R > 2b_2^D$.

$\Longleftarrow$: Let $b_2^R > 2b_2^D$. Choose $\sigma_2^R = \sigma_3^R > b_2^D$. Then D has no chance whatever the district. Hence, $\Pi_{D2} = 0$.

Note that with $b_2^R > 2b_2^D$ an equal split by both is equilibrium behavior. $\square$

Lemma 3 $\Pi_{D2} = \frac{1}{2} \iff b_2^R = 2b_2^D$.

Proof.

$\Longrightarrow$: Let $\Pi_{D2} = \frac{1}{2}$. Then, it follows from (18) that $b_2^D = \min\{\sigma_2^R, \sigma_3^R\}$. Suppose that $\sigma_2^R \neq \sigma_3^R$. Then it is budgetary feasible for R to chose both $\sigma_2^R > b_2^D$ and $\sigma_3^R > b_2^D$ yielding $\Pi_{D2} = 0$, contrary to the assumption of optimizing behavior. Thus, it must be that $\sigma_2^R = b_2^D$ and $\sigma_3^R = \sigma_2^D$ implying $b_2^R = 2b_2^D$.

$\Longleftarrow$: Let $b_2^R = 2b_2^D$. Then, by our previous argument $\Pi_{D2} = 0$ is out of reach for R and the best possible outcome is $\Pi_{D2} = \frac{1}{2}$. This can be reached with $\sigma_2^R = \sigma_3^R = b_{m-1}^D$. Then D has a chance of winning at most one district. In this district the winning probability is $\frac{1}{2}$.

Note that with $b_2^R = 2b_2^D$, equal split is equilibrium behavior by R, but not by D. $\square$

Lemma 4 $\Pi_{D2} = 1 \quad \Longleftrightarrow \quad 2b_2^D > b_2^R \geq b_2^D$.

Proof.

$\Longrightarrow$: Let $\Pi_{D2} = 1$. It follows from Lemmas 2 and 3 that $2b_2^D > b_2^R$. It follows from (18) that $b_2^R \geq b_2^D$.

$\Longleftarrow$: Let $2b_2^D > b_2^R \geq b_2^D$. Lemmas 2 and 3 imply that $\Pi_{D2} \geq 1$. It follows from (18) that R can induce $\Pi_{D2} = 1$ by choosing an equal split $\sigma_2^R = \sigma_3^R = \frac{1}{2}b_2^R$.

Note that with $2b_2^D > b_2^R \geq b_2^D$, equal split is eq. behavior by R. For D it is eq behavior only if $b_2^R = b_2^D$. If the inequality is strict D goes for one district and leaves the other district for R. $\square$

Lemma 5 $\Pi_{D2} = 2 \Longleftrightarrow b_2^D > b_2^R$.

Proof.

$\Longrightarrow$: Let $\Pi_{D2} = 2$. It follows from Lemmas 2-4 that $b_2^D > b_2^R$.

$\Longleftarrow$: Let $b_2^D > b_2^R$. It follows from (18) that $\Pi_{D2} = 2$.

Note that with $b_2^D > b_2^R$ an equal split by both is equilibrium behavior. $\square$

B.5 Proof of Proposition 6

District 1, second move by R. Party R's problem is to choose σ_1^R and b_2^R to maximize

$$\Pi_{R1} \quad = \quad \pi_{R1}(\sigma_1^D - \sigma_1^R) + \Pi_2^R(b_2^D - b_2^R)\ ,$$

subject to

$$\sigma_1^R + b_2^R = b_R\ .$$

Lemma 6

$$\Pi_{R1} = 3 \quad \Longleftrightarrow \quad b_R > 2b_2^D + \sigma_1^D\ .$$

Proof.

$\Longrightarrow$: Let $\Pi_{R1} = 3$. Then it must be that $\pi_{R1} = 1$ and $\Pi_{R2} = 2$. This requires that $\sigma_1^R > \sigma_1^D$ and $b_2^R > 2b_2^D$. Adding these inequalities and using the budget constraint yields $b^R > 2b_2^D + \sigma_1^D$.

$\Longleftarrow$: Let $b^R > 2b_2^D + \sigma_1^D$. Then it is budgetary feasible for R to have both $\sigma_1^R > \sigma_1^D$ and $b_2^R > 2b_2^D$ and hence $\pi_{R1} = 1$ and $\Pi_{R2} = 2$.

$\square$

Lemma 7

$$b^R = 2b_2^D + \sigma_1^D \quad \Longrightarrow \quad \Pi_{R1} = 2 \ .$$

Proof. Let $b^R = 2b_2^D + \sigma_1^D$. Then, $b_2^R > 2b_2^D$ implies $\sigma_1^R < \sigma_1^D$, $\pi_1^R = 0$, $\Pi_{R2} = 2$ and hence $\Pi_{R1} = 2$. Analogously, $b_2^R = 2b_2^D$ implies $\sigma_1^R = \sigma_1^D$, $\pi_{R1} = \frac{1}{2}$, $\Pi_{R2} = \frac{3}{2}$ and hence $\Pi_{R1} = 2$. Finally, suppose that $b_2^R < 2b_2^D$ implies $\sigma_1^R > \sigma_1^D$. Since it is possible to have $b_2^R \geq b_D^2$, this implies that $\pi_{R1} = 1$, $\Pi_{R2} = 1$ and hence $\Pi_{R3} = 2$. $\square$

Lemma 8

$$2b_2^D + \sigma_1^D > b^R > b^D \quad \Longrightarrow \quad \Pi_{R1} = 2 \ .$$

Proof. By Lemma 7, $\Pi_{R1} = 2$ is an upper bound for the payoff that can be reached if $2b_2^D + \sigma_1^D > b^R$. Reaching the upper bound is possible if b_2^R and σ_1^R can be chosen so that $b_2^R \geq b_2^D$ and $\sigma_1^R > \sigma_1^D$. This is possible only if $b^R > b^D$. $\square$

Lemma 9

$$b^R = b^D \quad \Longrightarrow \quad \Pi_{R1} = \frac{3}{2} \ .$$

Proof. With $b^R = b^D$, party R can choose $\sigma_1^R = \sigma_1^D$ and $b_2^R = b_2^D$ which yields $\pi_{R1} = \frac{1}{2}$ and $\Pi_{R2} = 1$. It remains to be shown that $\Pi_{R2} = 2$ is out of reach for R if D behaves optimally in the first move for district 1. Suppose D does an equal split so that $\sigma_1^D = \frac{1}{3}b^D$.

Suppose first that R chooses $\sigma_1^R < \sigma_1^D$ so that $\pi_1^R = 0$. On this assumption it is optimal to choose $\sigma_1^R = 0$ and $b_2^R = b^R$. We then have $b_2^D = \frac{2}{3}b_2^R$ and $b_2^D - b_2^R = -\frac{1}{3}b_2^D$. By equation (15) this implies $\Pi_{R2} = 1$. Hence, this is not better than mimicking D's strategy for R.

Finally, suppose that R chooses $\sigma_1^R > \sigma_1^D$ so that $\pi_{R1} = 1$. We then have $b_2^D - b_2^R > 0$ and hence $\Pi_2^R = 0$. Again, this is not better than mimicking D's strategy for R. Thus, if $b^D = b^R$, then an equal split is optimal for both R and D. $\square$

Lemma 10

$$b^R < b^D \quad \Longrightarrow \quad \Pi_{R1} \leq 1 \ .$$

Proof. The arguments in the proof of the previous Lemma imply that D can ensure that R wins at most one district by choosing an equal split. $\square$

Lemma 11

$$b^D > b^R \geq \frac{1}{2} b^D \quad \Longleftrightarrow \quad \Pi_{R1} = 1 \ .$$

Proof.

$\Longrightarrow$: Let $b^R \geq \frac{1}{2} b^D$. If $b^R \geq b_2^D$, then R can ensure $\Pi_{R1} = 1$ by choosing $b_2^R = b^R$. Hence, suppose that $b_2^D > b^R$. Then R can ensure $\Pi_{R1} = 1$ by choosing $\sigma_1^R = b^R > \sigma_1^D = b^D - b_2^D$. (Note that $b^R \geq \frac{1}{2} b^D$ and $b_2^D > b^R$ imply that $b^D - b_2^D < b^R$.)

$\Longleftarrow$: Suppose that $\frac{1}{2} b^D > b^R$ or, equivalently, that $b^D > 2b^R$. Then D can choose $\sigma_1^D > b^R$ and $b_2^D > b^R$. Hence, $\pi_{1D} = 1$ and $\Pi_{2D} = 2$. Thus, $\Pi_{R1} \neq 1$.

$\square$

Lemma 12

$$\frac{1}{2} b^D > b^R \quad \Longrightarrow \quad \Pi_{R1} = 0 \ .$$

Proof. Suppose that $\frac{1}{2} b^D > b^R$ or, equivalently, that $b^D > 2b^R$. Then D can choose $\sigma_1^D > b^R$ and $b_2^D > b^R$. Hence, $\pi_{D1} = 1$ and $\Pi_{D2} = 2$. Thus, $\Pi_{R1} = 0$. $\square$

To wrap up, we have

$$\Pi_{R1} \quad = \quad \begin{cases} 3, & \text{if} \quad b^R > 2b_2^D + \sigma_1^D \ , \\ 2, & \text{if} \quad 2b_2^D + \sigma_1^D \geq b^R > b^D \ , \\ \frac{3}{2}, & \text{if} \quad b^D = b^R \ , \\ 1, & \text{if} \quad b^D > b^R \geq \frac{1}{2} b^D \\ 0, & \text{if} \quad \frac{1}{2} b^D > b^R \ . \end{cases} \tag{20}$$

and

$$\Pi_{D1} = 3 - \Pi_{R1} \ .$$

District 1, first move by R. The analysis above leaves open what happens if $b^R > b^D$. In this case, D can reach at most $\Pi_{D1} = 1$. Reaching this upper bound is possible if and only if b_2^D and σ_1^D can be chosen so that $2b_2^D + \sigma_1^D \geq b^R$. This is possible if and only if $2b^D \geq b^R$. This observation completes the proof of Proposition 6.