Fair Seat Allocations in Plurality Systems with Implications for the Efficiency Gap and Detecting Partisan Gerrymandering

Jeffrey T. Barton
Department of Mathematics
Birmingham-Southern College

900 Arkadelphia Road
Birmingham, AL 35254
jbarton@bsc.edu

October 20, 2021

Abstract

We construct a baseline standard for determining fair seat proportions in plurality voting for two-party and multiparty elections. The construction is intuitive and computationally simple, and it produces sensible theoretical and empirical results when applied to U.S. and Canadian elections. It also includes the efficiency gap standard as its best linear approximation at parity. The new standard implies that plurality voting systems are inherently biased in favor of large parties, irrespective of any practical limitations or gerrymandering. Finally, we apply the standard to the detection of partisan gerrymanders in a way that remedies well-known deficiencies in measures like the efficiency gap.

Keywords: gerrymandering, efficiency gap, plurality voting, proportionality, multiparty, wasted votes

1 Introduction

Our first goal is to produce a baseline for fair seat proportions in first-past-the-post elections. We seek a standard for two-party and multiparty cases that is intuitive, simple to compute, consistent with real election results, and can be derived rigorously based on a handful of reasonable assumptions about fair maps. Second, we aim to apply the new standard to the problem of detecting partisan gerrymanders, an application that will require us to consider not only the unfairness in a map's allocation of seats, but the stability of such unfairness as well. The explicit consideration of the stability of a map's unfairness will be key to removing the limitations in measures such as the efficiency gap.

We begin by considering a districting plan, or map, as consisting of N equal-sized districts, $C_1, C_2, \ldots, C_N$. We let Party A be the majority party and Party B the minority party. The total number of votes for each party is given by V_A and V_B, respectively, so with V as the overall total vote, $V_A + V_B = V$. The corresponding overall vote proportions for Party A and Party B are $v_A = \frac{V_A}{V}$ and $v_B = \frac{V_B}{V}$, so that $v_A \geq 0.5$, $v_B \leq 0.5$, and $v_A + v_B = 1$. We denote the number of seats won by each party by S_A and S_B so that $S_A + S_B = N$, and we let the seat proportions won by each party be given by $s_A = \frac{S_A}{N}$ and $s_B = \frac{S_B}{N}$. We note that the minority party cannot win all of the seats, so that $0 \leq S_B < N$ and $0 < S_A \leq N$, or equivalently $0 \leq s_B < 1$, and $0 < s_A \leq 1$. We denote the **fair seat proportions** for each party by s_A^* and s_B^*, while S_A^* and S_B^* will be the fair number of seats for each. Unless the context makes clear otherwise, when referring to "proportionality" for seat allocations, we mean that parties' seat proportions must equal their overall vote proportions, so $s_A = v_A$ and $s_B = v_B$.

Within each district, C_n, we let p_n be the vote proportion for Party A and q_n the vote proportion for Party B so that $p_n + q_n = 1$ for all $n = 1, \ldots, N$. Our custom will be to represent a districting plan by specifying the minority party's district vote proportions from largest to smallest: $(q_1, q_2, \ldots, q_N)$ such that $q_1 \geq q_2 \geq \cdots \geq q_N$. Party B wins any seat for which $q_n > 0.5$, and Party A wins any seat for which $p_n > 0.5$. Our convention for specifying a district map means that we list all of the minority party's victories first (from least competitive to most), followed by all of the majority party's wins (from most competitive to least). Thus $q_n > 0.5$ for all $1 \leq n \leq S_B$, and $q_n < 0.5$ for all $S_B + 1 \leq n \leq N$. Finally, we will find it convenient to use the notation "0.50^+" and "0.50^-" to indicate district vote proportions for the winner and loser, respectively, in an election that is "maximally competitive." We treat such vote shares as indistinguishable from 0.50.

2 Fair Seat Proportions for Two-Party Elections

We begin by defining the fair seat proportion, s_B^*, for the minority party in a two-party, first-past-the-post election given its overall vote proportion, v_B. This in turn determines the fair seat proportion for the majority: $s_A^* = 1 - s_B^*$.

Many methods for defining a fair seat share for the minority party are based on trying to determine its fair share of *all* seats. The approach we take here is to first note that the theoretical maximum proportion of seats that the minority can win in any election is $2v_B$. Thus the minority party has *no claim* to any seats beyond this proportion. The question we address is "What share of the seats that it can actually win does the minority party deserve?" Below we argue that the minority party deserves proportionality in the seats it can actually win. In other words, we propose that

$$\frac{s_B^*}{2v_B} = v_B,$$

or

$$s_B^* = 2v_B^2. \tag{1}$$

Below we compare this notion of equitability to others before deriving the result.

2.1 Relationship to the Efficiency Gap

The *efficiency gap* by Stephanopolous and McGhee, provides a well-known method for determining fair seat proportions based on vote proportions. It defines a fair map as one where both parties have the same number of *wasted votes*, which are votes above 50% in districts a party wins and any votes in districts a party loses. This definition of a fair map implies that $s_B^* = 2v_B - 0.50$ [1]. A way to view this standard is as an adjustment to the minority party's maximum seat share. However, this standard suffers from two important limitations that we remedy in this article. The first is that the measure, like all measures that rely only on overall vote and seat proportions, does not return intuitive results for competitive sweeps. We address this issue in Section 4. The second limitation is that the efficiency gap is too generous to the majority in that it assigns *negative* seat proportions to the minority for $0 \leq v_B < 0.25$.

The *modified efficiency gap* addresses this issue by refining the idea of wasted votes [2]. As a result of simple mathematics, there are votes that *must* be wasted by each party in every election *independent of the map that is drawn*, and votes that must be wasted by every map should not be used in judging a particular one. Thus, for a given map the modified efficiency gap only counts each party's *unnecessarily* wasted votes, and this implies that an equitable map is one where the minority party should win half of its maximum number of seats. Taking rounding into account, this yields the formula $s_B^* = \frac{1}{2N}\lfloor 2Nv_B \rfloor$, where $\lfloor * \rfloor$ is the floor function. For small N, the explicit consideration of rounding in this definition is important for distinguishing the requirement from proportionality; however, for large N, the requirement reduces to $\frac{s_B^*}{2v_B} = 0.50$, which is of course proportionality and hence too generous towards the minority party.

Having the minority party win half of its maximum number of seats means that the minority ends up halfway between the two extremes of winning all $2v_B$ seats and winning 0 seats. But if we consider what is required by the two extremes, we can see why proportionality unreasonably favors the minority as v_B gets smaller. Specifically, in order for the minority party to win its maximum number of seats, it must use all of its votes with maximum efficiency: it's vote proportion must be 0.50^+ in every district it can win and 0 in the remaining districts. This efficiency requirement on the minority party does not change as v_B decreases. At the other extreme, the efficiency required of the majority party to sweep every seat decreases as v_B decreases. Thus, the proportion of the $2v_B$ seats deserved by the minority should not be fixed at 0.5, which continues to treat both extremes equally as v_B gets smaller. Instead the proportion should move toward the extreme of $s_B = 0$ as v_B gets smaller. Moreover, the proportion must be 0.50 when $v_B = 0.5$. These requirements are of course not prescriptive of a particular rule, but they are satisfied in a simple way by our suggestion that $\frac{s_B^*}{2v_B} = v_B$.

The requirement $s_B^* = 2v_B^2$ places the new measure in between those suggested by the efficiency gap and the modified efficiency gap. We will see that, similar to the way the modified efficiency gap results from a refinement of *which votes* count as wasted, our new measure results from an additional and crucial refinement of *which seats* we consider when counting them.

We observe that for a fairly wide range of values for v_B, the new measure is often not significantly different from the one implied by the efficiency gap. For example, with $v_B = 0.40$, the efficiency gap implies $s_B^* = 0.30$, while the new measure gives $s_B^* = 0.32$. However, due to rounding even this slight difference can lead to a different number of fair seats, as it would if $N = 8$ where the efficiency gap would prescribe 2 seats for the minority while the new standard would prescribe 3. We further note that at $v_B = 0.50$, both measures produce $s_B^* = 0.50$ (as they must), as well as the same derivative: $s_B^{*\prime} = 2$. Thus the standard implied by the efficiency gap is the first order approximation to the new standard at $v_B = 0.50$. This relationship may explain the effectiveness of the efficiency gap as a standard for minority vote shares near 0.50: it is the best linear approximation to fairness at that point.

2.2 Derivation

The derivation of our standard in both the two-party and multiparty cases is based on three assumptions:

1. Parties with equal vote shares should win the same proportion of seats: $v_A = v_B \implies s_A^* = s_B^*$.

2. There must be a seat that the largest party wins with a vote proportion at least as large as its overall vote proportion.

3. In seats that the largest party must win, the winning party and losing party should have equal numbers of unnecessarily wasted votes.

In the two-party case, Assumption #2 is always true; however, in the multiparty case that will no longer be so. We outline the idea behind the derivation before providing the mathematical details.

Let C_{n_1} be a district as guaranteed by Assumption #2 so that the majority must win it with $p_{n_1} \geq v_A$. Both parties would like to minimize the number of votes they use in it. By Assumption #2, the best case for Party A (and the worst for Party B) would be $p_{n_1} = v_A$ and $q_{n_1} = v_B$. The worst case for Party A would be $p_n = 1$ and $q_n = 0$. Thus any votes above v_A are unnecessarily wasted by Party A, while any votes above 0 are unnecessarily wasted by Party B. Invoking Assumption #3 leads us to split the difference so that $p_{n_1} = 0.5(1 + v_A)$ and $q_{n_1} = 0.5v_B$.

The fact that Party B wastes fewer than v_B votes in C_{n_1} means that in the remaining $N-1$ seats, Party B's overall vote proportion is greater than v_B. On the other hand, the fact that Party B does waste *some* votes in C_{n_1} means that Party B's ability to win the remaining $N-1$ seats is diminished. Even if Party B could have won $N-1$ seats at the outset, the fact that it used some of its votes on a hopeless cause may cost it the ability to win all $N-1$ remaining seats. If after allocating some votes to C_{n_1} Party A remains the largest party over the remaining $N-1$ seats, we repeat the process of assuming an equal number of unnecessarily wasted votes in seat C_{n_2} for both parties using their updated vote proportions over the $N-1$ seats.

As long as Party B's vote proportion remains less than Party A's, Party B cannot win all of the remaining seats. Thus we continue to update vote proportions and repeat the process of allocating equitable vote shares to seats that are not in dispute until both parties have an equal share of the vote in the remaining seats. By Assumption #1, that means that once the process terminates, each party will deserve half of the remaining seats.

Next we formalize the description above, and we work under the assumption that votes are allocated to seats continuously. For x in $[0,1]$. let $v_B(x)$ represent the vote proportion for Party B in the remaining $(1-x)$ seats after Party B has allocated its equitable share of votes to the first x seats. Define $v_A(x)$ similarly so that $v_A(x) = 1-v_B(x)$. Then $v_B(0) = v_B$, and the process of allocation terminates when $v_B(x_0) = v_A(x_0) = 0.5$ for some $0 \leq x_0 \leq 1$. We note that for small Δx,

$$0.5v_B(x)\Delta x = v_B(x)(1-x) - v_B(x+\Delta x)(1-x-\Delta x).$$

This simplifies to

$$(v_B(x+\Delta x) - v_B(x))(1-x) = (v_B(x+\Delta x) - 0.5v_B(x))\Delta x,$$

or

$$\frac{v_B(x+\Delta x) - v_B(x)}{\Delta x} = \frac{v_B(x+\Delta x) - 0.5v_B(x)}{1-x}.$$

Taking a limit as $\Delta x \to 0$ yields the differential equation

$$v'_B(x) = \frac{0.5v_B(x)}{1-x}, \tag{2}$$

which produces the solution

$$v_B(x) = \frac{v_B}{\sqrt{1-x}}, \text{ and } v_A(x) = 1 - \frac{v_B}{\sqrt{1-x}}. \tag{3}$$

Solving $v_B(x_0) = 0.5$ tells us that the process terminates when $x_0 = 1-4v_B^2$, at which point each party will have equal vote shares in the remaining $4v_B^2$ seats. We describe these seats as those that are *feasibly in dispute*, or *feasibly contestable*, and by Assumption #1, Party B deserves 50% of these seats. Thus the equitable seat proportion for the minority party is given by

$$s_B^* = 2v_B^2.$$

We emphasize that the proportion of seats that are feasibly in dispute, $4v_B^2$, is not necessarily the maximum proportion of seats that Party B *can* win (which would be $2v_B$).

They are seats that Party B can feasibly win if it devotes an equitable share of votes to the seats it cannot win. While not a theoretical maximum for the proportion of seats that Party B can win, it does provide a practical upper bound as we show using data from past elections in the next section.

Finally, we note that the process does not mandate that there be competitive districts among the feasibly contestable seats. It only requires that the two parties split the seats equally. The manner in which the seats are won is open.

2.3 Application to U.S. Elections

In this section we compare the new method for determining fair seat shares to real past elections in the U.S. The database of elections we use is thanks to Greg Warrington at the University of Vermont [5]. It consists of 1142 U.S. House of Representative elections between 1972 and 2016 as well as 647 state elections from the same time period [6]. As it is not possible to gerrymander a district plan with only one district, we restrict our attention to elections with at least two seats. There are 999 federal and 1646 total such elections in the database.

Figure 1 shows the relationship between the minority party's statewide vote proportion and its seat proportion for all elections in the database for which there were at least 2 seats. The solid curve represents our new standard for equitability while the dashed curve is the proportion of feasibly contestable seats. There were 472 out of 1646 elections where the seat proportion was equitable after rounding to the nearest seat; these are represented by the dark dots.

We note that the scatter plot evinces some curvature in the relationship and that our proposed equitability curve is visually a reasonable representation of it. We should not expect the equitability curve to be a "best fit" curve because of issues like gerrymandering and geography that can impact real election maps. The fact that there are more points below the curve than above it indicates that in general there appear to be more factors working against the minority than for it, as we might expect. Finally, we see that there are no examples where the minority party has won more than the feasibly contestable proportion of seats.

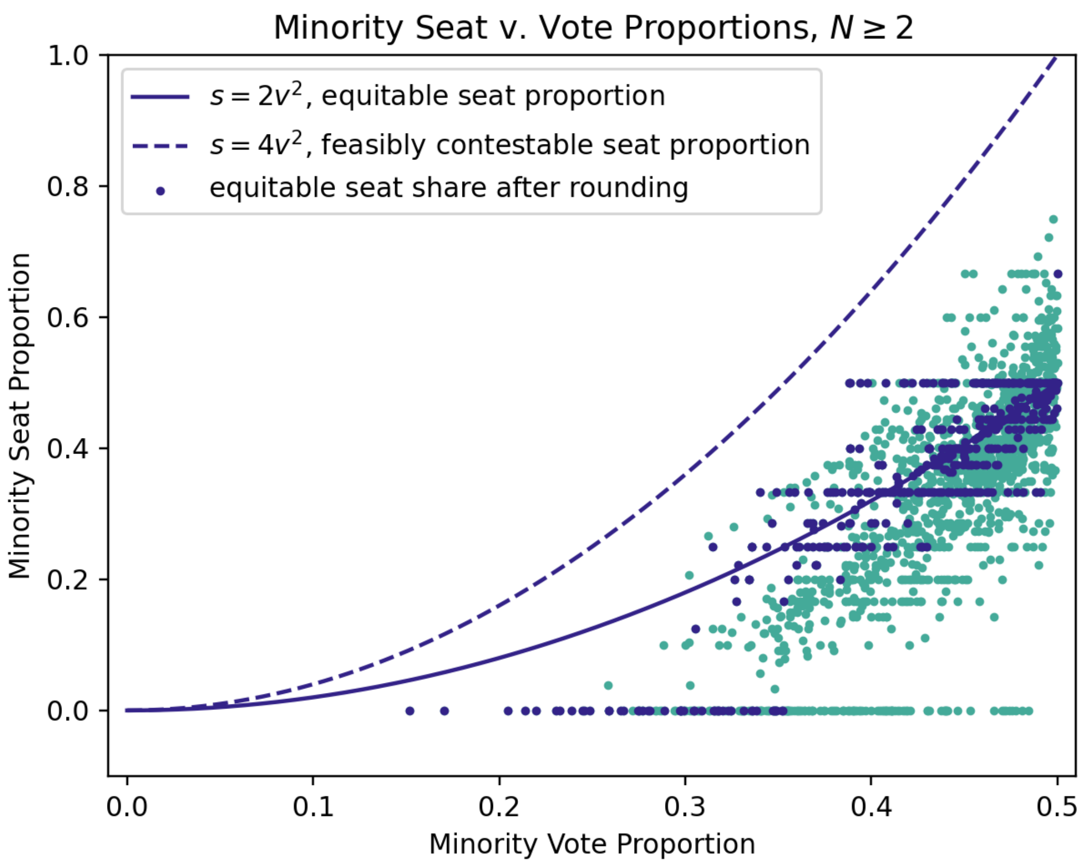


Figure 1: Historical seat vs. vote proportions for Warrington's database of elections 1972-2016 where $N \geq 2$. The proposed equitability standard appears to be a reasonable fit.

2.4 Discrepancy

Having determined the equitable seat proportion for the minority in a two-party election, we are now in a position to define the discrepancy, D, in a given map as the difference between the equitable and actual seat proportions with rounding taken into account.

With N districts, s_B will be equitable after rounding whenever

$$s_B - \tfrac{1}{2N} \leq s_B^* < s_B + \tfrac{1}{2N},$$

or equivalently, whenever

$$s_B^* - \tfrac{1}{2N} < s_B \leq s_B^* + \tfrac{1}{2N}.$$

In such a case we say that there is no discrepancy present, and we let $D = 0$.

If s_B lands outside of the interval $\left(s_B^* - \frac{1}{2N}, s_B^* + \frac{1}{2N}\right]$, then we say discrepancy is present, and we define the discrepancy to be the distance from s_B to the interval. Thus when there is discrepancy,

$$D = |s_B - s_B^*| - \tfrac{1}{2N}. \tag{4}$$

We can of course avoid the initial check for discrepancy by letting

$$D = \max\left\{|s_B - s_B^*| - \tfrac{1}{2N}, 0\right\}.$$

Defining D in this way guarantees that $D > 0$ implies at least a one-seat difference between S_B and the rounded S_B^*. It provides a correction for apparently large differences between s_B and s_B^* that may vanish after rounding with small values of N, and it also prevents small changes in s_B^* from causing large jumps in D. Whenever a party has $s > s^*$ we call that party the **advantaged party**, regardless of whether the difference is enough to cause a positive discrepancy.

If the majority is also the advantaged party, then we have

$$0 \leq D \leq 2v_B^2 - \tfrac{1}{2N}.$$

If the minority is the advantaged party, then as a theoretical maximum we have

$$0 \leq D \leq 2v_B - 2v_B^2 - \tfrac{1}{2N}.$$

However, if we consider the minority's largest seat share to be given by the feasibly contestable seats, then as a practical matter we get the same bound as for the majority:

$$0 \leq D \leq 4v_B^2 - 2v_B^2 - \tfrac{1}{2N} = 2v_B^2 - \tfrac{1}{2N}.$$

In all cases, the upperbound tends to $0.5 - \frac{1}{2N} < 0.5$ as $v_B \to 0.5$.

3 Extension to Multiparty Elections

While there are many methods for determining fair seat shares in two-party elections, the multiparty case is more difficult. In this section we show the natural extension of the differential equation derivation above to the case where there are multiple parties in a first-past-the-post election. We begin by extending our notation.

If there are K parties, $A_1, A_2, \ldots, A_K$, then we let v_{A_k}, s_{A_k}, and $s_{A_k}^*$ denote each party's vote share, seat share, and fair seat share, respectively, for $1 \leq k \leq K$. Our convention will be to order the parties by vote share so that

$$v_{A_1} \geq v_{A_2} \geq \cdots \geq v_{A_K}.$$

The same three assumptions from before guide our work here, though "largest party," "winning party," and "losing party" can all now be plural. The case of more than one largest party implies that all such parties have the same vote proportion, and as a result they split seats and the winners' share of the unnecessarily wasted votes evenly in a way that we specify below.

In a victory guaranteed by Assumption #2, the largest party's vote proportion must be at least v_{A_1} while the largest it can be is 1. We consider a fair initial allocation for the winning party to be one that splits the difference, i.e. a vote proportion of $\frac{1+v_{A_1}}{2} = 1 - \frac{1}{2}\sum_{k=2}^{K} v_{A_k}$. The corresponding equitable allocation for the losing parties is one that preserves the relationship among their vote proportions with each contributing $\frac{1}{2}v_{A_k}$ for $2 \leq k \leq K$.

As in the two-party case, we are led to a system of differential equations with initial solutions given by

$$
\begin{aligned}
v_{A_1}(x) &= 1 - \sum_{k=2}^{K} \frac{v_{A_k}}{\sqrt{1-x}} \\
v_{A_2}(x) &= \frac{v_{A_2}}{\sqrt{1-x}} \\
&\vdots \\
v_{A_K}(x) &= \frac{v_{A_K}}{\sqrt{1-x}}.
\end{aligned}
$$

When $J \geq 2$ parties are tied for largest, each party allocates $\frac{1}{J}\left(1 - \sum_{k=J+1}^{K} v_{A_k}(x)\right)$ as its share of the possible unnecessarily wasted votes.

Since the largest parties allocate more than their average vote proportion to seats they win, their vote proportions in the remaining seats decrease with x. Losing parties, on the other hand, allocate less than their average vote proportions to seats they must lose, and this causes their vote proportions to increase with x. The process terminates when the vote proportions of all parties coalesce to $\frac{1}{K}$, at which point by Assumption #1 each party deserves $\frac{1}{K}$ of the remaining seats. Following the steps from the two-party situation, it immediately follows that the smallest party, A_K, deserves seat proportion

$$s^*_{A_K} = K v^2_{A_K}. \tag{5}$$

Finding the remaining fair seat proportions is a straightforward matter of iteratively solving the equations above while keeping careful track of seats deserved:

1. Beginning with $v_{A_1}(x) \geq v_{A_2}(x)$, we allocate votes from all parties using the equations above. As $v_{A_1}(x)$ decreases, $v_{A_2}(x)$ increases until a point x_1 such that $v_{A_1}(x_1) = v_{A_2}(x_1)$. To find that point we solve the equation:

$$1 - \sum_{k=2}^{K} \frac{v_{A_k}}{\sqrt{1-x_1}} = \frac{v_{A_2}}{\sqrt{1-x_1}},$$

   which yields

$$x_1 = 1 - \left(v_{A_2} + \sum_{k=2}^{K} v_{A_k}\right)^2.$$

   On the interval $[0, x_1]$, Party A_1 is the largest party and so has sole claim to these x_1 seats.

2. Continuing from x_1, Parties A_1 and A_2 remain tied for the largest party until the point x_2 where $v_{A_1}(x_2) = v_{A_2}(x_2) = v_{A_3}(x_2)$. To find that point we solve the equation:

$$\frac{1}{2}\left(1 - \sum_{k=3}^{K} \frac{v_{A_k}}{\sqrt{1-x_2}}\right) = \frac{v_{A_3}}{\sqrt{1-x_2}},$$

which yields

$$x_2 = 1 - \left(2v_{A_3} + \sum_{k=3}^{K} v_{A_k}\right)^2.$$

From x_1 to x_2, Assumption #1 implies that Party A_1 and Party A_2 have an equal claim to those seats. Thus we adjust the seat totals of Party A_1 and Party A_2 by adding $\frac{1}{2}(x_2 - x_1)$ to each.

3. We continue the process until we solve for x_{K-1} such that

$$v_{A_1}(x_{K-1}) = v_{A_2}(x_{K-1}) = \cdots = v_{A_K}(x_{K-1}) = \frac{1}{K}.$$

At this point the process terminates. On the interval $[x_{K-2}, x_{K-1}]$, all parties up to Party A_{K-1}, which are all tied for the largest party, are entitled to $\frac{1}{K-1}(x_{K-1} - x_{K-2})$ seats.

4. Finally, on the interval $[x_{K-1}, 1]$, all K parties have the same vote proportions and hence have equal claim on the $(1 - x_{K-1}) = K^2 v_{A_K}^2$ seats remaining.

5. The end result once all x_k's are determined, is that the fair seat share for each party in an election with K parties is given by

$$\begin{aligned}
s^*_{A_1} &= x_1 + \frac{1}{2}(x_2 - x_1) + \frac{1}{3}(x_3 - x_2) + \cdots + \frac{1}{K-1}(x_{K-1} - x_{K-2}) + \frac{1}{K}(1 - x_{K-1}) \\
s^*_{A_2} &= \frac{1}{2}(x_2 - x_1) + \frac{1}{3}(x_3 - x_2) + \cdots + \frac{1}{K-1}(x_{K-1} - x_{K-2}) + \frac{1}{K}(1 - x_{K-1}) \\
&\vdots \\
s^*_{A_{K-1}} &= \frac{1}{K-1}(x_{K-1} - x_{K-2}) + \frac{1}{K}(1 - x_{K-1}) \\
s^*_{A_K} &= \frac{1}{K}(1 - x_{K-1}) = K v_{A_K}^2.
\end{aligned}$$

Note that it in cases where $v_{A_k} = v_{A_{k+1}}$, we get $x_{k-1} = x_k$ so some terms in the sums above can be 0. Additionally, it is evident that

$$\sum_{k=1}^{K} s^*_{A_k} = 1.$$

The construction also guarantees that the fair seat share for a party is unaffected by the addition of parties with zero votes.

Assumption #3 requires that the winning parties split the unnecessarily wasted votes evenly with the losing parties in seats that are not in dispute. It is instructive to consider the implications for two alternate extremes for this assumption. First, if we require losing parties to waste 100% of the unnecessarily wasted votes, we would end up with maximal cracking, resulting in a sweep by the winning parties with each party contributing its overall vote proportion to each district. Second, if we require winning parties to shoulder 100% of the unnecessarily wasted votes, then we end up with maximal packing of the largest parties into seats not in dispute, resulting in proportional representation for all parties, $s^*_{A_k} = v_{A_k}$. We note that we could replace the 50% requirement with a parameter, ρ, that controls the proportion of unnecessarily wasted votes allocated by the losing parties. Doing so would only impact the derivation minimally, and would allow straightforward justifications for the claims above where $\rho = 1$ and $\rho = 0$.

Thus proportional representation serves as a floor for the largest parties, and it is only achievable under the clearly unreasonable assumption that losing parties never waste any votes in seats they cannot win. We are led to conclude that plurality voting systems, even before taking into account the effects of geography or potential gerrymandering, are inherently biased in favor of the largest parties winning more than their proportional share of votes. This lends further credence to arguments that proportional representation elections are inherently fairer to small parties.

3.1 Examples

We illustrate the mechanics of the process and show that it returns sensible, intuitive results in the following examples.

Example 1. *For the 2019 Canadian parliamentary election, we consider the results from British Columbia [3]. There were 42 seats for BC and five major parties with vote proportions given by*

$$(Cons., Lib., NDP, Green, People's, Other) = (0.340, 0.262, 0.244, 0.125, 0.017, 0.012).$$

Thus our initial system of equations is

$$\begin{aligned}
v_{A_1}(x) &= 1 - \frac{0.660}{\sqrt{1-x}} \\
v_{A_2}(x) &= \frac{0.262}{\sqrt{1-x}} \\
v_{A_3}(x) &= \frac{0.244}{\sqrt{1-x}} \\
v_{A_4}(x) &= \frac{0.125}{\sqrt{1-x}} \\
v_{A_5}(x) &= \frac{0.017}{\sqrt{1-x}} \\
v_{A_6}(x) &= \frac{0.012}{\sqrt{1-x}}.
\end{aligned}$$

With x_1 such that $v_{A_1}(x_1) = v_{A_2}(x_1)$, we get $x_1 = 1 - (0.262 + 0.60)^2 \approx 0.150$. This represents the proportion of seats to which the Conservative Party has sole claim. Beginning at x_1 the Conservative Party and Liberal Party have the same proportion of the vote on the *remaining* $1 - x_1$ *seats. Thus from this point forward, the Conservatives and the Liberals deserve the same proportion of seats. Also, beginning at* x_1*, we have*

$$v_{A_1}(x) = v_{A_2}(x) = \frac{1}{2}\left(1 - \sum_{k=3}^{6} \frac{v_{A_k}}{\sqrt{1-x}}\right) = \frac{1}{2}\left(1 - \frac{0.398}{\sqrt{1-x}}\right).$$

To find x_2*, we set* $v_{A_1}(x_2) = v_{A_2}(x_2) = v_{A_3}(x_2)$ *to get*

$$\frac{1}{2}\left(1 - \frac{0.398}{\sqrt{1-x_2}}\right) = \frac{0.244}{\sqrt{1-x_2}}.$$

Solving yields $x_2 = 1-(2\cdot 0.244+0.398)^2 \approx 0.215$*. The interval from* x_1 *to* x_2 *represents the proportion of seats over which the Conservatives and Liberals were tied for the largest party and hence have equal claim. Thus each party deserves half of the interval* $(0.150, 0.215)$*, or 0.0325 seats each.*

Repeating the process produces $x_3 = 1 - (3 \cdot 0.125 + 0.154)^2 \approx 0.720$*,* $x_4 = 0.991$*, and* $x_5 = 0.995$*. After* x_5*, all parties share the same vote proportion,* $\frac{1}{6}$*, and hence each deserves* $\frac{1}{6}$ *of the seats on the interval* $[x_5, 1]$*. All that is left to do is sum the deserved seat proportions:*

$$\begin{aligned}
s^*_{A_1} &= x_1 + \frac{1}{2}(x_2 - x_1) + \frac{1}{3}(x_3 - x_2) + \cdots + \frac{1}{K-1}(x_{K-1} - x_{K-2}) + \frac{1}{K}(1 - x_{K-1}) \\
&= 0.150 + \frac{1}{2}0.065 + \frac{1}{3}0.505 + \frac{1}{4}0.271 + \frac{1}{5}0.004 + \frac{1}{6}0.005 \\
&= 0.420 \\
s^*_{A_2} &= \frac{1}{2}0.065 + \frac{1}{3}0.505 + \frac{1}{4}0.271 + \frac{1}{5}0.004 + \frac{1}{6}0.005 \\
&= 0.270 \\
s^*_{A_3} &= \frac{1}{3}0.505 + \frac{1}{4}0.271 + \frac{1}{5}0.004 + \frac{1}{6}0.005 \\
&= 0.238 \\
s^*_{A_4} &= \frac{1}{4}0.271 + \frac{1}{5}0.004 + \frac{1}{6}0.005 \\
&= 0.069 \\
s^*_{A_5} &= \frac{1}{5}0.004 + \frac{1}{6}0.005 \\
&= 0.002 \\
s^*_{A_6} &= \frac{1}{6}0.005 \\
&= 0.001
\end{aligned}$$

As there were 42 total seats in British Columbia, our equitable distribution of seats is

$$(Cons., Lib., NDP, Green, People's, Other) = (17.6, 11.4, 10.0, 2.9, 0.1, 0.0).$$

The actual seat shares were

$$(Cons., Lib., NDP, Green, People's, Other) = (17, 11, 11, 2, 0, 1),$$

an indication that the method returns sensible results. Results, including comparisons with proportionality, for all ten provinces that had more than one seat for 2019 are presented in the appendix.

The next examples show that the new measure is responsive not only to a party's vote proportion, but to its relationship to other parties' vote proportions as well.

Example 2. *As the largest party's advantage over the next largest party grows, so does its fair seat proportion. The three tables below show the results of the new construction for progressively less competitive elections where the largest party has vote share 0.55.*

Party:	A_1	A_2	A_3	A_4
Vote Prop.	0.55	0.43	0.01	0.01
Fair Seat Prop.	0.6124	0.3868	0.0004	0.0004

Table 1: Results for an election with two competitive parties.

Party:	A_1	A_2	A_3	A_4
Vote Prop.	0.55	0.35	0.05	0.05
Fair Seat Prop.	0.670	0.310	0.010	0.010

Table 2: Results with less competitive, second largest party.

Party:	A_1	A_2	A_3	A_4
Vote Prop.	0.55	0.15	0.15	0.15
Fair Seat Prop.	0.730	0.090	0.090	0.090

Table 3: Results with three uncompetitive parties.

Example 3. *The next three tables show that as the vote shares of all parties approach parity, the fair seat proportions do as well.*

Party:	A_1	A_2	A_3	A_4
Vote Prop.	0.40	0.30	0.20	0.10
Fair Seat Prop.	0.500	0.310	0.150	0.040

Table 4: Results with evenly spaced, moderately competitive parties.

Party:	A_1	A_2	A_3	A_4
Vote Prop.	0.300	0.267	0.233	0.200
Fair Seat Prop.	0.344	0.280	0.216	0.160

Table 5: Results for evenly spaced, more competitive parties.

Party:	A_1	A_2	A_3	A_4
Vote Prop.	0.25	0.25	0.25	0.25
Fair Seat Prop.	0.250	0.250	0.250	0.250

Table 6: Results for perfectly competitive parties.

3.2 Comparison to Canadian Parliamentary Elections

In this section we present evidence that our multiparty construction for fairness produces sensible results when applied to real elections. The data we use is taken from the Elections Canada website [3], and it includes all Canadian Parliamentary elections from 1997 to 2019. We focus on results by province, and we restrict our attention to provinces that awarded more than one seat.

In the two-party case it was easy to see from a simple scatterplot that the equitable seat share curve captured the relationship seen in real election data reasonably well. It is also apparent from that graph that proportionality is not a reasonable expectation of maps in the two-party case. In the multiparty case we do not have a simple way to picture the relationship between vote proportions and fair seat proportions, because the multiparty case depends not just on a party's vote proportion but also its relationship to those of the other parties.

We can, however, still provide some evidence that proportionality is not to be expected in the multiparty case and that our new standard for equitability is clearly a better baseline for fairness. We focus our attention on Canadian parliamentary election results. These are all multiparty elections with district, or riding, boundaries drawn by independent commissions[3]**. Thus we expect these maps to be reasonably free from gerrymandering, though geographical considerations may still be important.

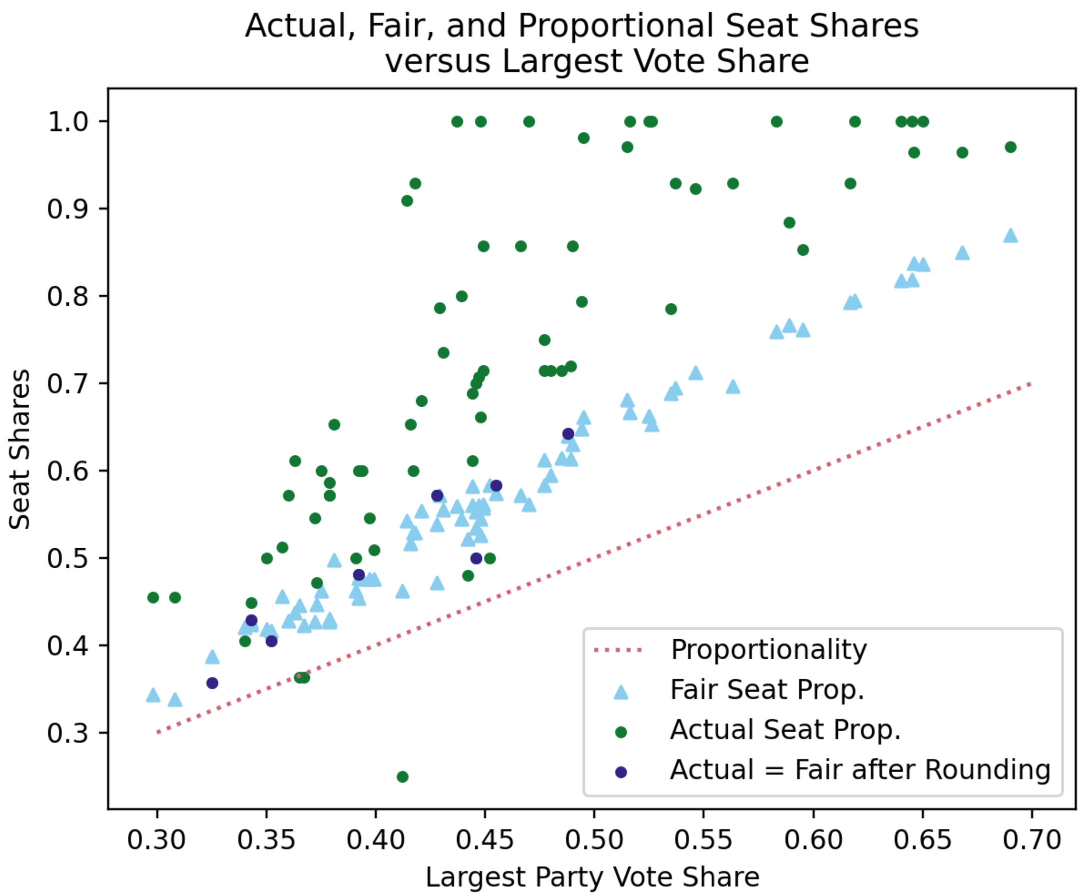


Figure 2: Actual, fair, and proportional seat shares by province versus the largest party's vote share for Canadian Parliamentary elections 1997-2019. Proportions are not rounded to the nearest seat. The dark dots are seat shares that would be fair after rounding. Sweeps by a single party are shown across the top, 6 of which were elections for Prince Edward Island that had 4 seats.

In Figure 2 we present a scatterplot of the seat share earned by each province's largest party for the Canadian Parliamentary elections 1997-2019. The plot shows a similar result to that in Figure 1. The new standard is clearly a better fit for actual election data than proportionality for the largest party. Of the 80 elections presented, there are five where the largest party's seat share was smaller than its fair share and eight where the largest party's seat share was equal to its fair share. As we should expect for a baseline of fairness for ideal elections, most of the elections have the largest party winning more than its fair share of seats. The obvious outlier is for the 2011 election in Prince Edward Island, where the Conservative Party won 41.2% of the vote but only 1 of 4 seats, while the Liberal party won 41% of the vote and 3 of the 4 seats. For small values of the maximum vote share, we note that the fair seat proportion gets closer to proportionality. This is to be expected as elections with small vote shares for the largest party will tend to have more evenly matched competitors.

Beyond just the largest party, the new standard is also demonstrably better as a baseline of fairness in an overall sense for these 80 elections. For each party in each election we compute the discrepancy as defined in Section 2.4 using both proportionality and the new standard for fairness. We then compare the average discrepancy for each election by averaging over the number of relevant parties in the province. In the elections considered here, the smallest party to win a seat had an overall vote share of 0.008 in the province. Thus, we consider a party to be relevant if its vote share in the province is at least 0.007. As in the two-party case we are not trying to find the best fit for the election data; however, we are claiming that our baseline standard for fairness is and must be a better fit than proportionality even if boundaries are drawn by an independent, well-intentioned commission.

Figure 3 compares the distributions of average discrepancy when using proportionality as the baseline for fairness versus using the new standard. It is clear that the new standard is a better overall fit. For example, the median and first quartile for the new standard are less than half the values for proportionality.

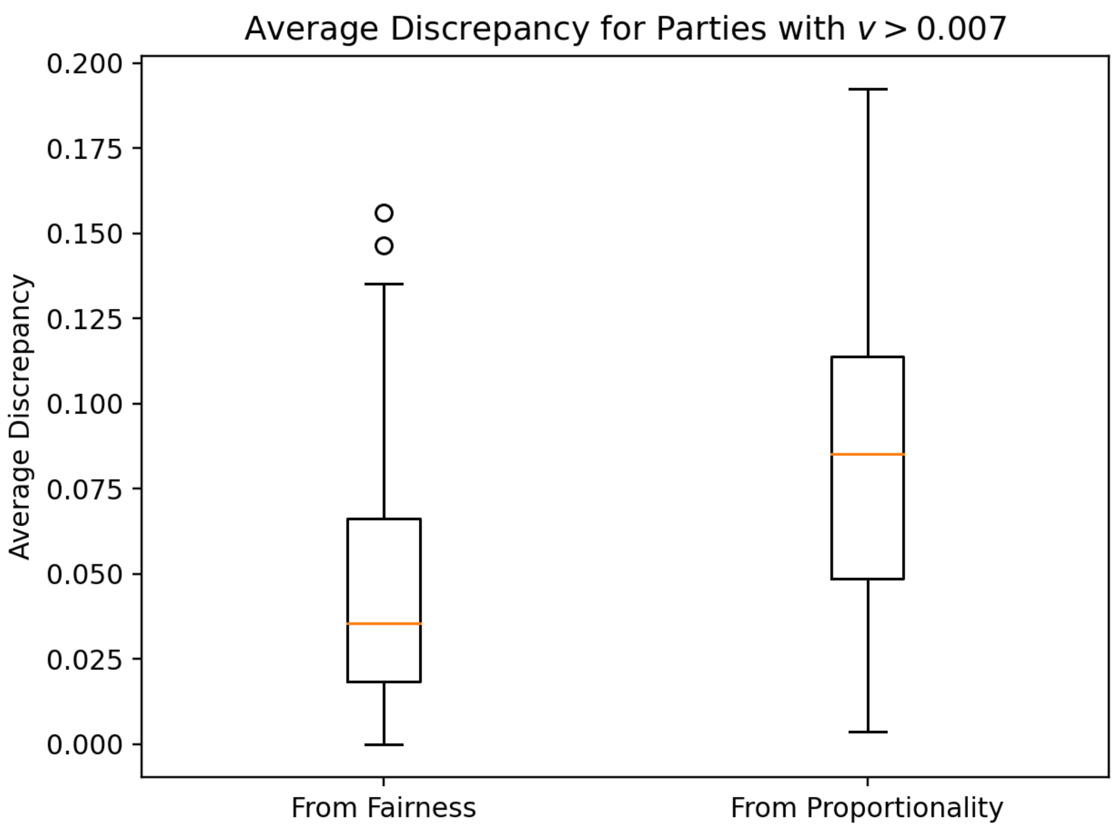


Figure 3: A comparison by of the distributions of average discrepancy between the new standard and that implied by proportionality. For each province, the average is taken over parties with at least 0.7% of the overall vote.

Figure 4 presents a scatterplot for the average discrepancy from fairness versus the average discrepancy from proportionality. The dotted line shows equality. Again, we see the same result. The new standard was a better fit for 95% of the elections, and in the 5% where proportionality was better, the values were very close.

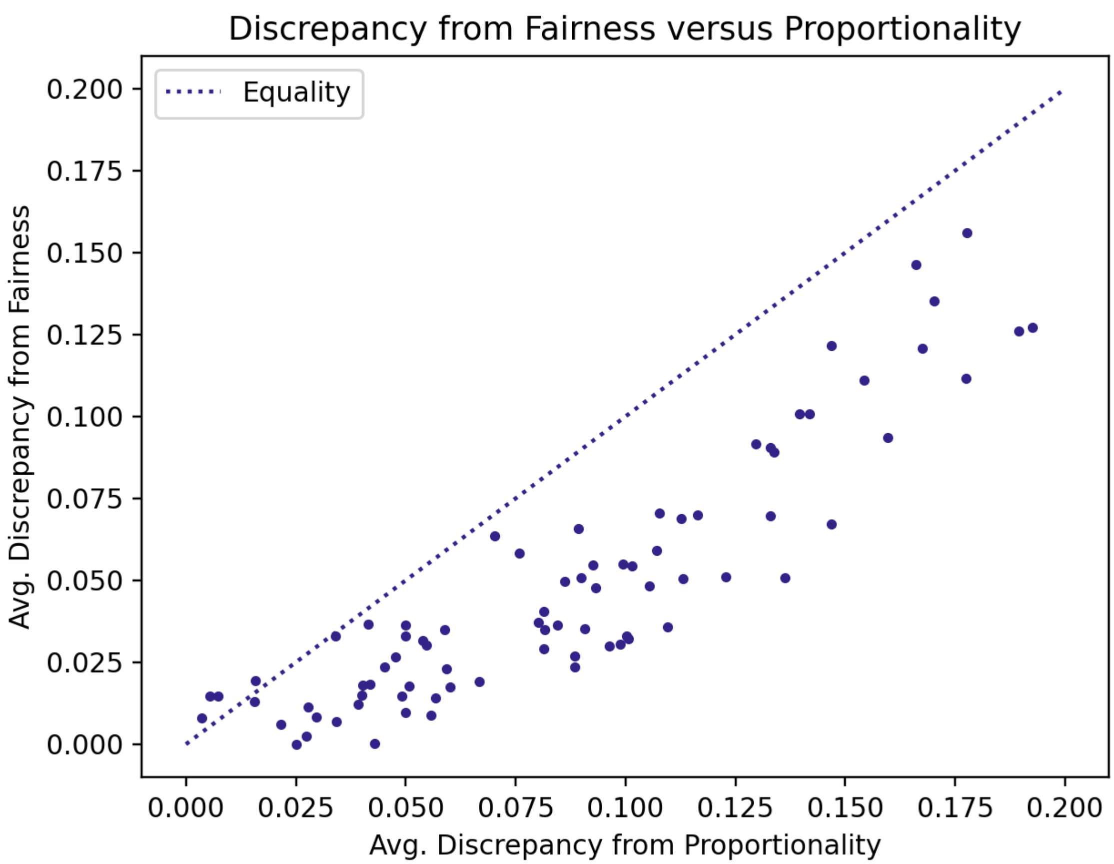


Figure 4: A scatterplot of average discrepancy based on the new standard and based on proportionality. The average is taken over parties with at least 0.7% of the overall vote.

Though our stated intent is to produce a reasonable baseline of fairness for multiparty elections and not necessarily a "best fit" to the election data, the new construction can be used to provide promising results to that end. We do so by adjusting Assumption #3 to include the parameter ρ that determines the proportion of the unnecessarily wasted votes expended by the losing parties in seats they cannot win. Details are included in the appendix.

4 Implications for Detecting Partisan Gerrymandering

A partisan gerrymander occurs when one party draws district boundaries that dilute the influence of the other party's votes through *packing*, the concentration of voters into few districts, and *cracking*, the dispersal of voters across many districts. In what follows we propose a method for detecting problematic partisan gerrymanders in the two-party case. Our method requires a map to satisfy two criteria in order to be flagged as problematic: not only must the seat distribution be unfair, but that unfairness must be likely to persist in the future. Having defined discrepancy as our measure for unfairness for seat distribution, we now provide a measure for the stability of that unfairness that is intuitive, easily computed, and provides sound results both in theoretical examples and in actual elections.

4.1 Seat Lean

We are not just interested in the size of the discrepancy in a district plan but also in the prospect of the discrepancy persisting in the next election. Our approach in this section is to start by providing a measure for the degree to which a map favors one party over another for the next election. We call this measure the **seat lean** of the map, and it is based on the likelihood of seats changing hands in the next election given each district's losing vote proportion in the current one.

We start by using Warrington's database of federal elections [5] to estimate the historical likelihood of a U.S. congressional seat flipping to the other party given the vote proportion of the loser in the current election. Out of the 1142 federal elections in the database, we consider only those for which the next election is under the same district plan as the current one. Thus we remove all elections from years 1980, 1990, 2000, 2010, and 2016 from consideration.

Among the remaining elections, there are 7823 district elections to consider. Out of these, 568 flipped to the other party in the following election for a proportion of about 0.073. As one would expect, if we restrict our attention to districts where the minority represents a larger proportion of the vote, we get a higher proportion of flips. For example, there were 198 district elections where the loser's vote proportion was between 0.47 and 0.48. Of these, 39 flipped in the next election for a proportion of about 0.20.

Figure 5 shows a plot of the proportion of seats in the database that have flipped for a given minority vote proportion, p, where the flip proportion is calculated over the interval $(p - 0.01, p + 0.01)$. Included on the plot is the graph of the function $y = 4p^4$.

The function $y = 4p^4 = (2p^2)^2$ provides a good, though not necessarily best, fit for the proportions derived from the data set. This is the function we will use to define the seat lean of a map. It provides a simple and reasonably accurate estimate for how often we expect a district to change hands given the district's minority vote proportion.

It is not important for our purposes to find the best approximation for the proportion of seat flips. Even if it were, there is a limit to how precise we can expect to be due to the limitations inherent in measuring vote proportions. As Warrington notes in [6], many district elections involve third party candidates or perhaps are uncontested, and in these cases the vote proportions in the data set have been imputed to estimate the two-party vote split. What is important here is that our approximation makes the proper distinctions among certain kinds of packing and cracking, and we will see in a moment that this is the case for our chosen function.

Thinking of a flip proportion as roughly an expected seat gain, we now define the seat lean, L, for a map. Party A loses seats C_n, $1 \le n \le S_B$, so from those seats in the next election Party A would expect to gain an estimated $\sum_{n=1}^{S_B} 4p_n^4$ seats. Similarly, Party B would expect to gain approximately $\sum_{n=S_B+1}^{N} 4q_n^4$ seats from its losses. With Party A as the advantaged party, we define the seat lean, L, to be the approximate expected net change in seat proportion. We have

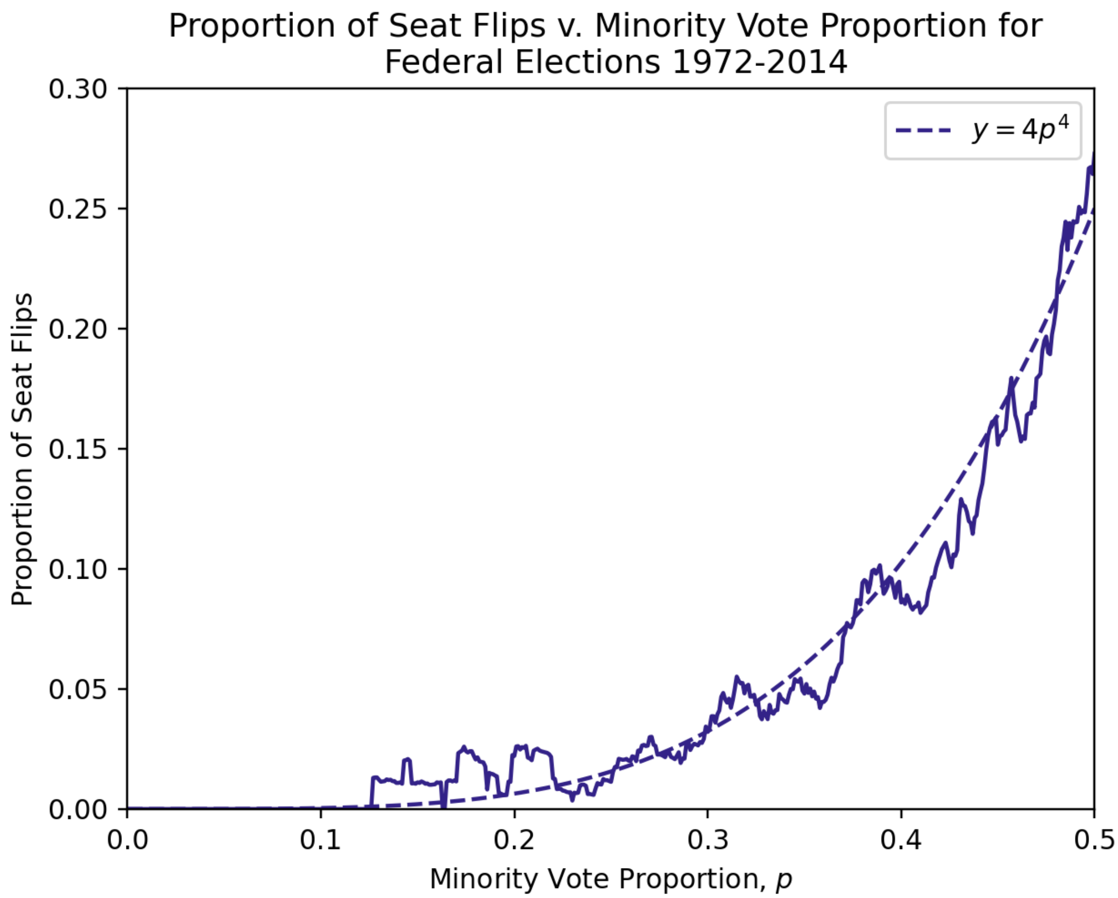


Figure 5: Proportion of seats with minority vote proportion in $(p - 0.01, p + 0.01)$ that changed hands in the next election. The dashed line approximates the relationship with the simple function, $f(p) = 4p^4$.

$$L = \frac{1}{N}\left\{\sum_{n=1}^{S_B} 4p_n^4 - \sum_{n=S_B+1}^{N} 4q_n^4\right\}. \tag{6}$$

The summations would be swapped if Party B had the advantage.

Positive values for L indicate that the advantaged party is poised to increase its advantage in the coming election, while negative values indicate that the disadvantaged party is poised to make up ground. Later we will suggest a threshold so that for a given level of discrepancy, if the disadvantaged party expects to make up *enough* ground in the next election, a map will not be considered problematic.

A crucial feature of the measure L is its ability to distinguish between maps like the following.

Example 4. *Consider an election where $v_B = 0.25$, $N = 2$, and the minority party is swept so $s_B = 0$. Map 1 is given by $(0.25, 0.25)$, while Map 2 is given by $(0.50^-, 0)$. The seat lean in Map 1 is $L = \frac{1}{2}\{0 - 2 \cdot 4(.25)^4\} \approx -0.016$. On the other hand, the seat lean in Map 2* *is* $L = \frac{1}{2}\{0 - (4(.50)^4 + 4(0)^4)\} \approx -0.125$. *The seat lean is much more negative for Map 2. Thus the measure recognizes Map 2 as superior for Party B since Party B is in a much better position to win back a seat in the next election.*

The ability to make the distinction in the previous example is the crucial feature that we require of our function for assessing relative seat advantage for the next election. A measure such as the "average margin of victory," for example, would be unable to make this distinction – both maps have an average margin of victory of 0.5. The example also illustrates a general principle that the measure of seat lean makes clear: *either party should always prefer packing to cracking in the seats that it loses.*

An advantage of this measure for seat lean is that it should be relatively resistant to changes in the way vote shares are imputed, particularly for uncontested elections, because small vote proportions contribute very little to the overall measure.

4.1.1 Relationship between Discrepancy and Seat Lean

Generally speaking, as discrepancy increases we should expect the seat lean to become more favorable to the disadvantaged party. This is because if the disadvantaged party wins fewer seats with the same vote proportion or if it wins the same number of seats with a larger vote proportion, then it will have more votes to allocate to losing seats and perhaps more opportunities to win seats. We see this relationship borne out in the data.

The scatter plot in Figure 6 presents the relationship between D and L for the federal elections ($N \geq 2$) in the Warrington database. Recent elections of interest (from 2012-16 in Massachusetts, Maryland, Michigan, North Carolina, Ohio, and Pennsylvania) are indicated by the dark dots.

Intuitively, the most problematic gerrymanders should be those with large discrepancy that are located near the top of the plot. These are maps where one party not only has a large seats advantage but also relatively strong prospects for the following election. Note that as we might expect, the recent elections of interest appear to be by and large the kinds of maps that should be problematic.

The relationship shown in Figure 6 points to an inherent difficulty in distinguishing problematic maps from those that are just inequitable. There is an element of "self-correction" inherent in an inequitable map: the higher the discrepancy, the better the prospects are for the disadvantaged party in the next election. We want to measure the degree to which an inequitable map is resistant to this self-correcting tendency.

We can get a sense for how changes in discrepancy might be related to changes in seat lean by considering the effect that flipping a seat would have on both. With all else being held constant, consider a map where one seat flips from a loss for the disadvantaged party at 0.50^- to a win at 0.50^+. Since v_B remains unchanged, only the net change in seat proportion effects the discrepancy, which decreases by $\frac{1}{N}$. Seat lean on the other hand increases by $\frac{1}{2N}$. This suggests $L = -0.5D$ as a candidate for the downward trend we see in Figure 6. In Figure 7 we add the line $L = -0.5D$ to the scatter plot of seat lean v. discrepancy and note that it does in fact do a good job of capturing the downward trend seen in the data.

We can remove most of the influence of D on L by subtracting the values given by $L = -0.5D$ from observed values in the data set. This is equivalent to taking $L + 0.5D$,

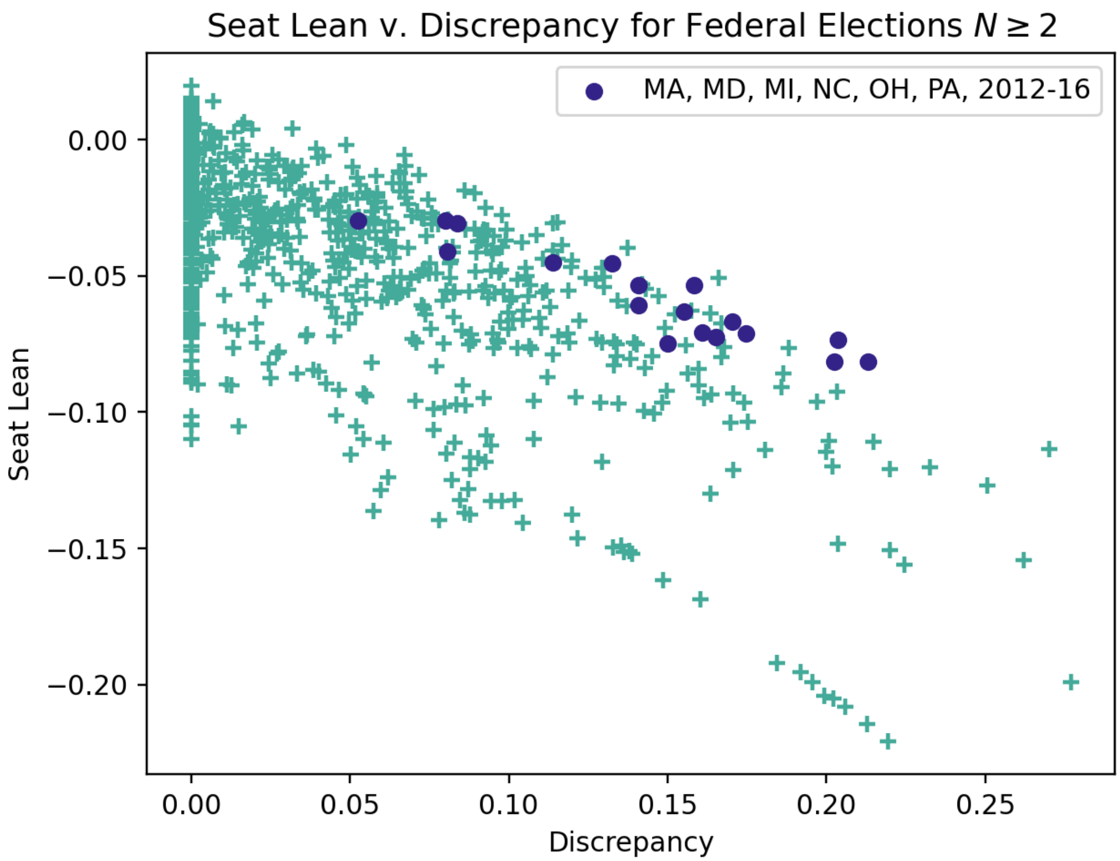


Figure 6: Relationship between discrepancy and seat lean with recent elections of interest highlighted. As discrepancy increases, prospects become more favorable for the disadvantaged party for the next election.

a quantity we will refer to as the **adjusted seat lean**, denoted AL. The results of this adjustment are shown in Figure 8. The correlation in the plot is 0.035.

4.1.2 A Standard for Problematic Seat Lean

The adjusted seat lean, $AL = L + 0.5D$, allows us to make fair comparisons of seat lean across different values of discrepancy. In this section we suggest that $AL = 0$ provides a reasonable cutoff for what should constitute a problematic adjusted seat lean for a given inequitable map. Thus any inequitable map whose seat lean places it on or above the line $L = -0.5D$ in Figure 7 will be considered to have a problematic lean. Equivalently, *maps where $AL \geq 0$ will have a problematic lean, while those for which $AL < 0$ will not be flagged as such.*

Though it is not justified to formally interpret a map's seat lean as its expected value for net change in seat proportion, it is perhaps helpful to think of our cutoff along those lines: roughly speaking, the cutoff represents maps where the disadvantaged party is in a position to reclaim on average half of the discrepancy in the next election. We provide both

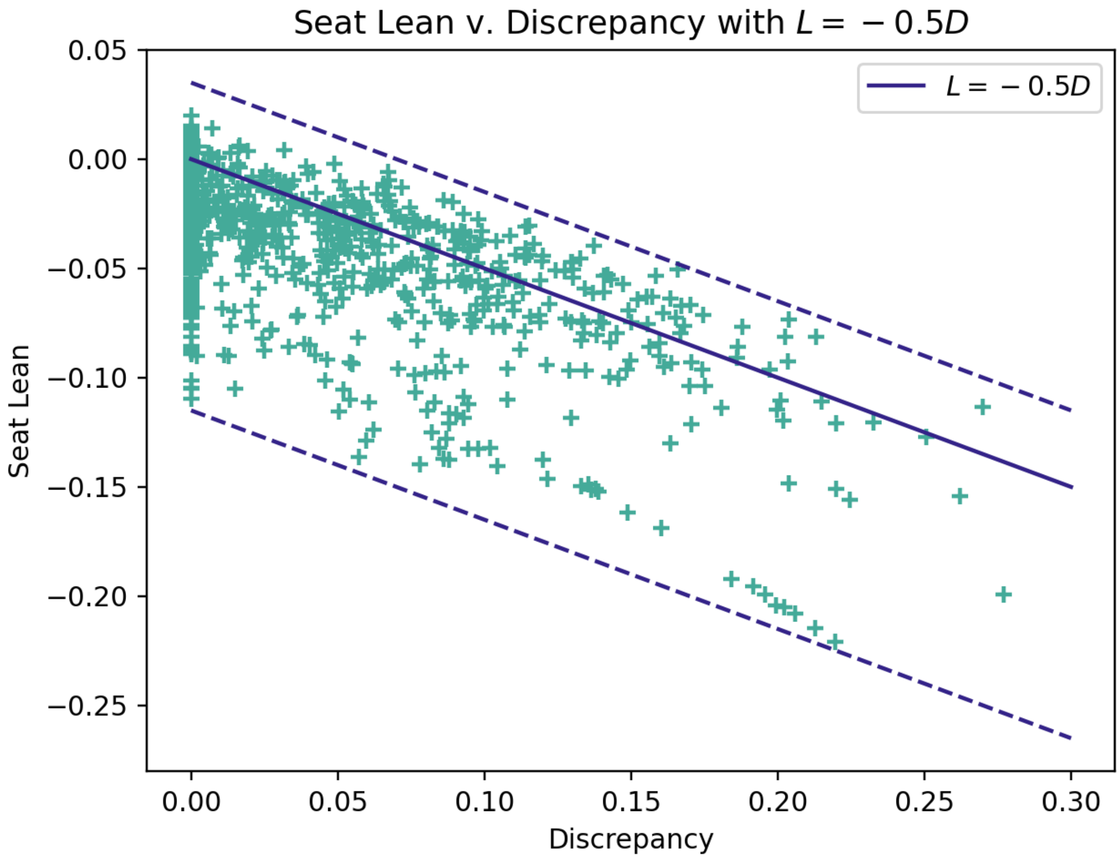


Figure 7: The line $L = -0.5D$ captures the relationship between D and L reasonably well.

theoretical and empirical justification for this cutoff below.

4.1.3 Edge Cases and Proportionality

In this section we apply the standard for seat lean to a few theoretical examples. In all cases, our standard returns a verdict that is consistent with our intuition. First we investigate a maximally competitive map.

Example 5. *Suppose we have a map where the statewide vote is evenly split so that $v_B = 0.5$, and every district is maximally competitive. Without knowing how many seats each party wins, such a map can be represented by*

$$(0.5^+, \ldots, 0.5^+, 0.5^-, \ldots, 0.5^-).$$

Assume Party B is not the advantaged party so that $s_B \leq 0.5$. The discrepancy in such a map of course depends on s_B and is given by $D = .5 - s_B - \frac{1}{2N}$; the seat lean is given by $L = \frac{1}{4}(s_B - s_A)$. Then

$$AL = .25 - .5s_B - \frac{1}{4N} + .25s_B - .25s_A = .25 - .25(s_A + s_B) - \frac{1}{4N} = -\frac{1}{4N} < 0.$$

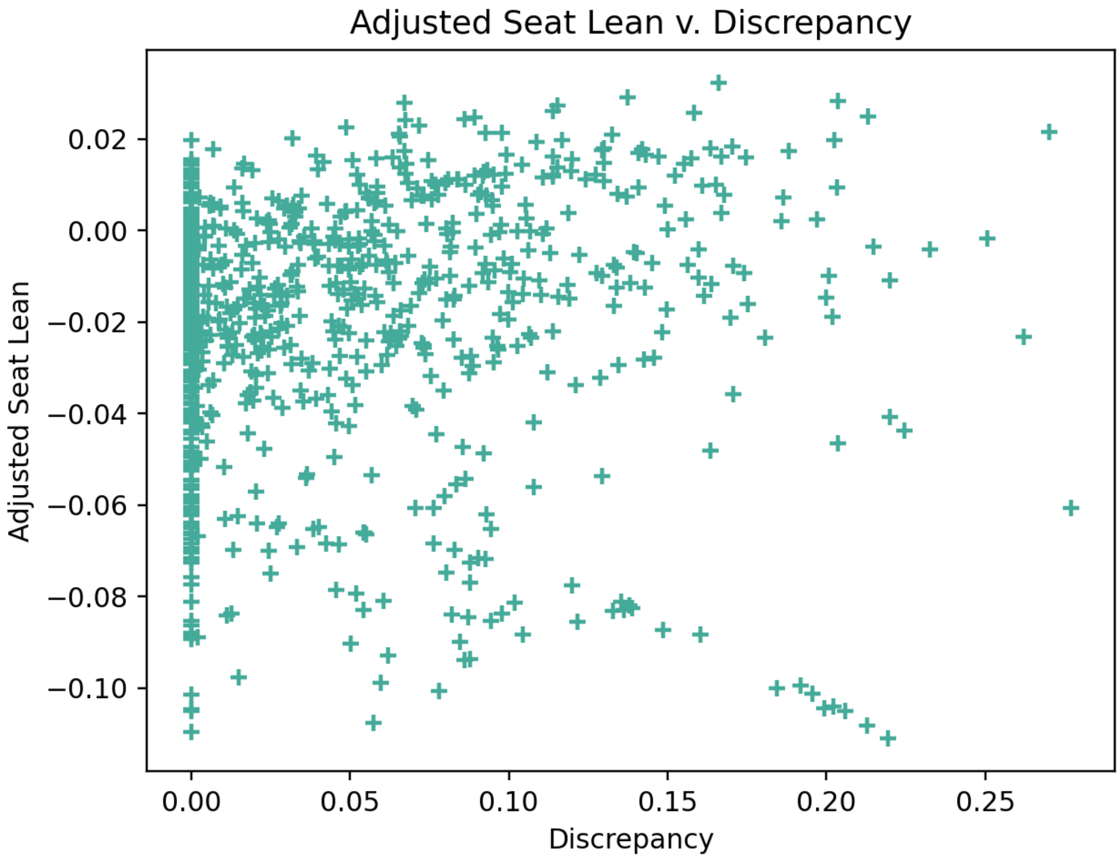


Figure 8: Relationship between discrepancy and adjusted seat lean. The correlation is 0.035.

This shows that the adjusted seat lean would be indifferent to the actual winners in maximally competitive districts (as it should be) and that such a map would never be flagged as having a problematic lean, regardless of the size of discrepancy present.

In fact, the way we define the adjusted seat lean guarantees the result of the previous example, because the adjustment calibrates the measure so that it is indifferent to the winner of a maximally competitive district. To reiterate a previous point, the particular function we use to define seat lean does not have to be the best possible fit to our data – any reasonable function we used could be calibrated in the same way.

Our next example takes a look at sweeps by the majority where the minority's vote is maximally cracked. This situation represents maximal discrepancy in favor of the majority while also arranging for the best possible seat lean for the majority among all sweeps. We show that for most cases this arrangement will be flagged as problematic; however, once the minority's statewide vote share is large enough, its prospects for the next election will be sufficiently favorable that we will no longer flag such a map as problematic.

Example 6. *Consider the case where $N = 10$, the majority sweeps the minority, and the minority's votes are maximally cracked. With v_B as the minority vote share, such a map* *is given by* $(v_B, v_B, \ldots, v_B)$*. The discrepancy in the map is* $D = 2v_B^2 - \frac{1}{20}$*, the seat lean is* $L = -4v_B^4$*, and* $AL = v_B^2 - 4v_B^4 - \frac{1}{40}$*, which is only greater than 0 when* v_B *is roughly between 0.17 and 0.47. Thus even the map with the largest discrepancy and the most favorable seat lean for the majority is not problematic when* $v_B > 0.47$*, since in that case the minority is positioned to gain sufficient seats in the next election.*

The previous example used the case $N = 10$, but the idea holds for all N. However, as N increases, so does the cutoff for v_B at which we no longer view the map as problematic. When $N = 20$, the cutoff is $v_B = 0.486$, while if $N = 50$, the cutoff is $v_B = 0.494$. We can make sense of this relationship by noting that as the number of districts increases, a clean sweep by the majority results in higher discrepancy so the minority must have more competitive losses to overcome it. As a practical matter, it seems correct to suggest that sweeps with more and more districts should be viewed with increasing suspicion.

Our next example shows that a sweep by the majority is never necessarily problematic regardless of the minority's vote proportion, v_B.

Example 7. *Consider the case where the majority sweeps the minority, and the minority's vote proportion is* v_B*. Assume that the minority's votes have been maximally packed in its losing seats. Such a map is given by*

$$(0.50^-, \ldots, 0.50^-, 0, \ldots, 0).$$

where the proportion of maximally competitive seats is $2v_B$*. The discrepancy in the map is* $D = 2v_B^2 - \frac{1}{2N}$*, the seat lean is* $L = -0.5v_B$*, and the adjusted seat lean is* $AL = v_B^2 - 0.5v_B - \frac{1}{4N}$*, which is always less than 0 since* $v_B^2 \leq .5v_B$*. Thus regardless of the value for* v_B*, a majority sweep may not be flagged as problematic if the minority's votes are sufficiently packed in its losing seats.*

The results in the previous two examples illustrate a necessary feature for any measure of partisan gerrymandering: such a measure must have the ability to distinguish problematic sweeps from those that are easily reversible. We further note that the efficiency gap and modified efficiency gap do not possess this ability, nor does the declination [6], which is undefined on sweeps.

We get a result similar to that in Example 6 when we consider the most inequitable map in favor of the minority party.

Example 8. *Consider the case where* $N = 10$*, and the minority wins its maximum number of seats. With* v_B *as the minority vote share, such a map is given by* $(0.5^+, 0.5^+, \ldots, 0.5^+, 0, \ldots, 0)$*, where the number of seats won by the minority is* $2Nv_B$*. The discrepancy in the map is* $D = 2v_B - 2v_B^2 - \frac{1}{20}$*, the seat lean is* $L = -\frac{v_B}{2}$*, and* $AL = \frac{v_B}{2} - v_B^2 - \frac{1}{40}$*, which is greater than 0 when* v_B *is roughly between 0.06 and 0.44. Thus the most inequitable map in favor of the minority is still not problematic when* $v_B > 0.44$*, since past that point the seat lean is sufficiently in favor of the majority.*

Under the current standard, proportionality is neither required nor necessarily disallowed. As the next two examples show, it depends on the packing and cracking among each party's losing seats.

Example 9. *Consider a map where* $N = 10$*,* $v_B = 0.3$*, and both parties are maximally packed. Such a map is given by* $(1, 1, 1, 0, 0, 0, 0, 0, 0, 0)$ *with discrepancy* $D = 0.3 - 0.18 - 0.05 = 0.07$*, seat lean* $L = 0$*, and* $AL = 0.035$*. Since* $AL > 0$*, this map's seat lean would be flagged as problematic. The majority party is at a disadvantage with very little prospect of overcoming that discrepancy in the next election.*

Example 10. *Next consider the case where* $N = 10$ *and* $v_B = 0.3$*, but now the map is given by* $(0.5^+, 0.5^+, 0.5^+, 0.214, 0.214, \ldots, 0.214)$*. Here the discrepancy is the same,* $D = 0.07$ *as the previous map, but the seat lean is given by* $L = 0.00587 - 0.075 = -0.069$ *and* $AL = -0.034$*. The minority party barely wins its seats, but the majority party wins its seats comfortably. This is reflected in the measure* $AL = -0.034$*, which indicates that the disadvantaged majority party is in a favorable position to turn things around in the next election. For this map, proportionality would not be flagged as problematic.*

4.1.4 Adjusted Seat Lean and Persistence of Discrepancy

Having verified that our proposed standard for seat lean produces satisfactory results on important theoretical examples, we now apply our standard to Warrington's database of federal elections. Our results with real elections will support our contention that when discrepancy is present, maps with higher adjusted seat lean should be viewed as more problematic as they are much more likely to have their discrepancy persist in the next election. We also show that this is a property not shared by discrepancy itself.

We consider the collection of maps in our database where discrepancy is present that occur during years where the same map is in effect for the subsequent election. Thus we again exclude the years 1980, 1990, 2000, 2010, and 2016, ending up with 440 eligible maps. For each of these maps we record whether the discrepancy persisted at some level in the following election. If the discrepancy is to the same party's advantage in the next election we say that the discrepancy persists whereas if the discrepancy goes to 0 or turns in favor of the other party, we say the discrepancy did not persist. Of the 440 maps, the discrepancy persisted in 272 of them for a proportion of about 0.62.

Figure 9 shows a scatter plot of adjusted seat lean versus discrepancy for our subset of 440 elections. Maps for which the discrepancy persisted in the next election are shown by the aqua pluses. Maps for which the discrepancy was relinquished are the dark dots.

It is visually clear that elections for which $AL \geq 0$ have a much higher proportion of discrepancy persistence than those for which $AL < 0$. If fact, among the 125 elections with $AL \geq 0$, the discrepancy persisted through the subsequent election in 107 of them for a proportion of 0.856. Among the 315 elections with $AL < 0$, the discrepancy persisted in 165 of them for a proportion of 0.524.

More generally, in Figure 10 we compare the effects of the adjusted seat lean and discrepancy on the persistence of discrepancy. Specifically, we compute the proportion of persistence for all maps whose adjusted lean/discrepancy is greater than or equal to a given level.

As the adjusted seat lean increases we see the expected behavior in the plot – the proportion of persistence increases, reflecting the fact that it becomes harder and harder for the disadvantaged party to overcome the discrepancy. On the other hand, increasing

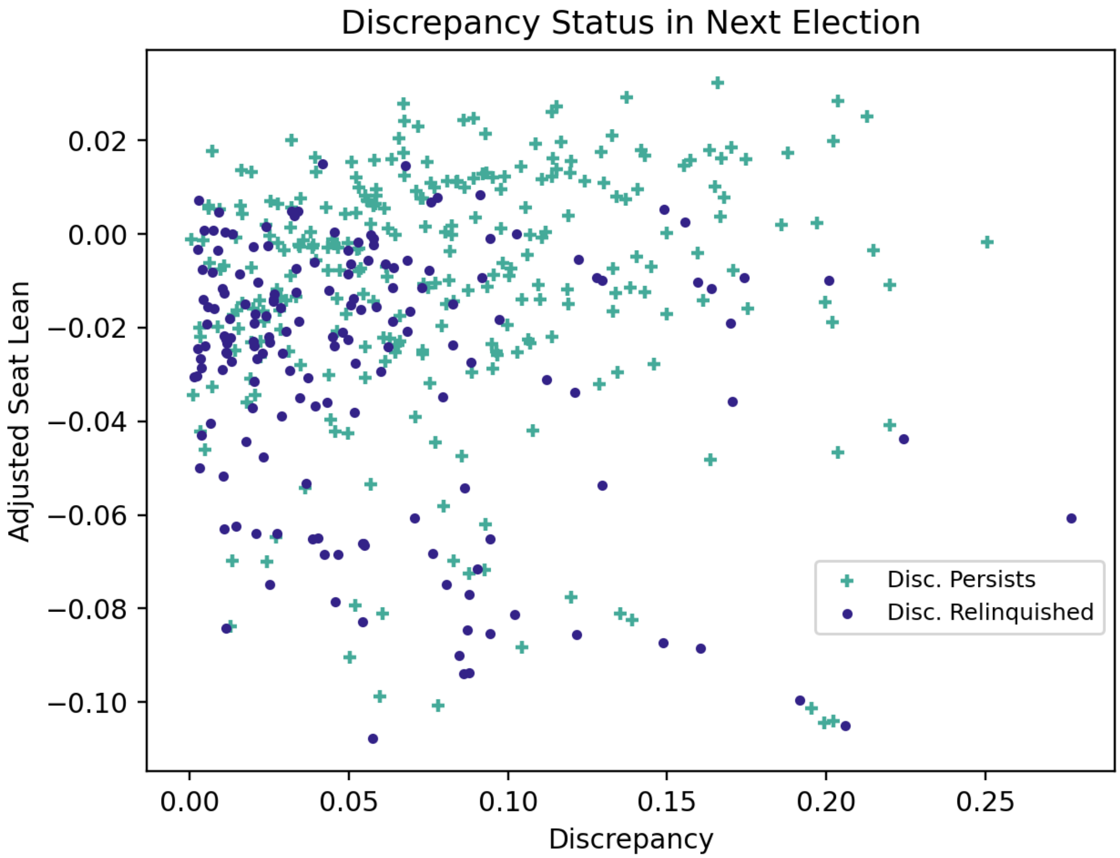


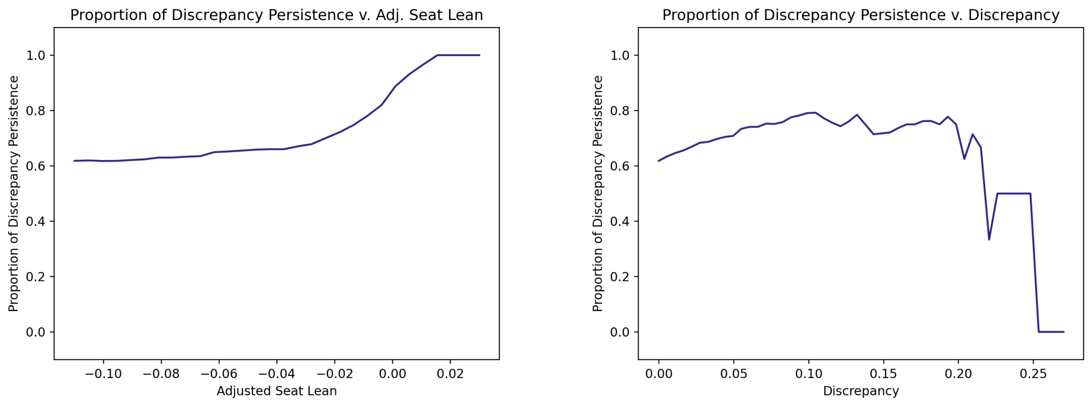
Figure 9: Discrepancy status in next election given D and AL.

discrepancy does not necessarily lead to higher persistence. In fact it is relatively common for very high levels of discrepancy to not persist in the next election. The steep plunge at the end of the discrepancy graph can be attributed to the fact that the largest discrepancy in the data set did not persist.

While discrepancy on its own is not particularly predictive of persistence, we note that given an adjusted seat lean greater than zero, increases in discrepancy do seem to contribute to increased persistence.

4.2 A Standard for Problematic Partisan Gerrymanders

We return to our earlier assertion that in order for a map to be considered a problematic gerrymander it must exhibit high discrepancy and that discrepancy must be very likely to persist in the subsequent election. Having a problematic seat lean on its own is not sufficient for a map to be a problematic gerrymander because the discrepancy in the map, while perhaps very likely to persist, may not be large enough to be of concern. Likewise, a high level of discrepancy is not sufficient because even a very high level of discrepancy may be unlikely to persist.



(a) The proportion of elections whose adjusted seat lean is larger than the given x-value and whose discrepancy persists through the next.

(b) The proportion of elections whose discrepancy is larger than the given x-value and whose discrepancy persists through the next.

Figure 10: A comparison of the effects of increasing adjusted seat lean and discrepancy on persistence of discrepancy. A high level of discrepancy tells us very little about whether the discrepancy will persist, but a high adjusted seat lean virtually guarantees that it will.

Having established $AL \geq 0$ as the cutoff for a problematic lean, what remains is to specify a companion level of discrepancy. There is a fairly wide range of values for D that could be used. In fact, any choice of D larger than about 0.05 in combination with $AL \geq 0$ produces a discrepancy persistence of at least 93% (with minor fluctuations caused by where flips occur). Thus we are really just left to decide how extreme the discrepancy should be allowed to be.

If we set the level at $D = 0.08$, then this represents the 80th percentile of discrepancy for federal elections. However, if we also require $AL \geq 0$, we would flag roughly 7% of all federal elections as being a problematic gerrymander. Moreover, of the 440 elections we examined for discrepancy persistence, 61 satisfy $D \geq 0.08$ and $AL \geq 0$. Of those 61, 58 saw their discrepancy persist to the next election for a rate of about 95%. Of course adjusting the cutoff for D allows us to control the percentage of elections that would be flagged. (For example, to flag 10% of federal elections we would set $D = 0.05$.) For the discussion that follows, we suggest that

$$D \geq 0.08 \text{ and } AL \geq 0 \tag{7}$$

together provide reasonable criteria for flagging a map as a problematic partisan gerrymander.

As with every attempt to provide a simple quantitative threshold for gerrymandering based only on seat and vote proportions, this standard should be viewed as a companion to other considerations. It does not take into account geographical issues, shapes of districts, gerrymandering aimed at minority populations, etc. There may even be instances where it is not possible to draw a district map that would not be considered problematic (in Massachusetts, for example [4].

To alleviate such concerns, the criteria set out here could be used in conjunction with a stipulation (not original to this author) that if a party challenges a map that is flagged by this criteria, then that party must submit a map that is less problematic and which meets any other legal criteria. If the plaintiffs are unable to do so, then the original map would stand.

4.3 Discussion of Examples from Past Elections

In this section we take a look at which federal elections would be flagged as problematic gerrymanders and which elections would be considered the "worst offenders."

Using 0.08 for the discrepancy cutoff, the collection of federal elections flagged by $D \geq 0.08$ and $AL \geq 0$ contains 82 elections. Below we present the fifteen most extreme of these 82 based on a variety of criteria.

Example 11. *Flagged elections with highest levels of discrepancy ($D > 0.165$):*

1. *2010 MA*
2. *2014 MA*
3. *2012 PA*
4. *1980 VA*
5. *2012 OH*
6. *1996 OK*
7. *1998 OK*
8. *2016 CT*
9. *1996 MA*
10. *2012 NC*
11. *2014 NC*
12. *2012 VA*
13. *1972 MO*
14. *2012 SC*
15. *2002 MA*

Example 12. *Flagged elections with highest levels of adjusted lean ($AL > 0.019$):*

1. *2002 MA*
2. *2004 MA*
3. *2012 PA*
4. *2008 MA*
5. *2006 MA*
6. *2016 PA*
7. *2014 MA*
8. *1996 NY*
9. *1974 IL*
10. *2010 MA*
11. *1972 TX*
12. *2000 MA*
13. *2014 PA*
14. *2012 OH*
15. *2006 IL*

Example 13. *Flagged somewhat competitive elections ($0.4 \leq v_B \leq 0.6$) with highest levels of discrepancy (out of 53, $D > 0.149$):*

1. *2010 MA*
2. *2012 PA*
3. *1980 VA*
4. *2012 OH*
5. *2012 NC*
6. *2014 NC*
7. *2012 VA*
8. *2012 SC*
9. *2012 MI*
10. *2016 NC*
11. *2016 PA*
12. *2006 MI*
13. *2006 VA*
14. *2014 MD*
15. *2006 OH*

Example 14. *Flagged somewhat competitive elections ($0.4 \leq v_B \leq 0.6$) with highest levels of adjusted seat lean (out of 53, $AL > 0.015$):*

1. *2012 PA*
2. *2016 PA*
3. *1974 IL*
4. *2010 MA*
5. *2014 PA*
6. *2012 OH*
7. *2006 IL*
8. *2014 NC*
9. *1980 TX*
10. *1992 TX*
11. *1974 OH*
12. *2000 IL*
13. *2010 IL*
14. *2012 NC*
15. *2006 MI*

We offer a few notes on the results:

- Of the recently drawn maps of interest (2012 MA, MD, MI, NC, OH, and PA), all were flagged except 2012 MD whose discrepancy was 0.053. Thus our standard considers the maps drawn for 2012 in MA, MI, NC, OH, and PA to be problematic gerrymanders but not the 2012 map from MD.
- The results for highest discrepancy in competitive elections are particularly striking as 6 of the top 9 maps were drawn for 2012.
- These lists are all dominated by relatively recent elections. This is an indication that gerrymandering has gotten worse, a finding that agrees with that of many authors.
- Pennsylvania's 2012 map appears to be particularly problematic as it scored among the top 3 for highest discrepancy and highest adjusted lean.

Next we take a closer look at some elections of interest. The first is a perfect example for why discrepancy alone would not be a sufficient test for a problematic gerrymander.

Example 15. *In the 1974 election in Nebraska, the Democrats were swept despite earning 47.1% of the statewide vote. This resulted in the highest level of discrepancy of all federal elections in the database,* $D = 0.28$. *The vote shares for the Democrats in the three districts were* $(0.498, 0.467, 0.448)$. *The presence of such close outcomes results in a map that leans heavily to the disadvantaged Democrats,* $L = -0.20$. *The seat lean is very favorable to the Democrats even given the large discrepancy that would have to be overcome:* $AL = -0.06$. *Hence, this map would not be considered problematic despite its large discrepancy. In fact, the discrepancy did not persist in the next election as the Democrats won back a seat, causing the discrepancy to go to 0.*

Our next election provides an example of why using adjusted seat lean alone would not be a sufficient test for a problematic gerrymander.

Example 16. *In New York in 2014, Democrats won 60.8% of the statewide vote and 18 of the 27 seats while their equitable seat share would have been 18.7 seats. The Democrats were the disadvantaged party, but the discrepancy was only* $D = 0.007$. *The election results for each party's four closest wins were (with Republicans as minority):*

$$(\ldots, 0.604, 0.596, 0.566, 0.544, 0.498, 0.490, 0.472, 0.452, \ldots).$$

The Democrats' four closest wins were by margins of 0.002, 0.010, 0.028, and 0.048, while for the Republicans they were 0.044, 0.066, 0.096, and 0.104. Given this, we should expect the Republicans were better positioned than the Democrats for the next election, and both the lean, $L = 0.014$, *and the adjusted lean,* $AL = 0.018$, *reflect this. This indicates that even given the discrepancy present, the map leans heavily in the Republicans' favor for the next election.*

In the subsequent election, we note that the discrepancy persisted. The Democrats increased their vote share to 64.7%, but they did not pick up any additional seats, a result that is consistent with our assessment that the map put Republicans in a favorable position. The 2016 NY map saw the discrepancy increase to 0.066 in favor of the Republicans.

Of the 61 elections that our standard would flag out of the 440 that were eligible to flip, only 3 saw the discrepancy fail to persist in the next election. We discuss one of them next.

Example 17. *The 1984 election in Florida saw the Democrats win 12 of 19 seats while winning 50.7% of the statewide vote – roughly a 2-seat advantage over equitability. With a discrepancy of* $D = 0.091$*, a seat lean of* $L = -0.037$*, and an adjusted seat lean of* $AL = 0.0084$ *we would expect it to be difficult for the Republicans to make up ground in the next election. The Republican district vote proportions for the four closest wins by each party were:*

$$(\ldots, 0.659, 0.647, 0.640, 0.619, 0.446, 0.436, 0.412, 0.395, \ldots),$$

so the map does not look promising for the Republicans for the next election. Indeed, the Republicans did not pick up any seats in 1986, but the discrepancy still went to 0. The reason is that the Democrats saw a surge in their statewide vote share to 56.1% that caused their equitable seat share to increase to match the share they won.

5 Conclusion

The construction presented here for finding fair seat allocations for plurality voting systems is a consequence of a small number of reasonable assumptions about fair maps. The implied baseline standard for fairness gives sensible theoretical results and performs as it should on real two-party and multiparty elections. Furthermore, the standard shows that first-past-the-post systems favor the largest parties while penalizing the smallest as a matter of course. Thus lack of proportionality in such elections is not simply an artifact of the difficulties involved in drawing boundaries or of intentional gerrymandering. If proportionality is a desired property of an electoral system, then plurality voting with winner take all districts is inherently unsuitable. This implication holds even if one doubts the 50/50 split prescribed in Assumption #3 as long as one believes that the smallest parties should have *some* votes in seats they cannot win.

The new standard presented for the detection of problematic partisan gerrymanders remedies known deficiencies in the efficiency gap and measures like it that are based on overall vote and seat proportions. It does so by taking into account not just the discrepancy between fair and actual seat shares but how likely that discrepancy is to persist through the next election. This combination of equitability and stability produces a gerrymandering standard in the two-party case that is simple to compute and sensible when applied both to important theoretical cases, such as maximally competitive sweeps, and real elections. Furthermore, the standard neither forbids nor requires proportionality, as it depends on the particular packing and cracking present at the district level. Finally, as it applies to small values of N and is defined for all elections, including sweeps, the new standard is free of some of the limitations that are present in other measures.

References

[1] N.O. Stephanopoulos and E.M. McGhee, Partisan gerrymandering and the efficiency gap. *The University of Chicago Law Review* (2015), 831-900.

[2] J. Barton, Improving the Efficiency Gap. *Math Horizons* (2018), Vol. 26, No. 1, 18-21.

[3] Elections Canada, Table 7 “Distribution of seats by political affiliation and sex” and Table 9 “Percentage of valid votes by political affiliation,” *Past Elections, 1997-2019*, https://www.elections.ca/.

[4] M. Duchin, T. Gladkova, E. Henninger-Voss, B. Klingensmith, H. Newman, and H. Wheelen, Locating the Representational Baseline: Republicans in Massachusetts. *Election Law Journal*, (2019), Vol. 18, No. 4, 388-401.

[5] G. Warrington, U.S. Election Data, https://github.com/GSWarrington/gerrymandering-declination

[6] G. Warrington, “Quantifying Gerrymandering Using the Vote Distribution,” *Election Law Journal: Rules, Politics, and Policy*, Mar 2018, 39-57, http://doi.org/10.1089/elj.2017.0447

A Results by Province for 2019 Canadian Parliament

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.340	0.262	0.244	0.125	0.017	0.012
Prop. Seats	14.3	11.0	10.2	5.3	0.7	0.5
Fair Seats	17.6	11.4	10.0	2.9	0.1	0.0
Actual Seats	17	11	11	2	0	1

Table 7: 2019 Canadian Parliament results compared to proportionality and the new standard for British Columbia with 42 total seats.

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.690	0.138	0.116	0.028	0.022	0.006
Prop. Seats	23.5	4.7	3.9	1.0	0.7	0.2
Fair Seats	29.6	2.4	1.8	0.1	0.1	0.0
Actual Seats	33	0	1	0	0	0

Table 8: 2019 Canadian Parliament results compared to proportionality and the new standard for Alberta with 34 total seats.

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.640	0.117	0.196	0.026	0.018	0.003
Prop. Seats	9.0	1.6	2.7	0.4	0.3	0.0
Fair Seats	11.4	0.7	1.8	0.1	0.0	0.0
Actual Seats	14	0	0	0	0	0

Table 9: 2019 Canadian Parliament results compared to proportionality and the new standard for Saskatchewan with 14 total seats.

Party:	Cons.	Lib.	NDP	BQ	Green	People's	Other
Vote Prop.	0.160	0.343	0.108	0.324	0.045	0.015	0.005
Prop. Seats	12.5	26.8	8.4	25.3	3.5	1.2	0.4
Fair Seats	9.2	33.1	4.6	30.1	0.9	0.1	0.0
Actual Seats	10	35	1	32	0	0	0

Table 10: 2019 Canadian Parliament results compared to proportionality and the new standard for Quebec with 78 total seats.

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.452	0.265	0.208	0.051	0.017	0.007
Prop. Seats	6.3	3.7	2.9	0.7	0.2	0.1
Fair Seats	8.2	3.4	2.2	0.2	0.0	0.0
Actual Seats	7	4	3	0	0	0

Table 11: 2019 Canadian Parliament results compared to proportionality and the new standard for Manitoba with 14 total seats.

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.331	0.416	0.168	0.062	0.016	0.007
Prop. Seats	40.1	50.3	20.3	7.5	1.9	0.8
Fair Seats	42.9	62.6	13.2	2.2	0.2	0.0
Actual Seats	36	79	6	0	0	0

Table 12: 2019 Canadian Parliament results compared to proportionality and the new standard for Ontario with 121 total seats.

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.328	0.375	0.094	0.172	0.020	0.011
Prop. Seats	3.3	3.8	0.9	1.7	0.2	0.1
Fair Seats	3.7	4.6	0.4	1.2	0.0	0.0
Actual Seats	3	6	0	1	0	0

Table 13: 2019 Canadian Parliament results compared to proportionality and the new standard for New Brunswick with 10 total seats.

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.257	0.414	0.189	0.110	0.012	0.018
Prop. Seats	2.8	4.6	2.1	1.2	0.1	0.2
Fair Seats	2.8	6.0	1.6	0.6	0.0	0.0
Actual Seats	1	10	0	0	0	0

Table 14: 2019 Canadian Parliament results compared to proportionality and the new standard for Nova Scotia with 11 total seats.

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.279	0.449	0.237	0.031	0.001	0.003
Prop. Seats	2.0	3.1	1.7	0.2	0.0	0.0
Fair Seats	1.8	3.9	1.3	0.0	0.0	0.0
Actual Seats	0	6	1	0	0	0

Table 15: 2019 Canadian Parliament results compared to proportionality and the new standard for Newfoundland with 7 total seats.

Party:	Cons.	Lib.	NDP	Green	People's	Other
Vote Prop.	0.273	0.437	0.076	0.209	0	0.005
Prop. Seats	1.1	1.7	0.3	0.8	0.0	0.0
Fair Seats	1.0	2.2	0.1	0.6	0.0	0.0
Actual Seats	0	4	0	0	0	0

Table 16: 2019 Canadian Parliament results compared to proportionality and the new standard for Prince Edward Island with 4 total seats.

B Best Fit ρ

If we let ρ be the proportion of the unnecessarily wasted votes that losing parties must allocate to seats they cannot win, then a straightforward adjustment to our derivation gives a system of differential equations with the following initial solutions:

$$
\begin{aligned}
v_{A_1}(x) &= 1 - \sum_{k=2}^{K} \frac{v_{A_k}}{(1-x)^{1-\rho}} \\
v_{A_2}(x) &= \frac{v_{A_2}}{(1-x)^{1-\rho}} \\
&\vdots \\
v_{A_K}(x) &= \frac{v_{A_K}}{(1-x)^{1-\rho}}.
\end{aligned}
$$

The process of finding seat allocations then looks very much the same, beginning with finding x_1. We solve the equation:

$$
1 - \sum_{k=2}^{K} \frac{v_{A_k}}{(1-x_1)^{1-\rho}} = \frac{v_{A_2}}{(1-x_1)^{1-\rho}},
$$

which yields

$$
x_1 = 1 - \left(v_{A_2} + \sum_{k=2}^{K} v_{A_k} \right)^{\frac{1}{1-\rho}}.
$$

Then the method proceeds as before.

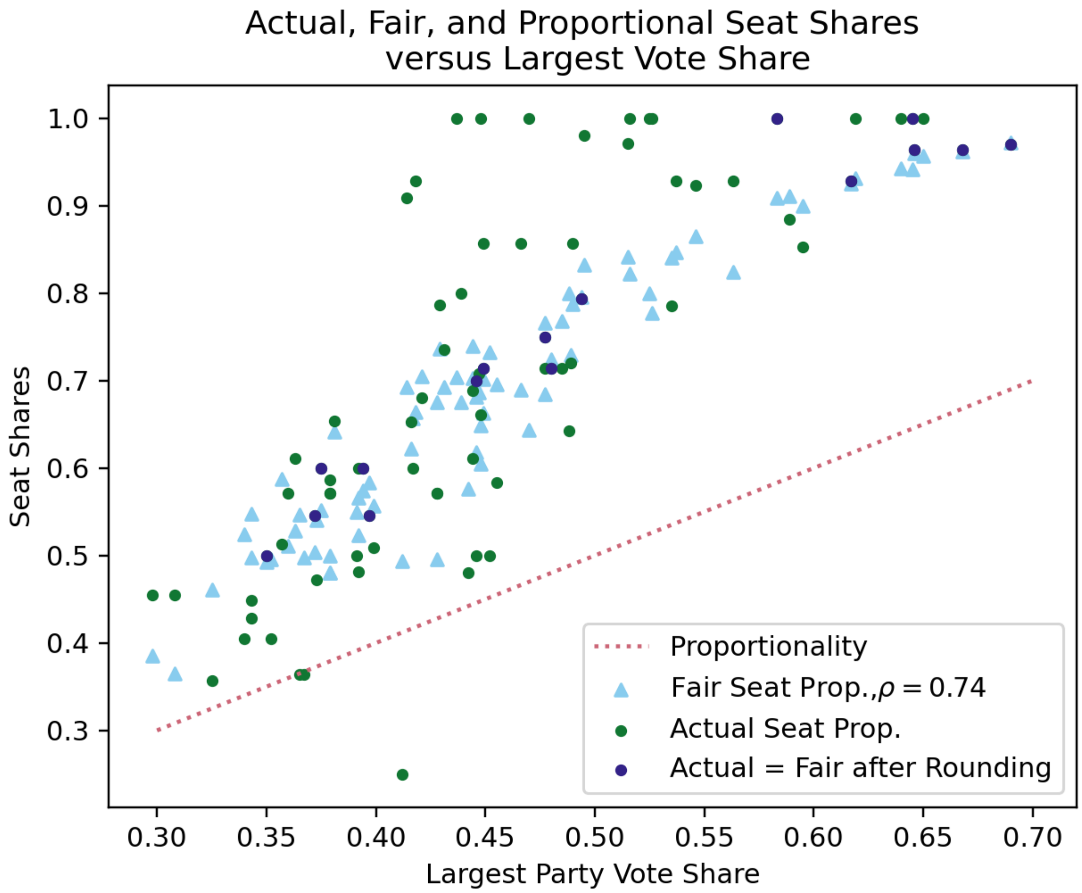


Figure 11: Actual, fair ($\rho = 0.74$), and proportional seat shares by province versus the largest party's vote share for Canadian Parliamentary elections 1997-2019. Proportions are not rounded to the nearest seat. The dark dots are seat shares that would be fair after rounding.

Figure 11 shows a clearly better fit for the election data for the largest party's seat share.

Figure 12 compares the distribution of average discrepancy measured from fairness (where $\rho = 0.74$) and from proportionality. The original standard ($\rho = 0.5$) had median discrepancy 0.036, while $\rho = 0.74$ yields a median of about 0.02.

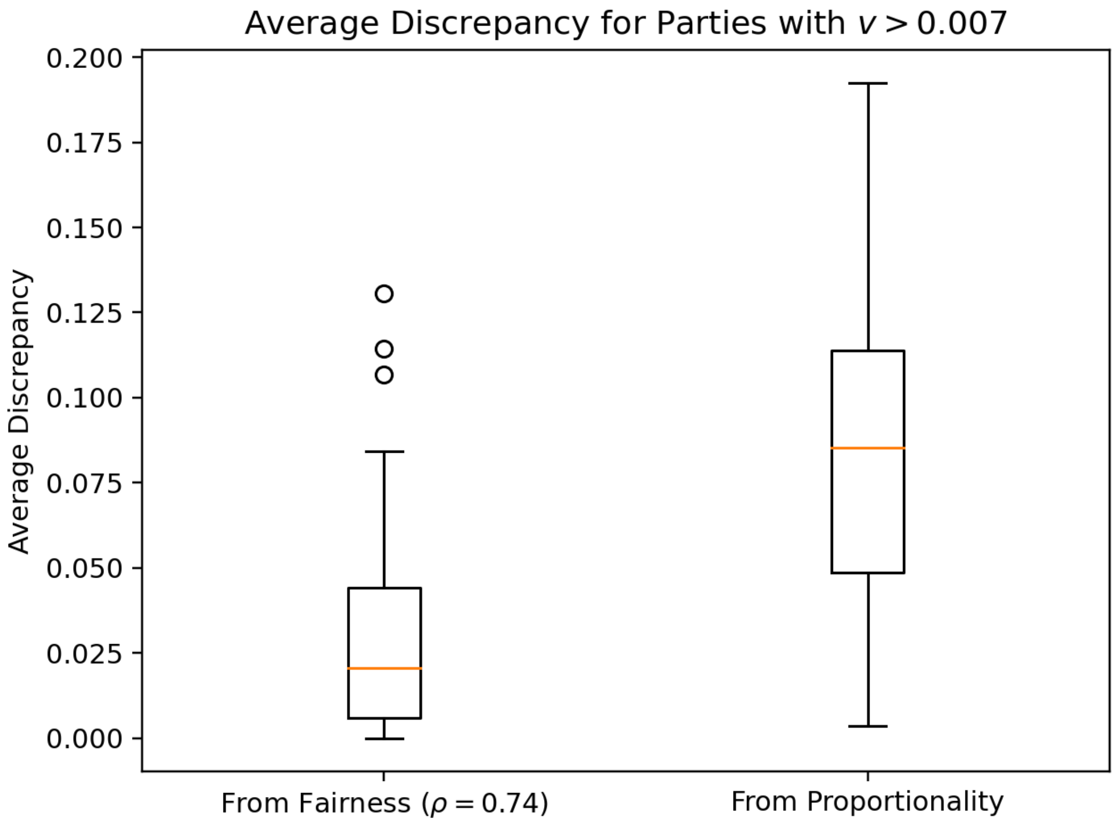


Figure 12: A comparison by of the distributions of average discrepancy between the new standard and that implied by proportionality. For each province, the average is taken over parties with at least 0.7% of the overall vote.

Finally, Figure 13 shows a scatterplot of average discrepancy taken from fairness (where $\rho = 0.74$) versus average discrepancy from proportionality. Compared to Figure 4 we see that the selection $\rho = 0.74$, while leading to a better overall fit for the election data, is bested by proportionality more often than for $\rho = 0.5$.

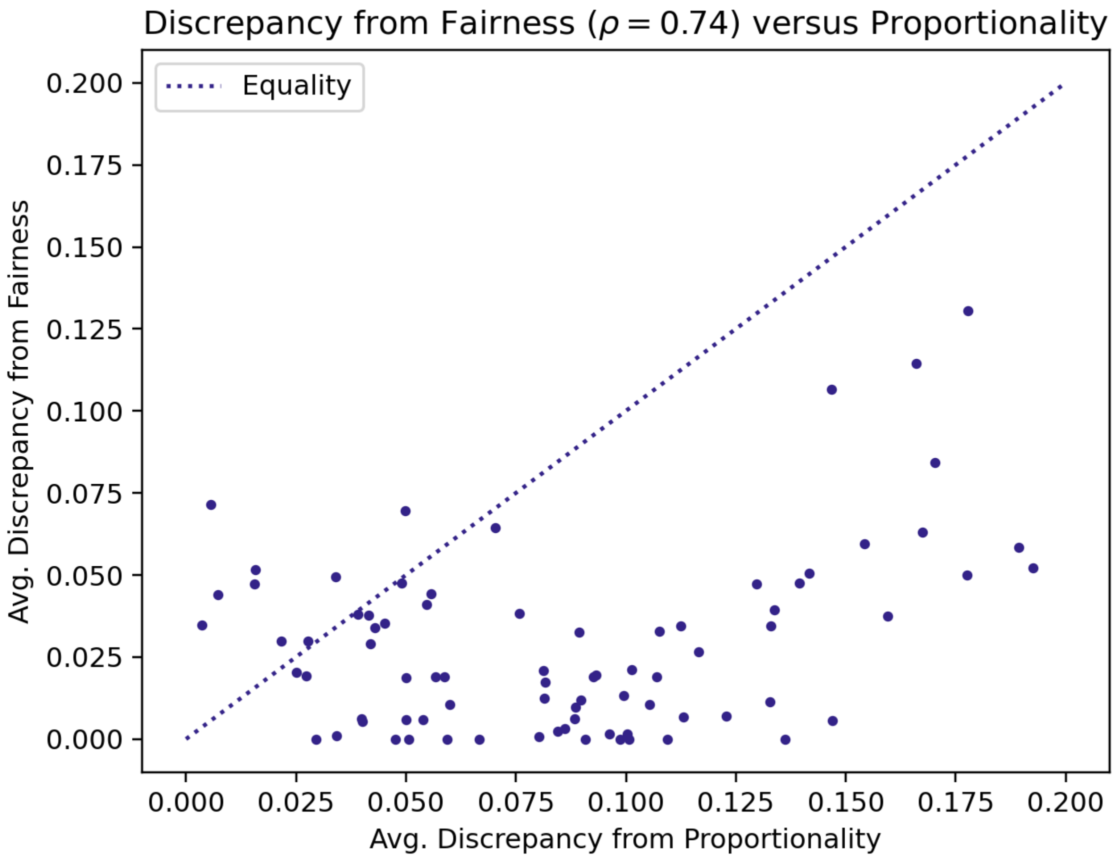


Figure 13: A scatterplot of average discrepancy based on the new standard ($\rho = 0.74$) and based on proportionality. The average is taken over parties with at least 0.7% of the overall vote.