

Assessing the Current Wisconsin State Legislative Districting Plan

Simon Jackman

July 7, 2015

Contents

1	Introduction	1
2	Qualifications, Publications and Compensation	2
3	Summary	2
4	Redistricting plans	6
4.1	Seats-Votes Curves	9
5	Partisan bias	11
5.1	Multi-year method	11
5.2	Uniform swing	13
5.3	Critiques of partisan bias	14
6	The Efficiency Gap	15
6.1	The efficiency gap when districts are of equal size	16
6.2	The seats-vote curve when the efficiency gap is zero	17
6.3	The efficiency gap as an excess seats measure	19
7	State legislative elections, 1972-2014	19
7.1	Grouping elections into redistricting plans	22
7.2	Uncontested races	22
8	Imputations for Uncontested Races	24
8.1	Imputation model 1: presidential vote shares	26
8.2	Imputation model 2	29
8.3	Combining the two sets of imputations	29
8.4	Seat and vote shares in 786 state legislative elections	32
9	The efficiency gap, by state and election	32
9.1	Are efficiency gap estimates statistically significant?	36
9.2	Over-time change in the efficiency gap	44
9.3	Within-plan variation in the efficiency gap	48
9.4	How often does the efficiency gap change sign?	53
10	A threshold for the efficiency gap	56
10.1	Conditioning on the first election in a districting plan	60
10.2	Conditioning on the first two elections in a districting plan	63
10.3	An actionable EG threshold?	63
10.4	Confidence in a given threshold	66
11	Conclusion: the Wisconsin plan	69

1 Introduction

My name is Simon Jackman. I am currently a Professor of Political Science at Stanford University, and, by courtesy, a Professor of Statistics. I joined the Stanford faculty in 1996. I teach classes on American politics and statistical methods in the social sciences.

I have been asked by counsel representing the plaintiffs in this lawsuit (the “Plaintiffs”) to analyze relevant data and provide expert opinions in the case titled above. More specifically, I have been asked

- to determine if the current Wisconsin legislative districting plan constitutes a partisan gerrymander;
- to explain a summary measure of a districting plan known as “the efficiency gap” ([Stephanopoulos and McGhee, 2015](#)), what it measures, how it is calculated, and to assess how well it measures partisan gerrymandering;
- to compare the efficiency gap to extant summary measures of districting plans such as partisan bias;
- to analyze data from state legislative elections in recent decades, so as to assess the properties of the efficiency gap and to identify plans with high values of the efficiency gap;
- to suggest a threshold or other measure that can be used to determine if a districting plan is an extreme partisan gerrymander;
- to describe how the efficiency gap for the Wisconsin districting plan compares to the values of the efficiency gap observed in recent decades elsewhere in the United States;
- to describe where the efficiency gap for the current Wisconsin districting plan lies in comparison with the threshold for determining if a districting plan constitutes an extreme partisan gerrymander.

My opinions are based on the knowledge I have amassed over my education, training and experience, and follow from statistical analysis of the following data:

- a large, canonical data set on candidacies and results in state legislative elections, 1967 to the present available from the Inter-University Consortium for Political and Social Research ([ICPSR study number 34297](#)); I use a release of the data updated through 2014, maintained by Karl Klarner (Indiana State University and Harvard University).
- presidential election returns, 2000-2012, aggregated to state legislative districts.

2 Qualifications, Publications and Compensation

My Ph.D. is in Political Science, from the University of Rochester, where my graduate training included courses in econometrics and statistics. My curriculum vitae is attached to this report.

All publications that I have authored and published in the past ten years appear in my curriculum vitae. Those publications include peer-reviewed journals such as: *The Journal of Politics*, *Electoral Studies*, *The American Journal of Political Science*, *Legislative Studies Quarterly*, *Election Law Journal*, *Public Opinion Quarterly*, *Journal of Elections*, *Public Opinion and Parties*, and *PS: Political Science and Politics*.

I have published on properties of electoral systems and election administration in *Legislative Studies Quarterly*, the *Australian Journal of Political Science*, the *British Journal of Political Science*, and the *Democratic Audit of Australia*. I am a Fellow of the Society for Political Methodology and a member of the American Academy of Arts and Sciences.

I am being compensated at a rate of \$250 per hour.

3 Summary

1. **Partisan gerrymandering and wasted votes.** In two-party, single-member district electoral systems, a partisan gerrymander operates by effectively “wasting” more votes cast for one party than for the other. Wasted votes are votes for a party in excess of what the party needed to win a given district or votes cast for a party in districts that the party doesn’t win. Differences

in wasted vote rates between political parties measure the extent of partisan gerrymandering.

2. **The efficiency gap (EG)** is a relative, wasted vote measure, the ratio of one party's wasted vote rate to the other party's wasted vote rate. EG can be computed directly from a given election's results, without recourse to extensive statistical modeling or assumptions about counter-factual or hypothetical election outcomes, unlike other extant measures of the fairness of an electoral system (e.g., partisan bias).
3. The efficiency gap is an "excess seats" measure, reflecting the nature of a partisan gerrymander. An efficiency gap in favor one party sees it wasting fewer votes than its opponent, thus translating its votes across the jurisdiction into seats more efficiently than its opponent. This results in the party winning more seats than we'd expect given its vote share (V) and if wasted vote rates were the same between the parties. $EG = 0$ corresponds to no efficiency gap between the parties, or no partisan difference in wasted vote rates. In this analysis (but without loss of generality) EG is normed such that negative EG values indicate higher wasted vote rates for Democrats relative to Republicans, and $EG > 0$ the converse.
4. A districting plan in which EG is consistently observed to be positive is evidence that the plan embodies a pro-Democratic gerrymander; the magnitudes of the EG measures speak to the severity of the gerrymander. Conversely, a districting plan with consistently negative values of the efficiency gap is consistent with the plan embodying a pro-Republican gerrymander.
5. **Performance of the efficiency gap in 786 state legislative elections.** My analysis of 786 state legislative elections (1972-2014) examines properties of the efficiency gap. EG is estimated with some uncertainty in the presence of uncontested districts (and uncontested districts are quite prevalent in state legislative elections), but this source of uncertainty is small relative to differences in the EG across states and across districting plans.
6. **Stability of the efficiency gap.** EG is stable in pairs of temporally adjacent elections held under the same districting plan. In 580 pairs of consecutive

EG measures, the probability that each *EG* measure has the same sign is 74%. In 141 districting plans with three or more elections, 35% have a better than 95% probability of *EG* being negative or positive for the entire duration of the plan; in about half of the districting plans the probability that *EG* doesn't change sign is above 75%.

7. **Recent decades show more pro-Republican gerrymandering, as measured by the efficiency gap.** Efficiency gap measures in recent decades show a pronounced shift in a negative direction, indicative of an increased prevalence of districting plans favoring Republicans. Among the 10 most pro-Democratic *EG* measures in my analysis, *none* were recorded after 2000.
8. **The current Wisconsin state legislative districting plan** (the "Current Wisconsin Plan"). In Wisconsin in 2012, the average Democratic share of district-level, two-party vote (V) is estimated to be 51.4% (± 0.6 , the uncertainty stemming from imputations for uncontested seats); recall that Obama won 53.5% of the two-party presidential vote in Wisconsin in 2012. Yet Democrats won only 39 seats in the 99 seat legislature ($S = 39.4\%$), making Wisconsin one of 7 states in 2012 where we estimate $V > 50\%$ but $S < 50\%$. In Wisconsin in 2014, V is estimated to be 48.0% (± 0.8) and Democrats won 36 of 99 seats ($S = 36.4\%$).
9. Accordingly, Wisconsin's *EG* measures in 2012 and 2014 are large and negative: -.13 and -.10 (to two digits of precision). The 2012 estimate is the largest *EG* estimate in Wisconsin over the 42 year period spanned by this analysis (1972-2014).
10. Among 79 *EG* measures generated from state legislative elections after the 2010 round of redistricting, Wisconsin's *EG* scores rank 9th (2012, 95% CI 4 to 13) and 18th (2014, 95% CI 14 to 21). Among 786 *EG* measures in the 1972-2014 analysis, the magnitude of Wisconsin's 2012 *EG* measure is surpassed by only 27 (3.4%) other cases.
11. Analysis of efficiency gaps measures in the post-1990 era indicates that conditional on the magnitude of the Wisconsin 2012 efficiency gap (the first election under the Current Wisconsin Plan), there is a 100% probability

that *all subsequent elections* held under that plan will also have efficiency gaps disadvantageous to Democrats.

12. **The Current Wisconsin Plan presents overwhelming evidence of being a pro-Republican gerrymander.** In the entire set of 786 state legislative elections and their accompanying *EG* measures, there are *no precedents* prior to this cycle in which a districting plan generates an initial two-election sequence of *EG* scores that are each as large as those observed in *WI*.
13. The Current Wisconsin Plan is generating *EG* measures that make it *extremely likely* that it has a systematic, historically large and enduring, pro-Republican advantage in the translation of votes into seats in Wisconsin's state legislative elections.
14. **An actionable threshold based on the efficiency gap.** Historical analysis of the relationship between the first *EG* measure we observe under a new districting plan and the subsequent *EG* measures lets us assess the extent to which that first *EG* estimate is a *reliable* indicators of a *durable* and hence *systematic* feature of the plan. In turn, this let us assess the *confidence* associated with a range of possible *actionable EG thresholds*.
15. My analysis suggests that *EG* greater than .07 in absolute value be used as an actionable threshold. Relatively few plans produce a first election with an *EG* measure in excess of this threshold, and of those that do, the historical analysis suggests that most go on to produce a sequence of *EG* estimates indicative of systematic, partisan advantage consistent with the first election *EG* estimates, At the 0.07 threshold, 95% of plans would be either (a) undisturbed by the courts, or (b) struck down because we are sufficiently confident that the plan, if left undisturbed, would go on to produce a one-sided sequence of *EG* estimates, consistent with the plan being a partisan gerrymander. In short, our "confidence level" in the 0.07 threshold is 95%.
16. **The Current Wisconsin Plan is generating estimates of the efficiency gap far in excess of this proposed, actionable threshold.** In 2012 elections to the Wisconsin state legislature, the efficiency gap is estimated to be -.13; in

2014, the efficiency gap is estimated to be $-.10$. Both measures are separately well beyond the conservative $.07$ threshold suggested by the analysis of efficiency gap measures observed from 1972 to the present.

A vivid, graphical summary of my analysis appears in Figure 1, showing the average value of the efficiency gap in 206 districting plans, spanning 41 states and 786 state legislative elections from 1972 to 2014. The Current Wisconsin Plan has been in place for two elections (2012 and 2014), with an average efficiency gap of $-.115$. Details on the interpretation and calculation of the efficiency gap come later in my report, but for now note that negative values of the efficiency gap indicate a districting plan favoring Republicans, while positive values indicate a plan favoring Democrats. Note that *only four other districting plans have lower average efficiency gap scores than the Current Wisconsin Plan*, and these are also from the post-2010 round of redistricting. That is, Wisconsin's current plan is generating the 5th lowest average efficiency gap observed in over 200 other districting plans used in state legislative elections throughout the United States over the last 40 years. The analysis I report here documents why the efficiency gap is a valid and reliable measure of partisan gerrymandering and why are confident that the current Wisconsin plan exceeds even a conservative definition of partisan gerrymandering.

4 Redistricting plans

A districting plan is an exercise in map drawing, partitioning a jurisdiction into districts, typically required to be contiguous, mutually exclusive and exhaustive regions, and — at least in the contemporary United States — of approximately the same population size. In a single-member, simple plurality (SMSP) electoral system, the highest vote getter in each district is declared the winner of the election. Partisan gerrymandering is the process of drawing districts that favor one party, typically by creating a set of districts that help the party win an excess of seats (districts) relative to its jurisdiction-wide level of support.

What might constitute evidence of partisan gerrymandering? One indication might be a series of elections conducted under the same districting plan in which a party's seat share (S) is unusually large (or small) relative to its vote share (V).

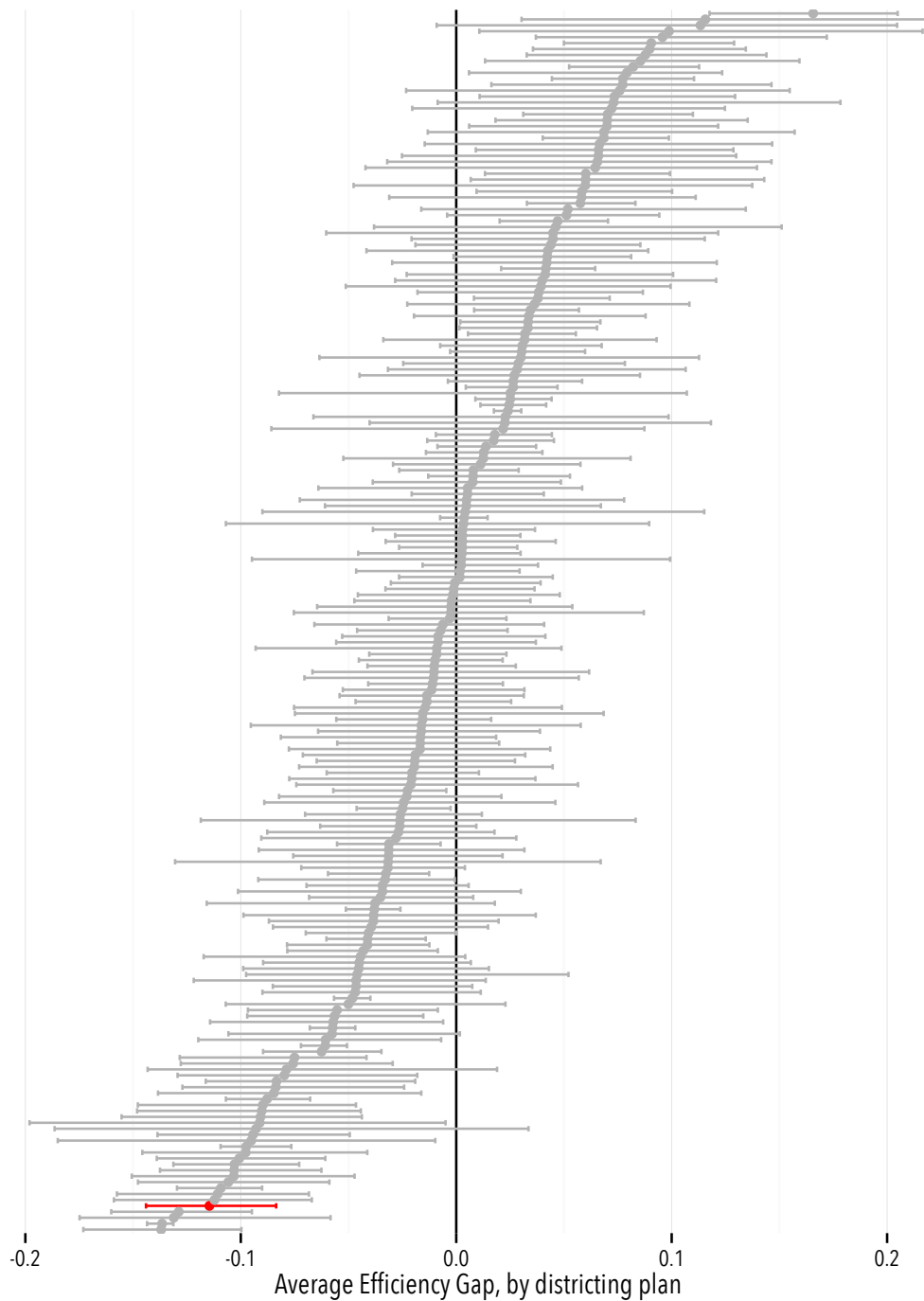


Figure 1: Average efficiency gap score, 206 districting plans, 1972-2014. Plans have been sorted from low average *EG* scores to high. Horizontal lines cover 95% confidence intervals. Negative efficiency gap scores are plans that disadvantage Democrats; positive efficiency gap scores favor Democrats. The Current Wisconsin Plan is shown in red. See also Figure 36.

There may be elections where a party wins a majority of seats (and control of the jurisdiction's legislature) despite not winning a majority of votes: $S > .5$ while $V < .5$ and vice-versa. In fact, there are numerous instances of mismatches between the party winning the statewide vote and the party controlling the state legislature in recent decades. I estimate that since 1972 there have been 63 cases of Democrats winning a majority of the vote in state legislative elections, while not winning a majority of the seats, and 23 cases of the reverse phenomenon, where Democrats won a majority of the seats with less than 50% of the statewide, two-party vote.

Geographic clustering of partisans is typically a prerequisite for partisan gerrymandering. This is nothing other than partisan "packing": a gerrymandered districting plan creates a relatively small number of districts that have unusually large proportions of partisans from party *B*. The geographic concentration of party *B* partisans might make creating these districts a straightforward task. In other districts in the jurisdiction, party *B* supporters never (or seldom) constitute a majority (or a plurality), making those districts "safe" for party *A*. This districting plan helps ensure party *A* wins a majority of seats even though party *B* has a majority of support across the jurisdiction, or at the very least, the districting plan helps ensure that party *A*'s seat share exceeds its vote share in any given election.

It is conventional in political science to say that such a plan allows party *A* to "more efficiently" translate its votes into seats, relative to the way the plan translates party *B*'s votes into seats. This nomenclature is telling, as we will see when we consider the *efficiency gap* measure, below.

Assessing the partisan fairness of a districting plan is fundamentally about measuring a party's excess (or deficit) in its seat share relative to its vote share. The efficiency gap is such a summary measure. To assess the properties of the efficiency gap, I first review some core concepts in the analysis of districting plans: vote shares, seat shares, and the relationship between the two quantities in single-member districts.

4.1 Seats-Votes Curves

Electoral systems translate parties' vote shares (V) into seat shares (S). Both V and S are proportions. Plotting the two quantities V and S against one another yields the “seats-votes” curve, a staple in the analysis of electoral systems and districting plans. Two seats-votes curves are shown in Figure 2, one showing a non-linear relationship between seats and votes typical of single-member district systems,¹ the other showing a linear relationship between seats and votes observed under proportional representation systems.

In pure proportional representation (PR) voting systems, seats-votes curves are 45 degree lines by design, crossing the $(V, S) = (.5, .5)$ point: i.e., under PR, $S = V$ and a party that wins 50% of the vote will be allocated 50% of the seats. Absent a deterministic allocation rule like pure PR, seats-votes curves are most usefully thought of in probabilistic terms, due to the fact that there are many possible configurations of district-specific outcomes corresponding to a given jurisdiction-wide V , and hence uncertainty — represented by a probability *distribution* — over possible values of S given V .

In single-member, simple plurality (SMSP) systems, we often see non-linear, “S”-shaped seats-votes curves. With an approximately symmetric mix of districts (in terms of partisan leanings), large changes in seat shares (S) can result from relatively small changes in votes shares (V) at the middle of the distribution of district types. This presumes a districting plan such that both parties have a small number of “strongholds,” with extremely large changes in vote shares needed to threaten these districts, and so the seats-votes curve tends to “flatten out” as jurisdiction-wide vote share (V) takes on relatively large or small values. Other shapes are possible too: e.g., bipartisan, incumbent-protection plans generate seats-votes curves that are largely flat for most values of V , save for the constraint that the curve run through the points $(V, S) = (0, 0)$ and $(1, 1)$; i.e., relatively large movements in V generates relatively little change in seats shares.

¹The curve labeled “Cube Law” in Figure 2 is generated assuming that $S/(1-S) = [V/(1-V)]^3$, an approximation for the lack of proportionality we observe in single-member district systems, though hardly a “law.”

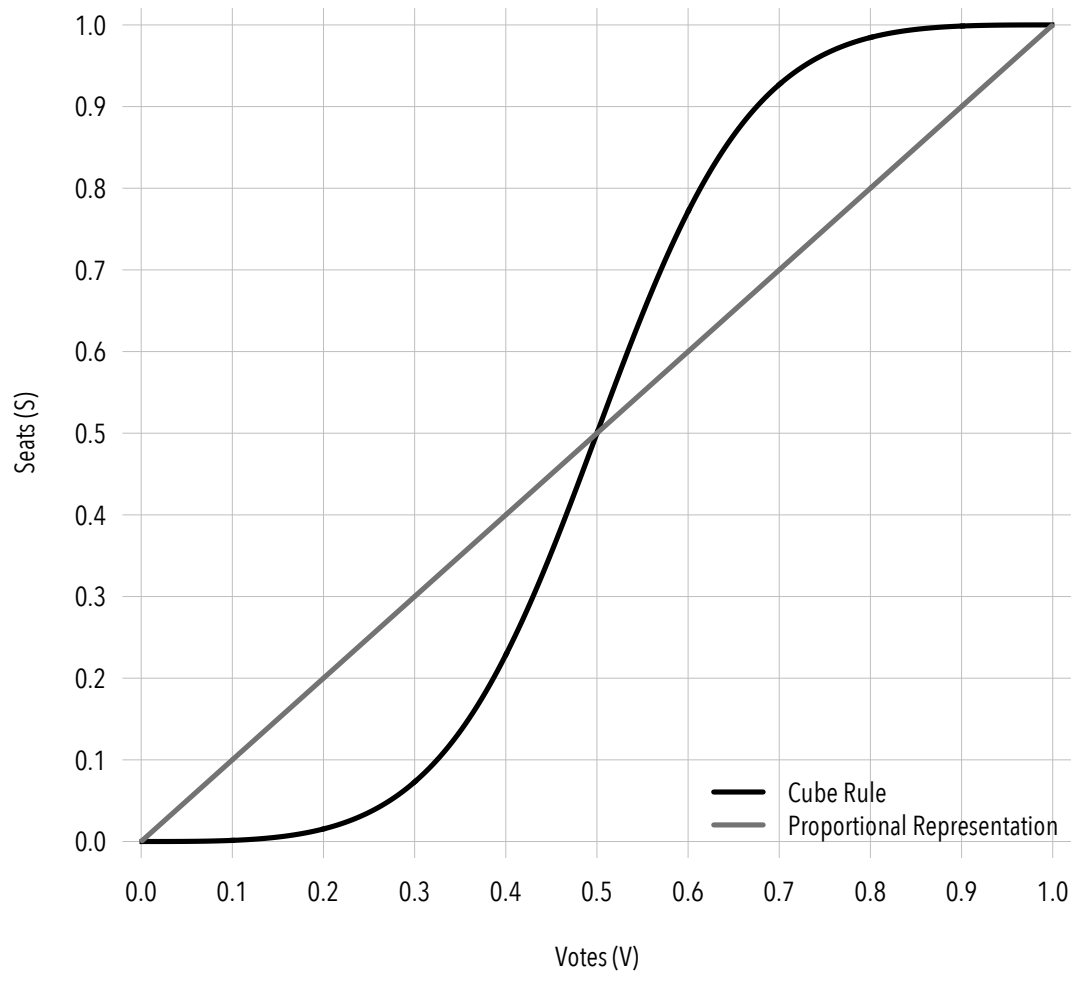


Figure 2: Two Theoretical Seats-Votes Curves

5 Partisan bias

Both of the hypothetical seats-votes curves in Figure 2 run through the “50-50” point, where $V = .5$ and $S = .5$. An interesting empirical question is whether *actual* seats-votes curves run through this point, or more generally, whether the seats-votes curve is symmetric about $V = .5$. Formally, symmetry of the seats-vote curve is the condition that $E(S|V) = 1 - E(S|1 - V)$, where E is the expectation operator, averaging over the uncertainty with respect to S given V . The vertical offset from the $(.5, .5)$ point for a seats-votes curve is known as *partisan bias*: the extent to which a party’s expected seat share lies above or below 50%, conditional on that party winning 50% of the jurisdiction-wide vote.

Figure 3 shows three seats-votes curves, with the graph clipped to the region $V \in [.4, .6]$ and $S \in [.4, .6]$ so as to emphasize the nature of partisan bias. The blue, positive bias curve “lifts” the seats-votes curve; it crosses $S = .5$ with $V < .5$ and passes through the upper-left quadrant of the graph. That is, with positive bias, a party can win a majority of the seats with *less* than a majority of the jurisdiction-wide or average vote; equivalently, if the party wins $V = .5$, it can expect to win *more* than 50% of the seats. Conversely, with negative bias, the opposite phenomenon occurs: the party can’t expect to win a majority of the seats until it wins more than a majority of the jurisdiction-wide or average vote.

5.1 Multi-year method

With data from multiple elections under the same district plan, partisan bias can be estimated by fitting a seats-votes curve to the observed seat and vote shares, typically via a simple statistical technique such as linear regression; this approach has a long and distinguished lineage in both political science and statistics (e.g., Edgeworth, 1898; Kendall and Stuart, 1950; Tufte, 1973). Niemi and Fett (1986) referred to this method of estimating the partisan bias of an electoral system as the “multi-year” method, reflecting the fact that the underlying data comes from a sequence of elections.

This approach is of limited utility when assessing a new or proposed districting plan. More generally, it is of no great help to insist that a sequence of elections must be conducted under a redistricting plan before the plan can be properly assessed. Indeed, few plans stay intact long enough to permit reliable analysis in

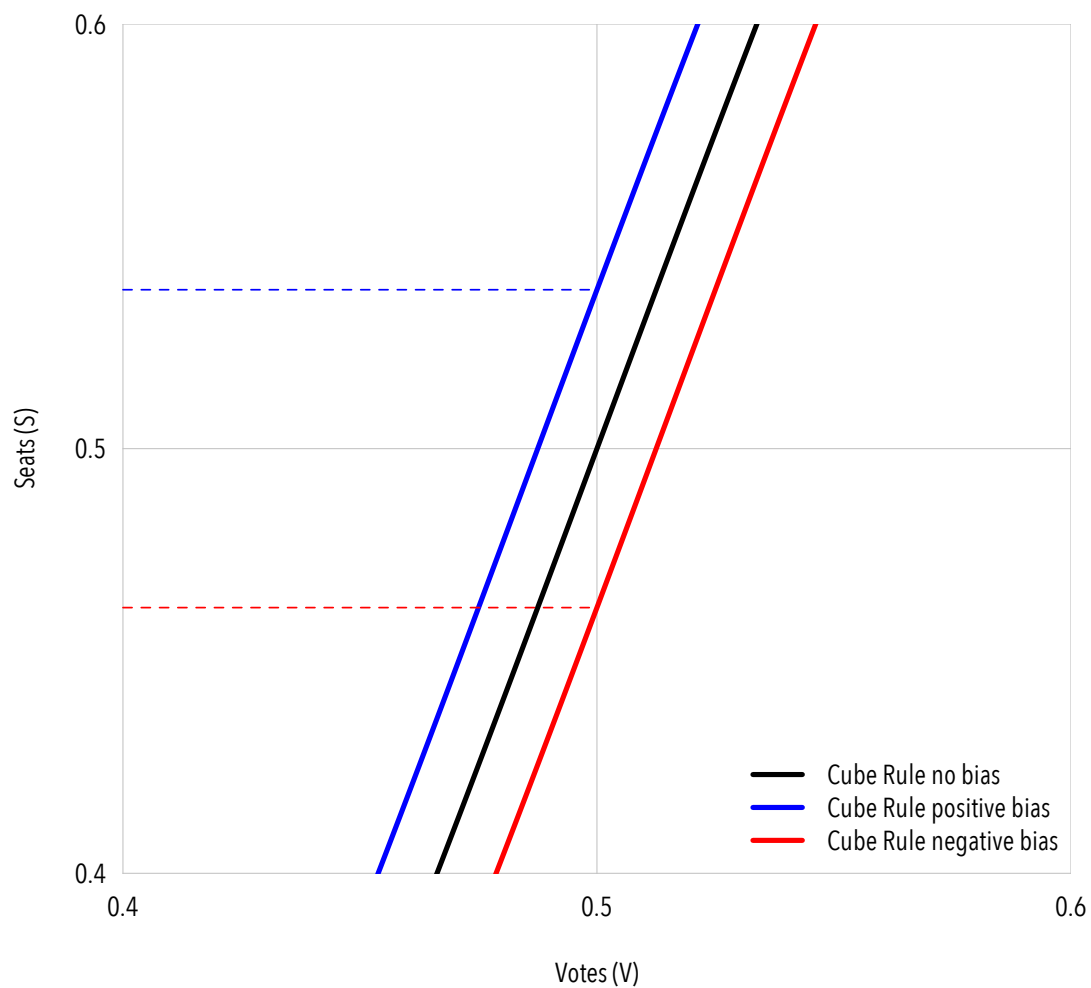


Figure 3: Theoretical seats-votes curves, with different levels of partisan bias. This graph is “zoomed in” on the region $V \in [.4, .6]$ and $S \in [.4, .6]$; the seats-votes “curves” are approximately linear in this region.

this way. State-level plans in the United States might generate as many five elections between decennial censuses. Accordingly, many uses of the “multi-year” method pool multiple plans and/or across jurisdictions, so as to estimate average partisan bias. For instance, [Niemi and Jackman \(1991\)](#) estimated average levels of partisan bias in state legislative districting plans, collecting data spanning multiple decades and multiple states, and grouping districting plans by the partisanship of the plan’s authors (e.g., plans drawn under Republican control, Democratic control, mixed, or independent).

Assessing the properties of a districting plan after a tiny number of elections — or *no* elections — requires some assumptions and/or modeling. A single election yields just a single (V, S) data point, through which no unique seats-vote curve can be fitted and so partisan bias can’t be estimated without further assumptions. Absent *any* actual elections under the plan, we might examine votes from a previous election, say, with precinct level results re-aggregated to the new districts.

5.2 Uniform swing

One approach—dating back to Sir David Butler’s (1974) pioneering work on British elections—is the uniform partisan swing approach. Let $\mathbf{v} = (v_1, \dots, v_n)'$ be the set of vote shares for party *A* observed in an election with n districts. Party *A* wins seat i if $v_i > .5$, assuming just two parties (or defining v as the share of two-party vote); i.e., $s_i = 1$ if $v_i > .5$) and otherwise $s_i = 0$. Party *A*’s seat share is $S = \frac{1}{n} \sum_{i=1}^n s_i$. V is the jurisdiction-wide vote share for party *A*, and if each district had the same number of voters $V = \bar{v} = \frac{1}{n} \sum_{i=1}^n v_i$, the average of the district-level v_i . Districts are never *exactly* equal sized, in which case we can define V as follows: let t_i be the number of voters in district i , and $V = \sum_{i=1}^n t_i v_i / \sum_{i=1}^n t_i$.

The uniform swing approach perturbs the observed district-level results \mathbf{v} by a constant factor δ , corresponding to a hypothetical amount of *uniform swing* across all districts. For a given δ , let $v_i^* = v_i + \delta$ which in turn generates $V^* = V + \delta$ and an implied seat share S^* . Now let δ vary over a grid of values ranging from $-V$ to $1 - V$; then V^* varies from 0 to 1 and a corresponding value of S^* can also be computed at every grid point. The resulting set of (V^*, S^*) points are then plotted to form a seats-vote curve (actually, a step function). Partisan bias is

simply “read off” this set of results, computed as $S^*|(V^* = .5) - .5$.

There is an elegant simplicity to this approach, taking an observed set of district-level vote shares \mathbf{v} and shifting them by the constant δ . The observed distribution of district level vote shares observed in a given election is presumed to hold under *any* election we might observe under the redistricting plan, save for the shift given by the uniform swing term δ .

5.3 Critiques of partisan bias

Among political scientists, the uniform swing approach was criticized for its determinism. Swings are never exactly uniform across districts. There are many permutations of observed vote shares that generate a statewide vote share of 50% other than simply shifting observed district-level results by a constant factor. A less deterministic approach to assessing partisan bias was developed over a series of papers by Gary King and Andrew Gelman in the early 1990s (e.g., [Gelman and King, 1990](#)). This approach fits a statistical model to district-level vote shares — and, optionally, utilizing available predictors of district-level vote shares — to model the way particular districts might exhibit bigger or smaller swings than a given level of state-wide swing. Perhaps one way to think about the approach is that it is “approximate” uniform swing, with statistical models fit to historical election results to predict and bound variation around a state-wide average swing. The result is a seats-vote curve and an estimate of partisan bias that comes equipped with uncertainty measures, reflecting uncertainty in the way that individual districts might plausibly deviate from the state-wide average swing yet still produce a state-wide average vote of 50%.

The King and Gelman model-based simulation approaches remain the most sophisticated methods of generating seats-votes curves, extrapolating from as little as one election to estimate a seats-votes curve and hence an estimate of partisan bias. Despite the technical sophistication with which we can estimate partisan bias, legal debate has centered on a more fundamental issue, the *hypothetical* character of partisan bias itself. Recall that partisan bias is defined as “seats in excess of 50% *had the jurisdiction-wide vote split 50-50.*” The premise that $V = .5$ is the problem, since this will almost always be a counter-factual or hypothetical scenario. The further V is away from $.5$ in a given election, the

counter-factual we must contemplate (when assessing the partisan bias of a districting plan) becomes all the more speculative.

In no small measure this is a marketing failure, of sorts. Partisan bias (at least under the uniform swing assumption) is essentially a measure of skew or asymmetry in *actual* vote shares. Partisan bias garners great rhetorical and normative appeal by directing attention to what happens at $V = .5$; it seems only “fair” that if a party wins 50% or more of the vote it should expect to win a majority of the districts.

Yet this distracts us from the fact that *asymmetry* in the distribution of vote shares across districts is the key, operative feature of a districting plan, and the extent to which it advantages one party or the other. Critically, we need not make appeals to counter-factual, hypothetical elections in order to assess this asymmetry.

6 The Efficiency Gap

The efficiency gap (*EG*) is also an asymmetry measure, as we see below. But unlike partisan bias, the interpretation of the efficiency gap is *not* explicitly tied to any counter-factual election outcome. In this way, the efficiency gap provides a way to assess districting plans that is free of the criticisms that have stymied the partisan bias measure.

Stephanopoulos and McGhee (2015) derive the *EG* measure with the concept of wasted votes. A party only needs $v_i = 50\% + 1$ of the votes to win district i . Anything more are votes that could have been deployed in other districts. Conversely, votes in districts where the party doesn’t win are “wasted,” from the perspective of generating seats: any districts with $v_i < .5$ generate no seats.

Wasted votes get at the core of what partisan gerrymandering is, and how it operates. A gerrymander against party A creates a relatively small number of districts that “lock up” a lot of its votes (“packing” with $v_i > .5$) and a larger number of districts that disperse votes through districts won by party B (“cracking” with $v_i < .5$). To be sure, both parties are wasting votes. But partisan advantage ensues when one party is wasting fewer votes than the other, or, equivalently, more efficiently translating votes into seats. Note also how the efficiency gap measure is also closely tied to asymmetry in the distribution of v_i .

Some notation will help make the point more clearly. If $v_i > .5$ then party A wins the district and $s_i = 1$; otherwise $s_i = 0$. The efficiency gap is defined by McGhee (2014, 68) as “relative wasted votes” or

$$EG = \frac{W_B}{n} - \frac{W_A}{n}$$

where

$$W_A = \sum_{i=1}^n s_i(v_i - .5) + (1 - s_i)v_i$$

is the sum of wasted vote proportions for party A and

$$W_B = \sum_{i=1}^n (1 - s_i)(.5 - v_i) + s_i(1 - v_i)$$

is the sum of wasted vote proportions for party B and n is the number of districts in the jurisdiction. If $EG > 0$ then party B is wasting more votes than A , or A is translating votes into seats more efficiently than B ; if $EG < 0$ then the converse, party A is wasting more votes than B and B is translating votes into seats more efficiently than A .

6.1 The efficiency gap when districts are of equal size

Under the assumption of equally sized districts McGhee (2014, 80) re-expresses the efficiency gap as:

$$EG = S - .5 - 2(V - .5) \tag{1}$$

recalling that $S = n^{-1} \sum_{i=1}^n s_i$ is the proportion of seats won by party A and $V = n^{-1} \sum_{i=1}^n v_i$ is the proportion of votes won by party A .

The assumption of equally-sized districts is especially helpful for the analysis reported below, since the calculation of EG in a given election then reduces to using the jurisdiction-level quantities S and V as in equation 1. For the analysis of historical election results reported below, it isn’t possible to obtain measures of district populations, meaning that we really have no option other than to rely on the jurisdiction-level quantities S and V when estimating the EG .

I operationalize V as the average (over districts) of the Democratic share of the two-party vote, in seats won by either a Democratic or Republican candidate;

this set of seats includes uncontested seats, where I will use imputation procedures to estimate two-party vote share. If districts are of equal size (and ignoring seats won by independents and minor party candidates) then this average over districts will correspond to the Democratic share of the state-wide, two-party vote.

6.2 The seats-vote curve when the efficiency gap is zero

This simple expression for the efficiency gap implies that *if the efficiency gap is zero*, we obtain a particular type of seats-votes curve, shown in Figure 4:

1. the seats-votes curve runs through the 50-50 point. If the jurisdiction wide vote is split 50-50 between party *A* and party *B* then with an efficiency gap of zero, $S = .5$.
2. conditional on $V = .5$ (an even split of the vote), the efficiency gap is the same as partisan bias: $V = .5 \iff EG = S - .5$, the seat share for party *A* in excess of 50%. That is, the efficiency gap reduces to partisan bias *under the counter-factual scenario* $V = .5$ that the partisan bias measure requires us to contemplate. On the other hand, the efficiency gap is not premised on that counter-factual holding, or any other counter-factual for that matter; the efficiency gap summarizes the distribution of observed district-level vote shares v_j .
3. the seats-votes curve is linear through the 50-50 point with a slope of 2. That is, with $EG = 0$, $S = 2V - .5$. Or, with a zero efficiency gap, each additional percentage point of vote share for party *A* generates *two* additional percentage points of seat share. A zero efficiency gap does not imply proportional representation (a seats-votes that is simply a 45 degree line).
4. a party winning 25% or less of the jurisdiction-wide vote should win zero seats under a plan with a zero efficiency gap; a party winning 75% or more of the jurisdiction-wide vote should win all of the seats under a plan with a zero efficiency gap. This is a consequence of the “2-to-1” seats/vote ratio and the symmetry implied by a zero efficiency gap. A party that wins an extremely low share of the vote ($V < .25$) can only be winning any seats if it enjoys an efficiency advantage over its opponent.

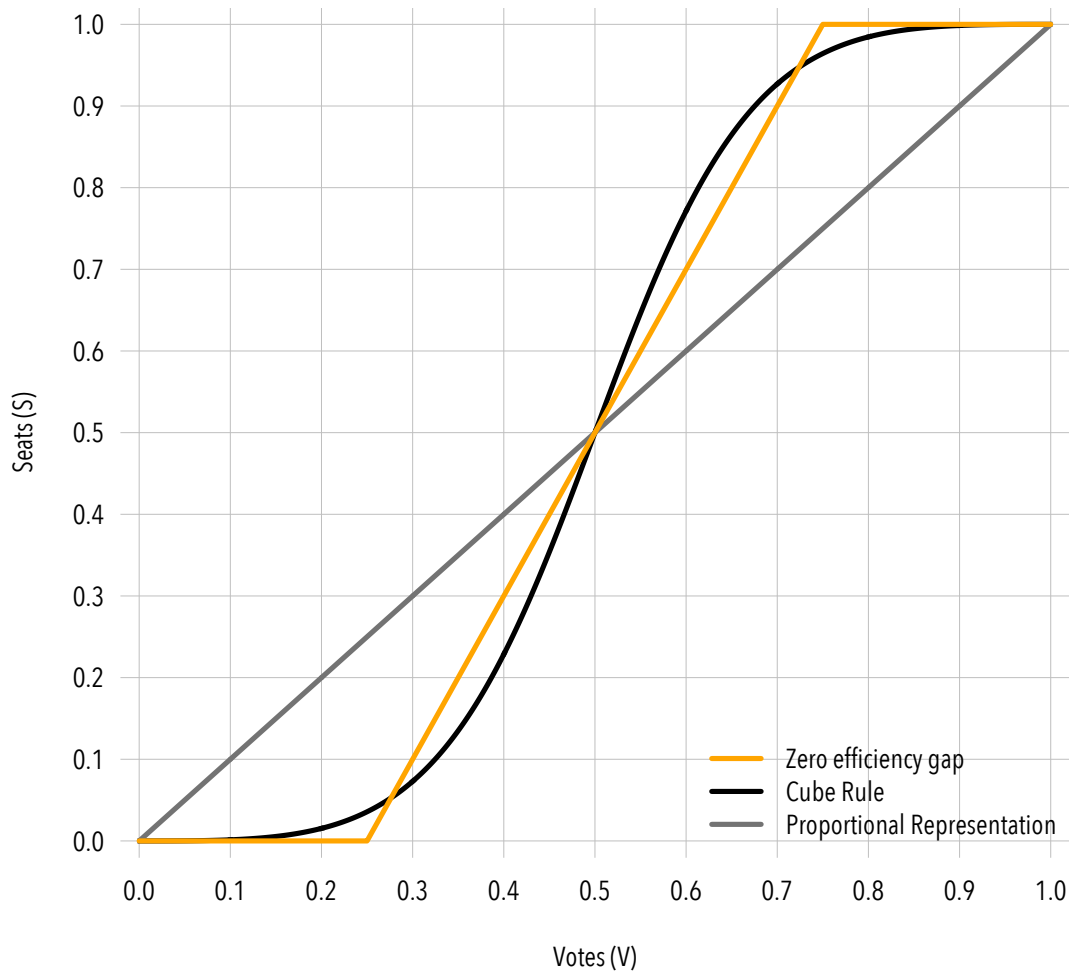


Figure 4: Theoretical seats-votes curves. The $EG = 0$ curve implies that (a) a party winning less than $V = .25$ jurisdiction-wide should not win any seats; (b) symmetrically, a party winning more than $V = .75$ jurisdiction-wide should win all the seats; and (c) the relationship between seat shares S and vote shares V over the interval $V \in [.25, .75]$ is a linear function with slope two (i.e., for every one percentage point gain in vote share, seat share should go up by two percentage points).

Moreover, the efficiency gap is trivial to compute once we have V and S for a given election. We don't need a sequence of elections under a plan in order to compute EG , nor do we need to anchor ourselves to a counter-factual scenario such as $V = .5$ as we do when computing partisan bias. For any given observed V , the hypothesis of zero efficiency gap tells us what level of S to expect.

6.3 The efficiency gap as an excess seats measure

In this sense the efficiency gap can be interpreted even more simply as an “excess seats” measure. Recall that $EG = 0 \iff S = 2V - .5$. In a given election we observe $EG = S - .5 - 2(V - .5)$. The efficiency gap can be computed by noting how far the observed S lies above or below the orange line in Figure 4.

A positive EG means “excess” seats for party A relative to a zero efficiency gap standard given the observed V in that election; conversely, a negative EG mean a deficit in seats for party A relative to a zero efficiency gap standard given the observed V .

7 State legislative elections, 1972-2014

We estimate the efficiency gap in state legislative elections over a large set of states and districting plans, covering the period 1972 to 2014. We begin the analysis in 1972 for two primary reasons: (a) state legislative election returns are harder to acquire prior to the mid-1960s, and not part of the large, canonical data collection we rely on (see below); and (b) districting plans and sequences of elections from 1972 onwards can be reasonably considered to be from the post-malappportionment era.

For each election we recover an estimate of the efficiency gap based on the election results actually observed in that election. To do this, I compute two quantities for each election:

1. V , the statewide share of the two-party vote for Democratic candidates, formed by averaging the district-level election results v_i (the Democratic share of the two-party vote in district i) in seats won by major party candidates, including uncontested seats, and

2. S , the Democratic share of seats won by major parties.

Recall that these quantities are the inputs required when computing the efficiency gap (equation 1).

The analysis that follows relies on a data set widely used in political science and freely available from the Inter-University Consortium for Political and Social Research ([ICPSR study number 34297](#)). The release of the data I utilize covers state legislative election results from 1967 to 2014, updated by Karl Klarner (Indiana State University and Harvard University). I subset the original data set to general election results since 1972 in states whose lower houses are elected via single-member districts, or where single-member districts are the norm. Multi-member districts “with positions” are treated as if they are single-member districts.

Figure 5 provides a graphical depiction of the elections that satisfy the selection criteria described above.

- Arizona, Idaho, Louisiana, Maryland, Nebraska, New Hampshire, New Jersey, North Dakota and South Dakota all drop out of the analysis entirely, because of exceedingly high rates of uncontested races, using multi-member districts, non-partisan elections, or the use of a run-off system (Louisiana).
- Alaska, Hawaii, Illinois, Indiana, Kentucky, Maine, Minnesota, Montana, North Carolina, Vermont, Virginia, West Virginia and Wyoming do not supply data over the entire 1972-2014 span; this is sometimes due to earlier elections being subject to exceedingly high rates of uncontestedness, the use of multi-member districts or non-partisan elections.
- Alabama and Mississippi have four-year terms in their lower houses, contributing data at only half the rate of the vast bulk of states with two-year legislative terms.
- Twenty-three states supply data every two years from 1972 to 2014, including Michigan and Wisconsin.
- Data is more abundant in recent decades. For the period 2000 to 2014, 41 states contribute data to the analysis at two or four year intervals.

In summary, the data available for analysis span 83,269 district-level state legislative contests, from 786 elections across 41 states.

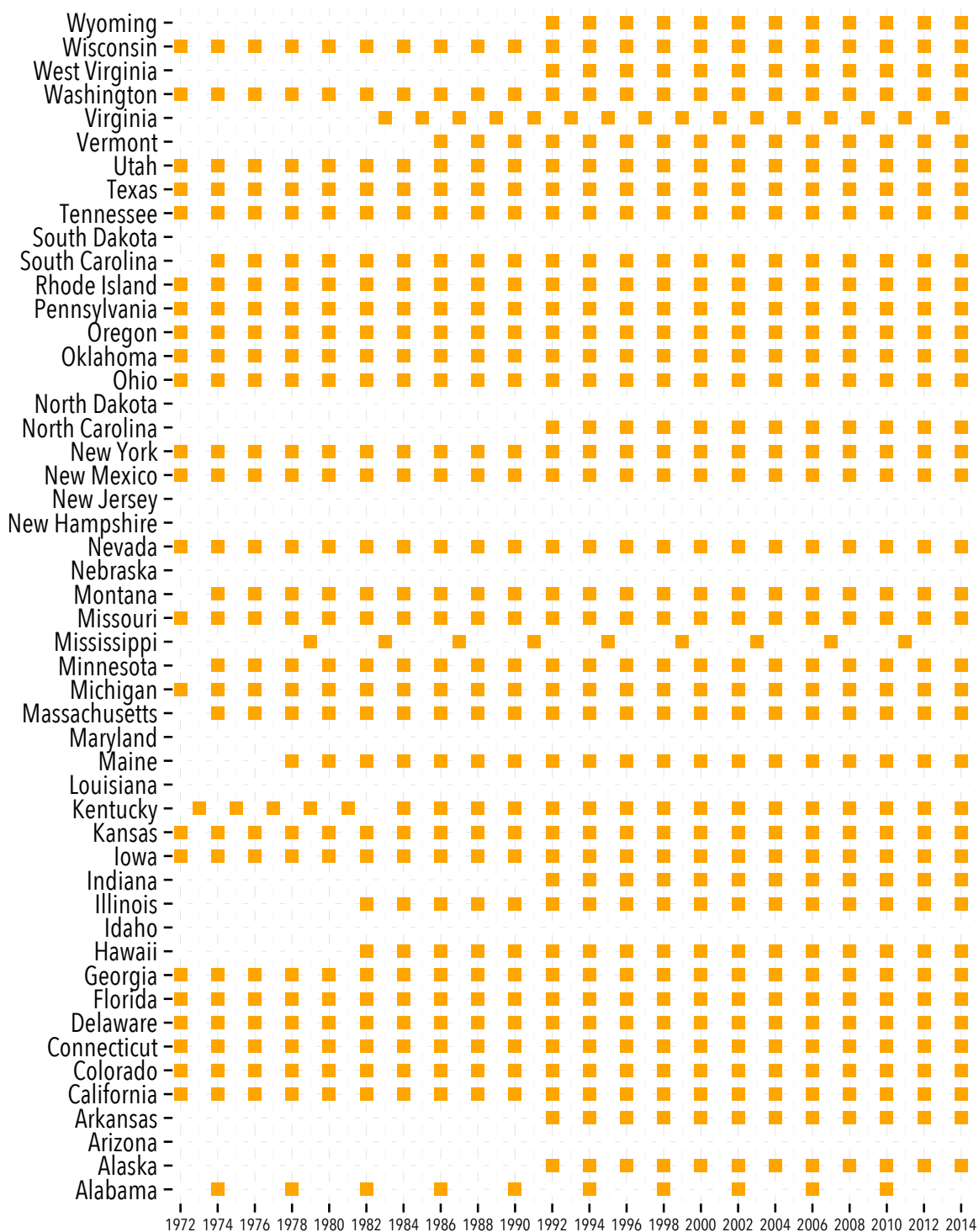


Figure 5: 786 state legislative elections available for analysis, 1972-2014, by state.

7.1 Grouping elections into redistricting plans

Districting plans remain in place for sequences of elections. An important component of my analysis involves tracking the efficiency gap across a series of elections held under the same districting plan. A key question is how much variation in the *EG* do we observe *within* districting plans, versus variation in the *EG between* districting plans.

To the extent that the *EG* is a feature of a districting plan per se, we should observe a small amount of within-plan variation relative to between plan variation. To perform this analysis we must group sequences of elections within states by the districting plan in place at the time.

[Stephanopolous and McGhee \(2015\)](#) provide a unique identifier for the districting plan in place for each state legislative election, for which I adopt here.

Figure 6 displays how the elections available for analysis group by districting plan. Districts are typically redrawn after each decennial census; the first election conducted under new district boundaries is often the “2” election (1982, 1992, etc). Occasionally we see just one election under a plan: examples include Alabama 1982, California, Hawaii 1982, Tennessee 1982, Ohio 1992, South Carolina 1992, North Carolina 2002, and South Carolina 2002.

Alaska, Kentucky, Pennsylvania and Texas held just one election under their respective districting plans adopted after the 2010 Census. In each of those states a different plan was in place for 2014 state legislative elections. Alabama’s state legislature has a four year term and we observe only the 2014 election under its post-2010 plan. The last election from Mississippi was in 2011 and was held under the plan in place for its 2003 and 2007 elections.

7.2 Uncontested races

Uncontested races are common in state legislative elections, and are even the norm in some states. For 38.7% of the district-level results in this analysis, it isn’t possible to directly compute a two-party vote share (v_i), either because the seat was uncontested or not contested by both a Democratic and Republican candidate, or (in a tiny handful of cases) the data are missing.

In some states, for some elections, the proportion of uncontested races is so high that we drop the election from the analysis. As noted earlier, examples

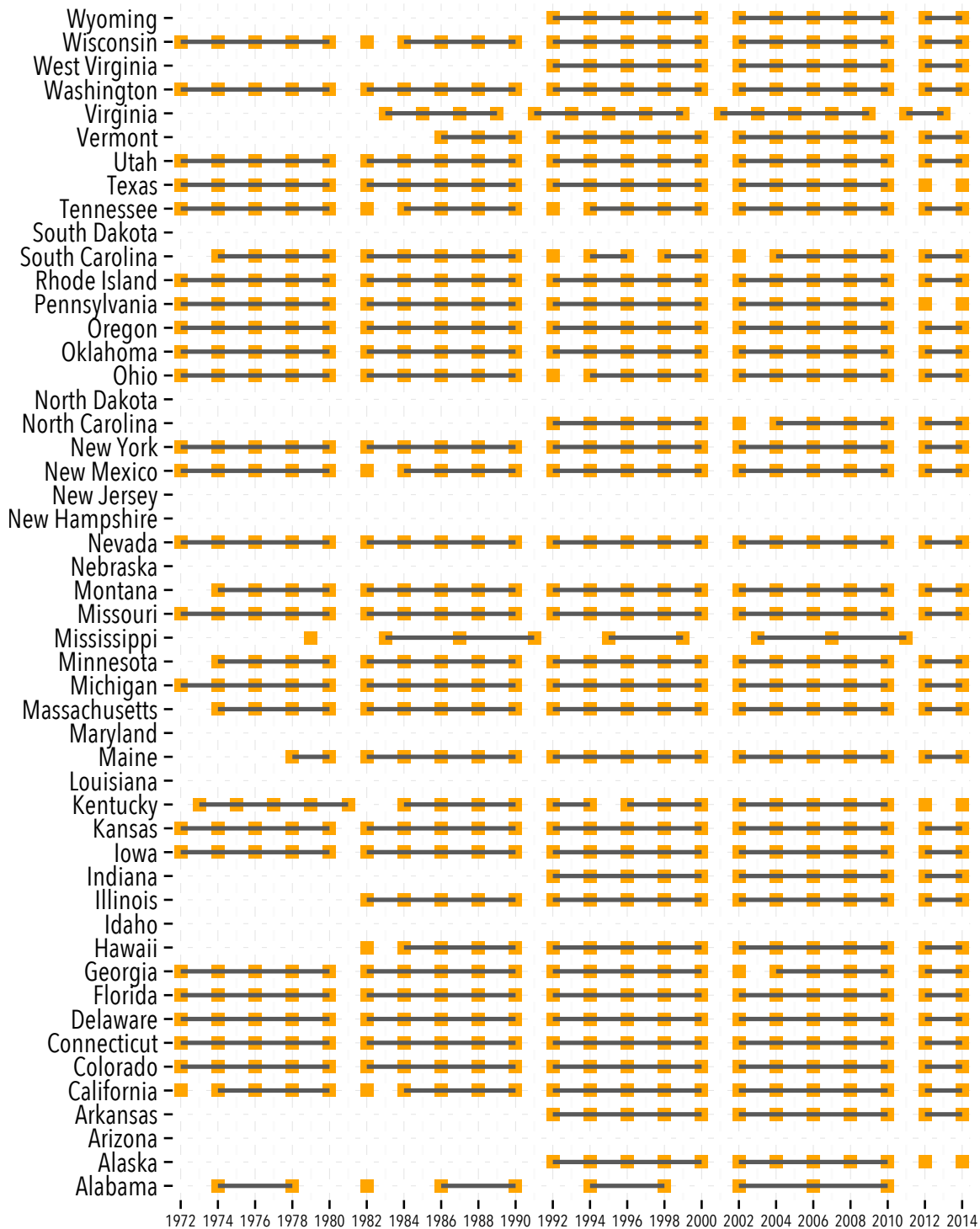


Figure 6: 786 state legislative elections available for analysis, 1972-2014, by state, grouped by districting plan (horizontal line).

include Arkansas elections prior to 1992 and South Carolina in 1972.

Even with these elections dropped from the analysis, the extent of uncontestedness in the remaining set of state legislative election results is too large to be ignored. Of the remaining elections, 31% have missing two-party results in at least half of the districts.

A graphical summary of the prevalence of uncontested districts appears in Figure 7, showing the percentage of districts without Democratic and Republican vote counts, by election and by state. Uncontested races are the norm in a number of Southern states: e.g., Georgia, South Carolina, Mississippi, Arkansas, Texas, Alabama, Virginia, Kentucky and Tennessee record rates of uncontestedness that seldom, if ever, drop below 50% for the period covered by this analysis. Wyoming also records a high proportion of districts that do not have Democratic versus Republican contests. States that lean Democratic also have high levels of uncontestedness too: see Rhode Island, Massachusetts, Illinois and, in recent decades, Pennsylvania.

Michigan and Minnesota are among the states with the lowest levels of uncontested districts in their state legislative elections. Over the set of 786 state legislative elections we examine, there are just *three* instances of elections with Democrats and Republicans running candidates in every district: Michigan supplies two of these cases (2014 and 1996) and Minnesota the other (2008).

8 Imputations for Uncontested Races

[Stephanopolous and McGhee \(2015\)](#) note the prevalence of uncontested races and report using a statistical model to impute vote shares to uncontested districts. They write:

We strongly discourage analysts from either dropping uncontested races from the computation or treating them as if they produced unanimous support for a party. The former approach eliminates important information about a plan, while the latter assumes that coerced votes accurately reflect political support.

I concur with this advice, utilizing an imputation strategy for uncontested districts with *two* distinct statistical models, predicting Democratic, two-party

Percent single-member districts without D and R candidates/vote counts, by state & election

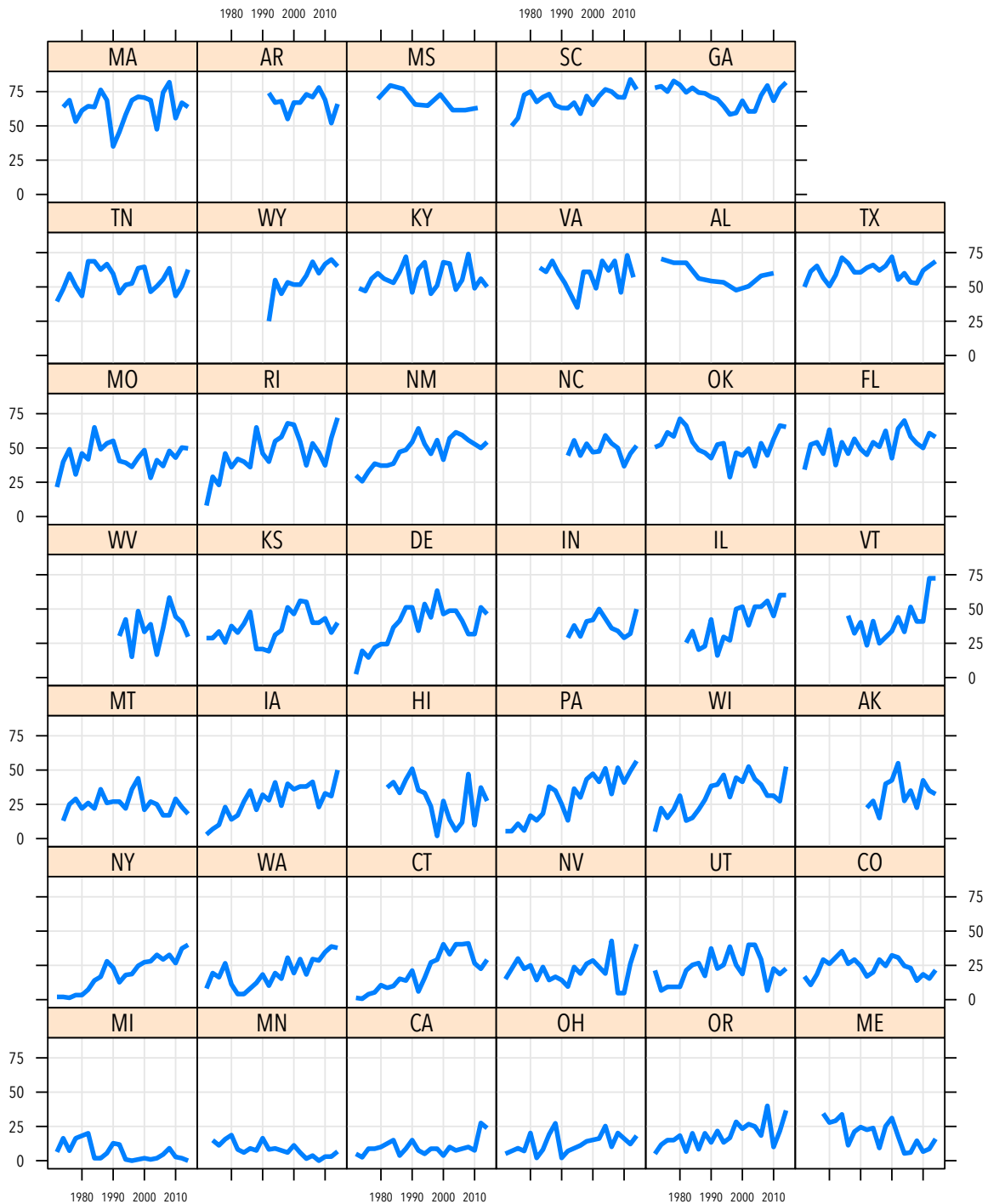


Figure 7: Percentage of districts missing two-party vote shares, by election, in 786 state legislative elections, 1972-2014. Missing data is almost always due to districts being uncontested by both major parties.

vote share in state legislative districts (v_i).

8.1 Imputation model 1: presidential vote shares

The first imputation model relies on presidential election returns reported at the level of state legislative districts. Presidential election returns are excellent predictors of state legislative election outcomes and observed even when state legislative elections are uncontested. I fit a series of linear regressions of v_i on the Democratic share of the two-party vote for president in district i , as recorded in the most temporally-proximate presidential election for which data is available and for which the current election's districting plan was in place; separate slopes and intercepts are estimated depending on the incumbency status of district i (Democratic, Open/Other, Republican).

The model also embodies the following assumptions in generating imputations for unobserved vote shares in uncontested districts. In districts where a Republican incumbent ran unopposed, we assume that the Democratic share of the two-party vote would have been less than 50%; conversely, where Democratic incumbents ran unopposed, we assume that the Democratic share of the vote would have been greater than 50%.

In most states the analysis predicts 2014 and 2012 state legislative election results v_i using 2012 presidential vote shares; 2006, 2008 and 2010 v_i is regressed on 2008 presidential vote shares, and so on. Some care is needed matching state and presidential election results in states that hold their state legislative elections in odd-numbered years, or where redistricting intervenes. In a small number of cases, presidential election returns are not available, or are recorded with district identifiers that can't be matched in the state legislative elections data. We lack data on presidential election results by state legislative district prior to 2000, so 1992 is the earliest election with which we can match state legislative election results to presidential election results at the district level.

The imputation model generally fits well. Across the 447 elections, the median r^2 statistic is 0.82. The cases fitting less well include Vermont in 2012 ($r^2 = 0.29$), with relatively few contested seats and multi-member districts with positions.

We examine the performance of the imputation model in a series of graphs, below, for six sets of elections: Wisconsin in 2012 and 2014, Michigan in 2014

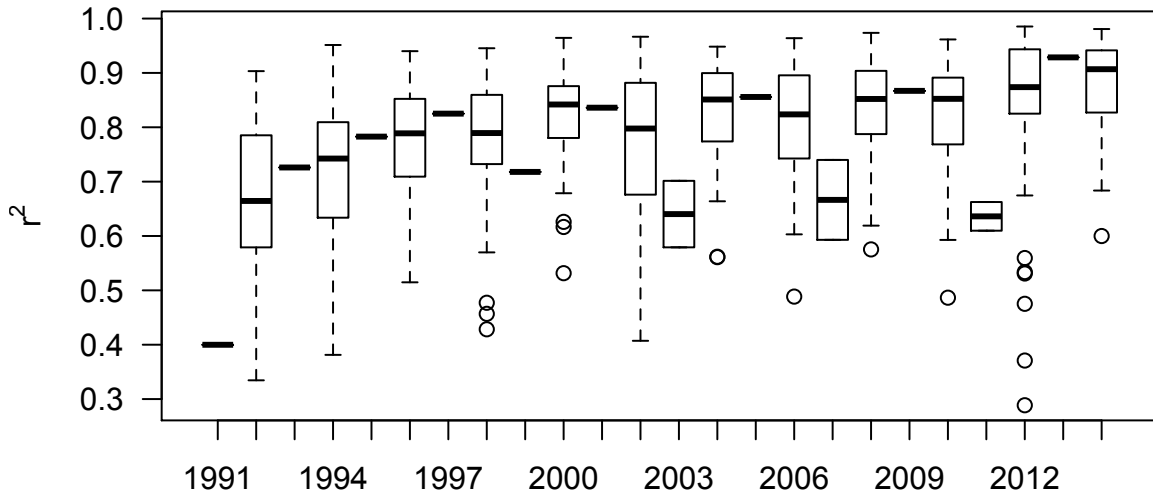


Figure 8: Distribution of r^2 statistics, regressions of Democratic share of two-party vote in state legislative election outcomes on Democratic share of the two-party for president.

(with no uncontested districts), South Carolina in 2012 (with the highest proportion of uncontested seats in the 2012 data), Virginia in 2013 and Wyoming in 2012 (the latter two generating extremely large, negative values of the efficiency gap). Vertical lines indicate 95% confidence intervals around imputed values for the Democratic share of the two-party vote in state legislative elections (vertical axis). Separate slopes and intercepts are fit for each incumbency type. Note also that the imputed data almost always lie on the regression lines.

Imputations for uncontested districts are accompanied by uncertainty. Although the imputation models generally fit well, like any realistic model they provides less than a perfect fit to the data. Note too that in any given election, there is only a finite amount of data and hence a limit to the precision with which we can make inferences about unobserved vote shares based on the relationship between observed vote shares and presidential vote shares.

Uncertainty in the imputations for v in uncontested districts generates uncertainty in “downstream” quantities of interest such as statewide Democratic vote share V and the efficiency gap measure EG . This is key, given the fact that uncontestedness is so pervasive in these data. We want any conclusions about the efficiency gap’s properties or inferences about particular levels of the efficiency gap to reflect the uncertainty resulting from imputing vote shares in uncontested districts.

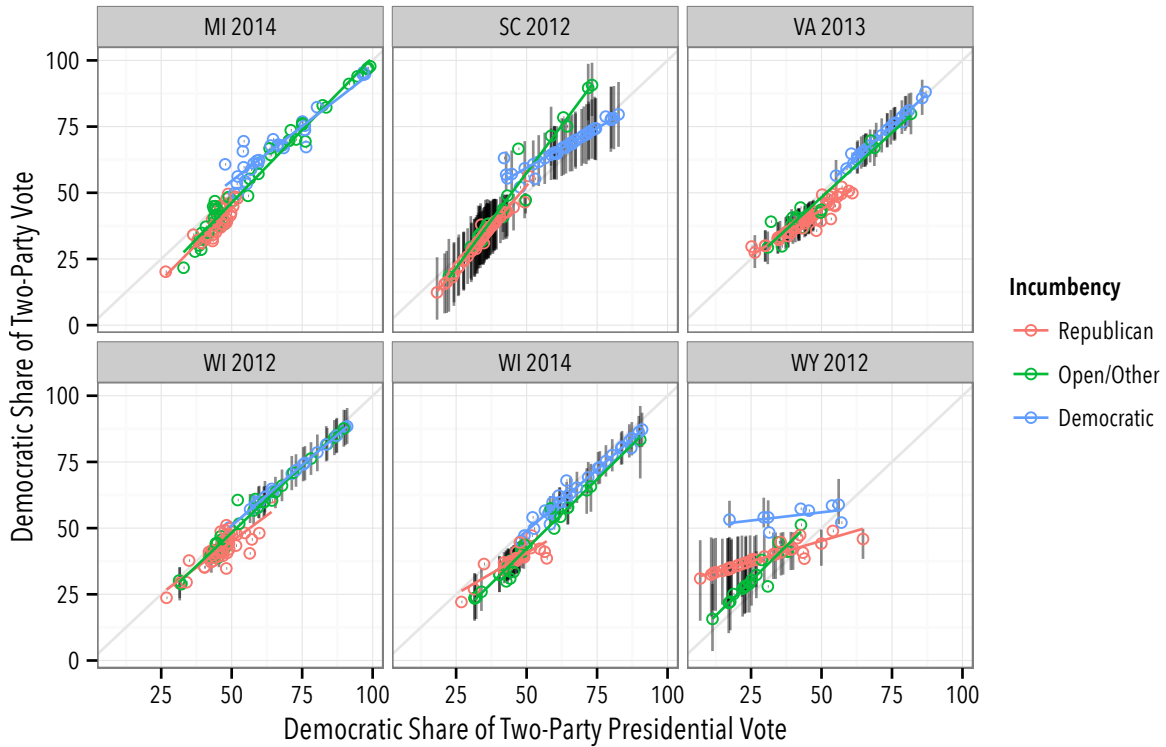


Figure 9: Regression model for imputing unobserved vote shares in 6 selected elections. Vertical lines indicate 95% confidence intervals around imputed values for the Democratic share of the two-party vote in state legislative elections (vertical axis). Separate slopes and intercepts are fit for each incumbency type. Note also that the imputed data almost always lie on the regression lines.

8.2 Imputation model 2

We rely on imputations based on presidential election returns when they are available. But presidential vote isn't always available at the level of state legislative districts (not before 1992, in this analysis). To handle these cases, we rely on a second imputation procedure, one that models sequences of election results observed under a redistricting plan, interpolating unobserved Democratic vote shares given (1) previous and future results for a given district; (2) statewide swing in a given state election; and (3) change in the incumbency status of a given district. This model also embodies the assumption that unobserved vote shares would nonetheless be consistent with what we *did* observe in a given seat: where a Democrat wins in an uncontested district, any imputation for v in that district must lie above 50%, and where a Republican wins an uncontested district, any imputation for v must lie below 50%.

8.3 Combining the two sets of imputations

We now have two sets of imputations for uncontested districts: (1) using presidential vote as a basis for imputation, where available (447 state legislative elections from 1992 to 2014); and (2) the imputation model that relies on the trajectory of district results over the history of a districting plan, including incumbency and estimates of swing, which supplies imputations for uncontested districts in all years.

When there are no uncontested districts, obviously the two imputations must agree, for the trivial reason that there are no imputations to perform. As the number of uncontested districts rises, the imputations from the two models have room to diverge. Where the two sets of imputations are available for a given election (elections where presidential vote shares by state legislative districts are available) we generally see a high level of agreement between the two methods.

The two sets of imputations for V correlate at .99. With only a few exceptions (see Figure 10), the discrepancies are generally small relative to the uncertainty in the imputations themselves. As the proportion of districts with missing data increases, clearly the scope for divergence between the two models increases.

To re-iterate, we prefer the imputations from "Model 1" based on the regressions utilizing presidential vote shares in state legislative districts, and use them

whenever available (i.e., for most states in the analysis, the period 1992-2014). We only rely on “Model 2” when presidential vote shares are not available. We model the difference between the two sets of imputations, adjusting the “Model 2” imputations of V to better match what we have obtained from “Model 1”, had the necessary presidential vote shares by state legislative district been available.

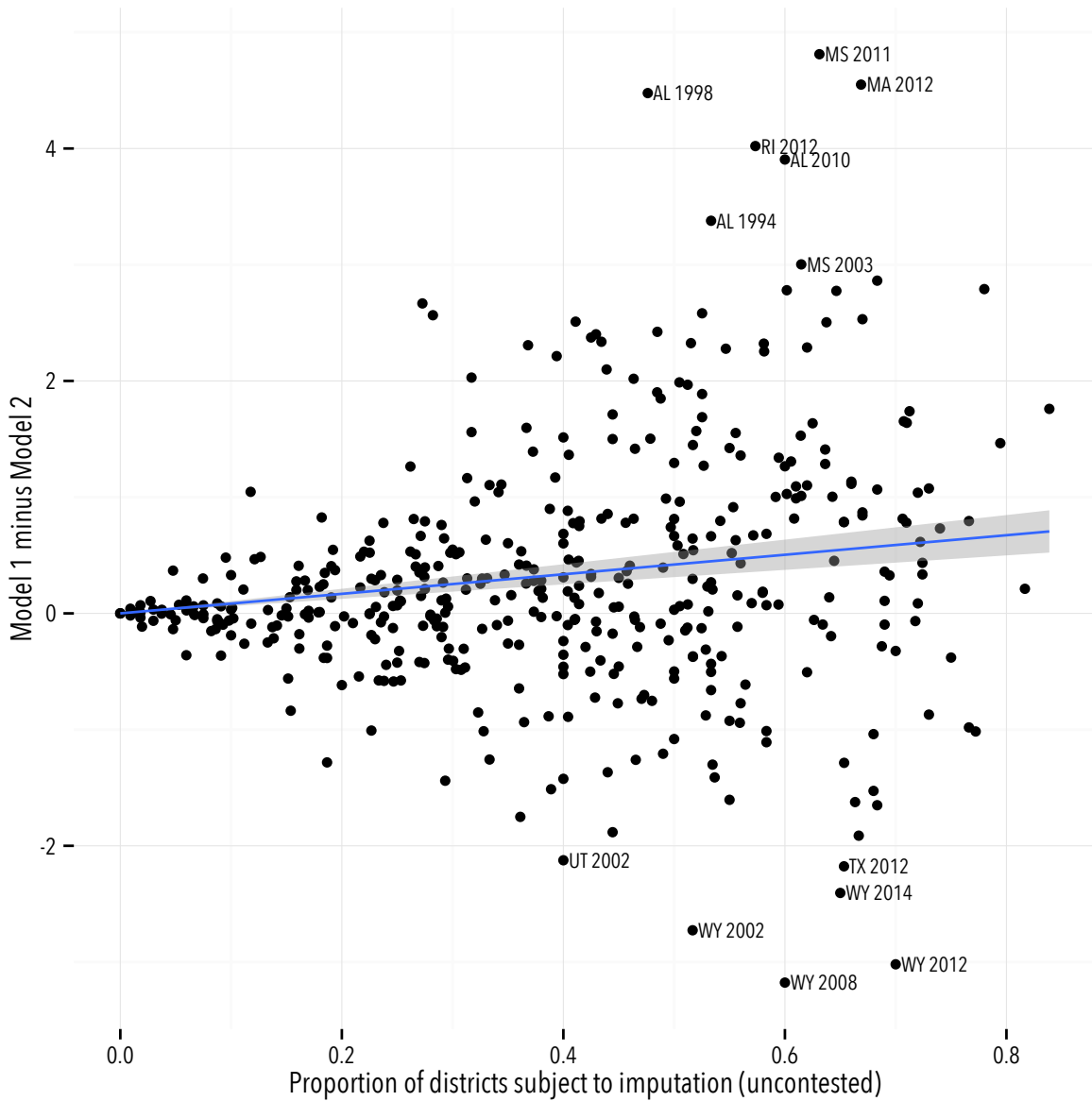


Figure 10: Difference between imputations for V by proportion of uncontested seats. The fitted regression line is constrained to respect the constraint that the imputations must coincide when there are no uncontested seats.

8.4 Seat and vote shares in 786 state legislative elections

After imputations for missing data, each election generates a seats-votes (V, S) pair. In Figure 11 we plot *all* of the V and S combinations over the 786 state elections in the analysis. We also overlay the seats-vote curve corresponding to an efficiency gap of zero. This provides us with a crude, visual sense of how often we see large departures from the zero EG benchmark.

The horizontal lines around each plotted point show the uncertainty associated with each estimate of V (statewide, Democratic, two-party vote share), given the imputations made for uncontested and missing district-level vote shares. Uncontested seats do not generate uncertainty with respect to the party winning the seat, and so the resulting uncertainty is with respect to vote shares, on the horizontal axis in Figure 11.

The efficiency gap in each election is the vertical displacement of each plotted (V, S) point from the orange, zero-efficiency gap line in Figure 11. Uncertainty as to the horizontal co-ordinate V (due to imputations for uncontested races) generates uncertainty in determining how far each point lies above or below the orange, zero efficiency gap benchmark.

9 The efficiency gap, by state and election

We now turn to the centerpiece of the analysis: assessing variation in the efficiency gap across districting plans.

We have 786 efficiency gap measures in 41 states, spanning 43 election years. These are computed by substituting each state election's estimate of V and the corresponding, observed seat share S into equation 1.

Figure 12 shows the efficiency gap estimates for each state election, grouped by state and ordered by year; vertical lines indicate 95% credible intervals arising from the fact that the imputation model for uncontested seats induces uncertainty in V and any quantity depending on V such as EG (recall equation 1). In many cases the uncertainty in EG stemming from imputation for uncontested seats is small relative to variation in EG both between and within districting plans.

We observe considerable variation in the EG estimates across states and elections. Some highlights:

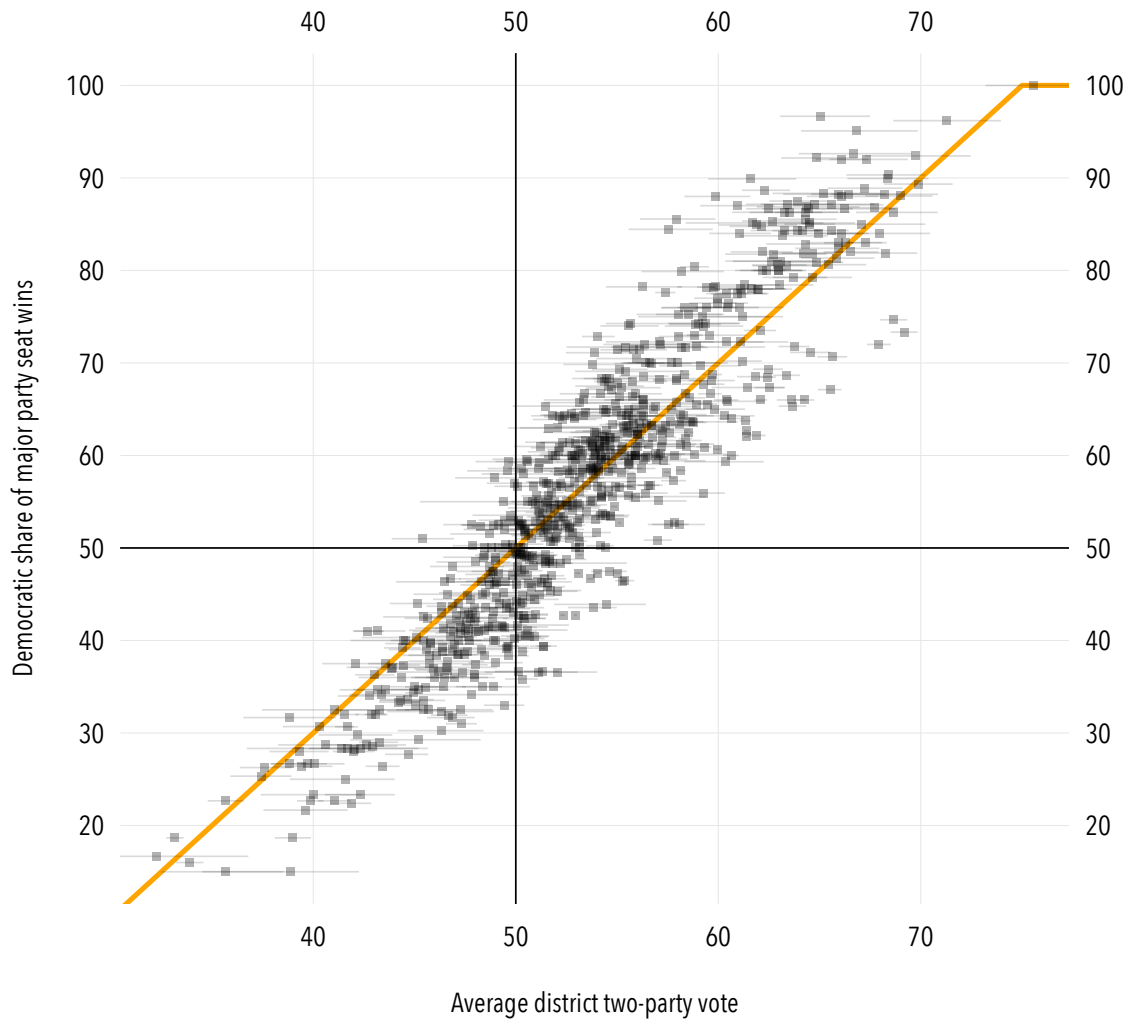


Figure 11: Democratic seat shares (S) and vote shares (V) in 786 state legislative elections, 1972-2014, in 41 states. Seat shares are defined with respect to single-member districts won by either a Republican or a Democratic candidate, including uncontested districts. Vote shares are defined as the average of district-level, Democratic share of the two-party vote, in the same set of districts used in defining seat shares. Horizontal lines indicate 95% credible intervals with respect to V , due to uncertainty arising from imputations for district-level vote shares in uncontested seats. The orange line shows the seats-votes relationship we expect if the efficiency gap were zero. Elections below the orange line have $EG < 0$ (Democratic disadvantage); points above the orange line have $EG > 0$ (Democratic advantage).

Efficiency gap, by state and year

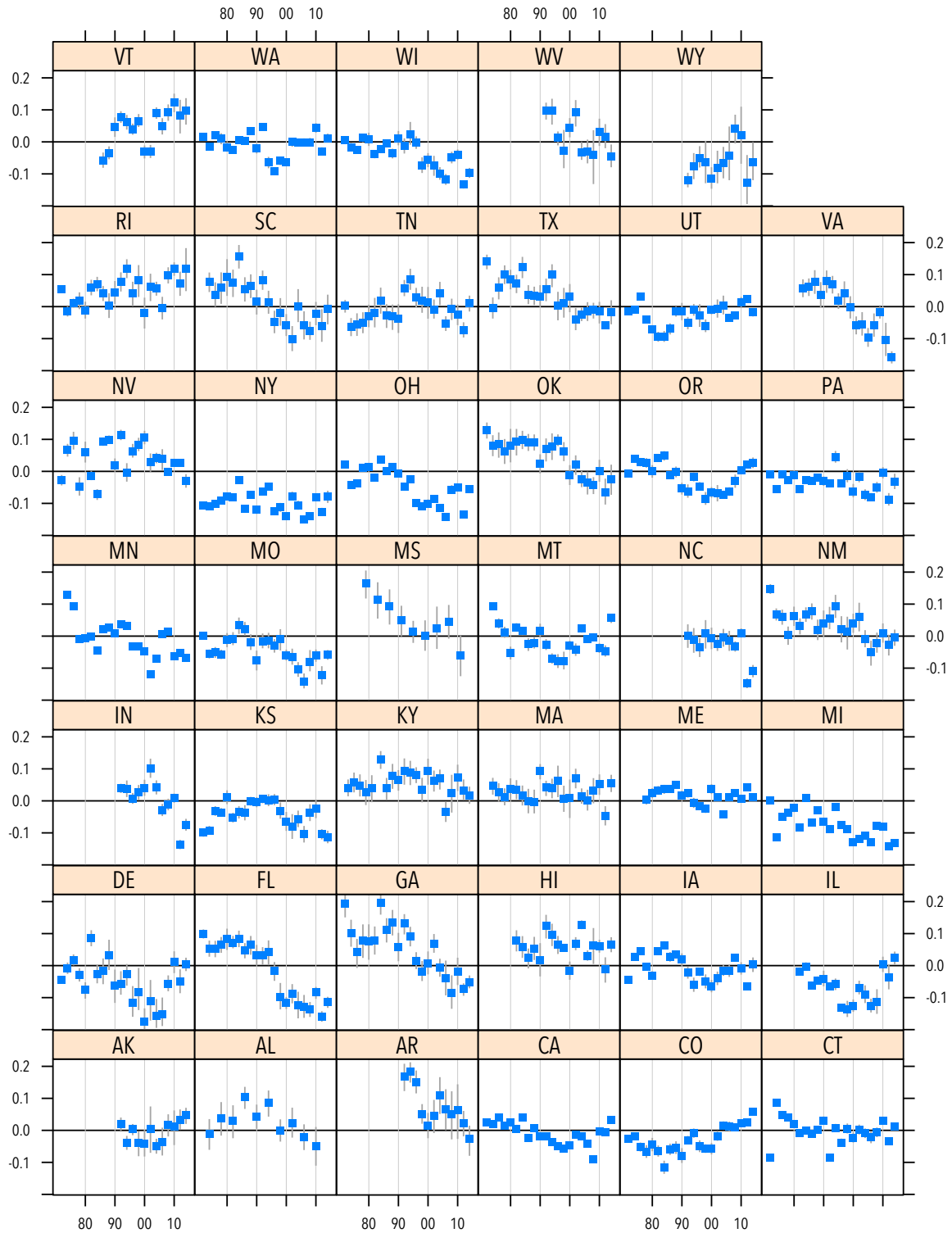


Figure 12: Efficiency gap estimates in 786 state legislative elections, 1972-2014. Vertical lines cover 95% credible intervals.

1. estimates of EG range from -0.18 to 0.20 with an average value of -0.005 .
2. The lowest value, -0.18 is from Delaware in 2000. There were 19 uncontested seats in the election to the 41 seat state legislature. Democrats won 15 seats ($S = 15/41 = 36.6\%$). I estimate V to be 52.1%. Via equation 1, this generates $EG = -0.18$. Considerable uncertainty accompanies this estimate, given the large number of uncontested seats. The 95% credible interval for V is ± 2.03 percentage points, and the 95% credible interval for the accompanying EG estimate is ± 0.04 .
3. The highest value of EG is 0.20 is from Georgia in 1984. There were 140 uncontested seats in the election to the 180 seat state legislature. Democrats won 154 seats ($S = 154/180 = 85.6\%$). I estimate V to be 57.9%. Again, using equation 1, this generates $EG = 0.2$. Considerable uncertainty also accompanies this estimate, given the large number of uncontested seats. The 95% credible interval for V is ± 1.89 percentage points, and the 95% credible interval for the accompanying EG estimate is ± 0.04 . Figure 13 contrasts the seats and votes recorded in Georgia against those for the entire data set, putting Georgia's large EG estimates in context.
4. New York has the lowest median EG estimates, ranging from -0.15 (2006) to -0.028 (1984). Statewide V ranges from 53.7% to 69.2%, but Democrats only win 70 (1972) to 112 (2012) seats in the 150 seat state legislature, so S ranges from .47 to .75, considerably below that we'd expect to see given the vote shares recorded by Democrats if the efficiency gap were zero. See Figure 15.
5. Arkansas has the highest median EG score by state, .10; see Figure 14.
6. Connecticut has the median, within-state median EG score of approximately zero; Figure 16 shows Connecticut's seats and votes have generally stayed close to the $EG = 0$ benchmark.
7. Michigan has the third lowest median EG scores by state, surpassed only by New York and Wyoming. Michigan's EG scores range from -0.14 (2012) to .01 (1984). V ranges from 50.3% to 60.6%, a figure we estimate confidently given low and occasionally even zero levels of uncontested districts

in Michigan state legislative elections. Yet S ranges from 42.7% (Democrats won 47 out of 110 seats in 2002, 2010 and 2014) to 63.6% (Democrats won 70 out of 110 seats in 1978). See Figure 17.

8. Wisconsin's EG estimates range from $-.14$ (2012) to $.02$ (1994). Although the EG estimates for WI are not very large relative to other states in other years, Wisconsin has recorded an unbroken run of negative EG estimates from 1998 to 2014 and records two very large estimates of the efficiency gap in elections held under its current plan: $-.13$ (2012) and $-.10$ (2014). In short, Democrats are underperforming in state legislative elections in Wisconsin, winning fewer seats than a zero efficiency gap benchmark would imply, given, their statewide level of support. See Figure 18.

9.1 Are efficiency gap estimates statistically significant?

Recall that $EG < 0$ means that Democrats are disadvantaged, with relatively more wasted votes than Republicans; conversely $EG > 0$ means that Democrats are the beneficiaries of an efficiency gap, in that Democrats have fewer wasted votes than Republicans. But EG does vary from election to election, even with the same districting plan in place and EG is almost always not measured perfectly, but is estimated with imputations for uncontested seats.

In Figure 19 we plot the imprecision of each efficiency gap estimate (the half-width of its 95% credible interval) against the estimated EG value itself. Points lying inside the cones have EG estimates that are small relative to their credible intervals, such that we would not distinguish them from zero at conventional levels of statistical significance. Not all EG estimates can be distinguished from zero at conventional levels of statistical significance, nor should they. But many estimates of the EG are unambiguously non-zero. Critically, the two most recent Wisconsin EG estimates ($-.13$ in 2012, $-.10$ in 2014) are clearly non-negative, lying far away from the “cone of ambiguity” shown in Figure 19; the 95% credible interval for the 2012 estimates runs from $-.146$ to $-.121$ and from $-.113$ to $-.081$ for the 2014 estimate.

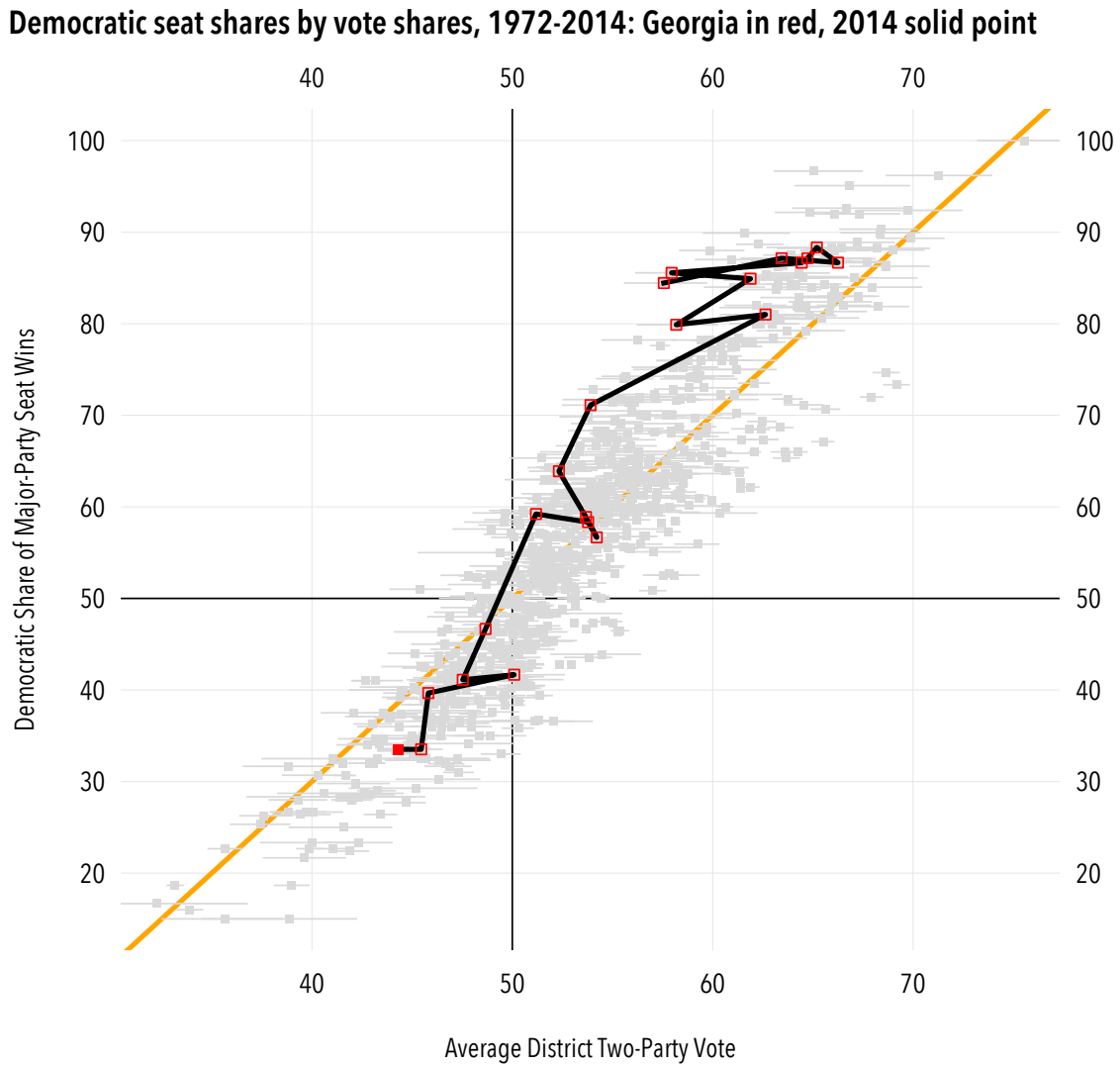


Figure 13: Georgia, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

Democratic seat shares by vote shares, 1972-2014: Arkansas in red, 2014 solid point

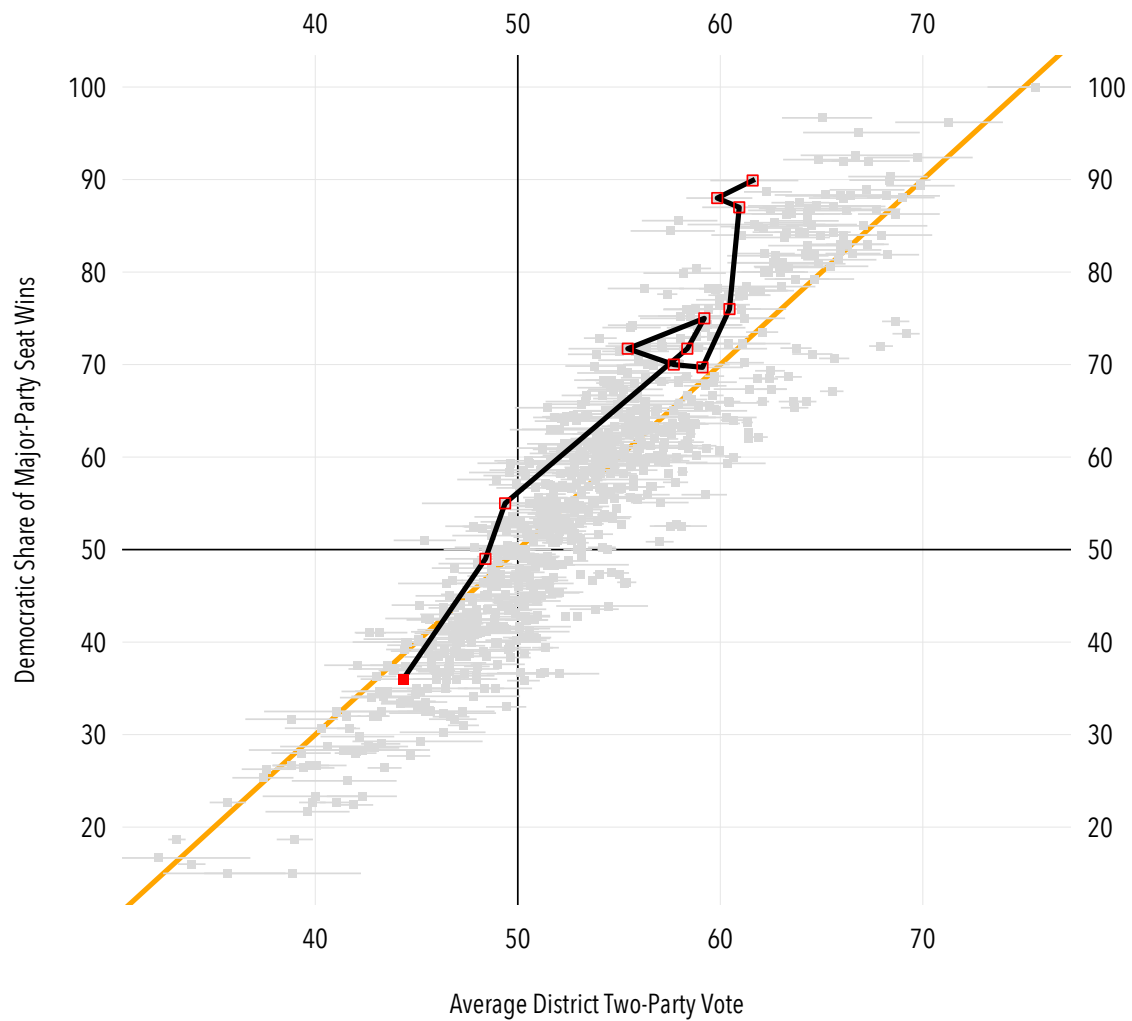


Figure 14: Arkansas, Democratic seat share and average district two-party vote share, 1992-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

Democratic seat shares by vote shares, 1972-2014: New York in red, 2014 solid point

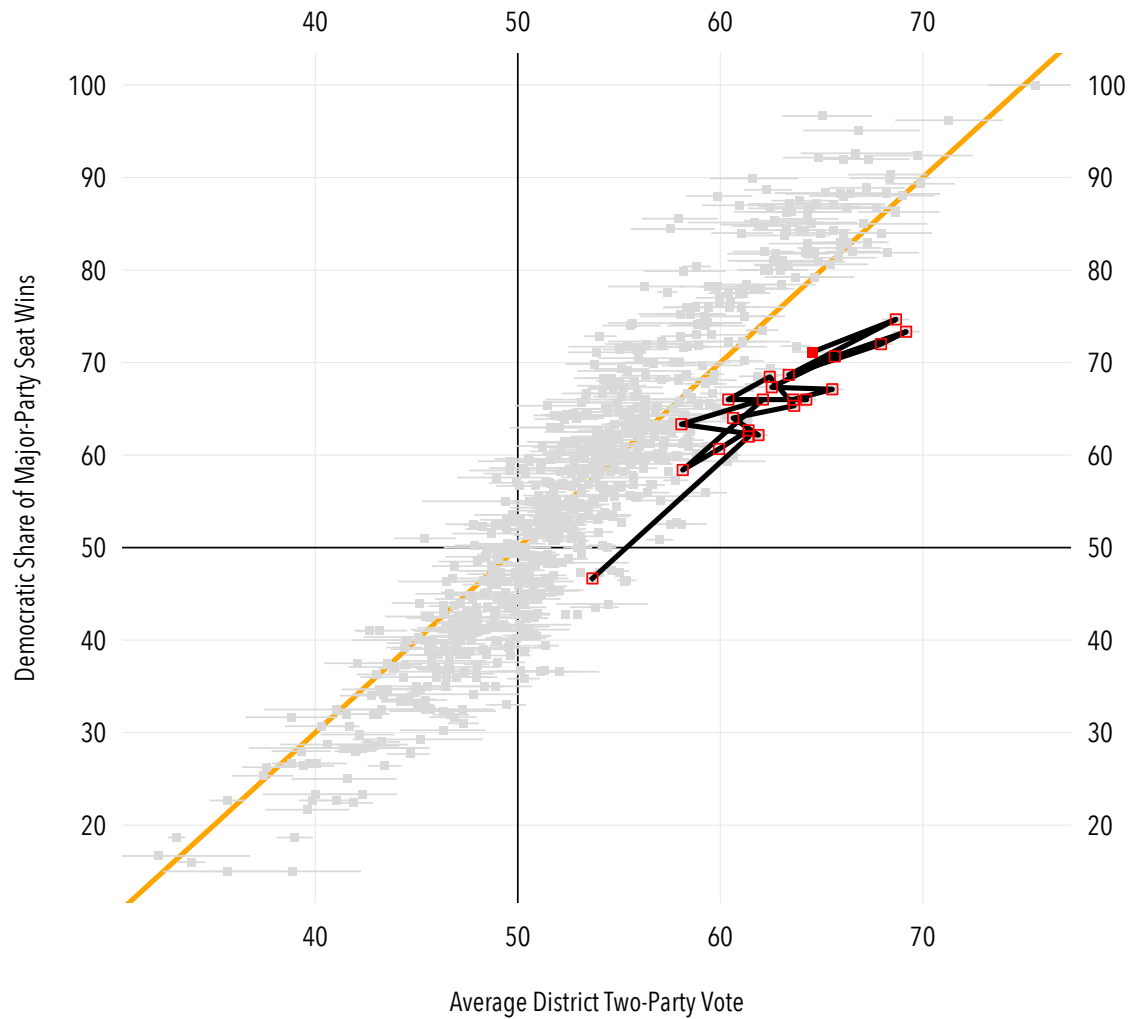


Figure 15: New York, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

Democratic seat shares by vote shares, 1972-2014: Connecticut in red, 2014 solid point

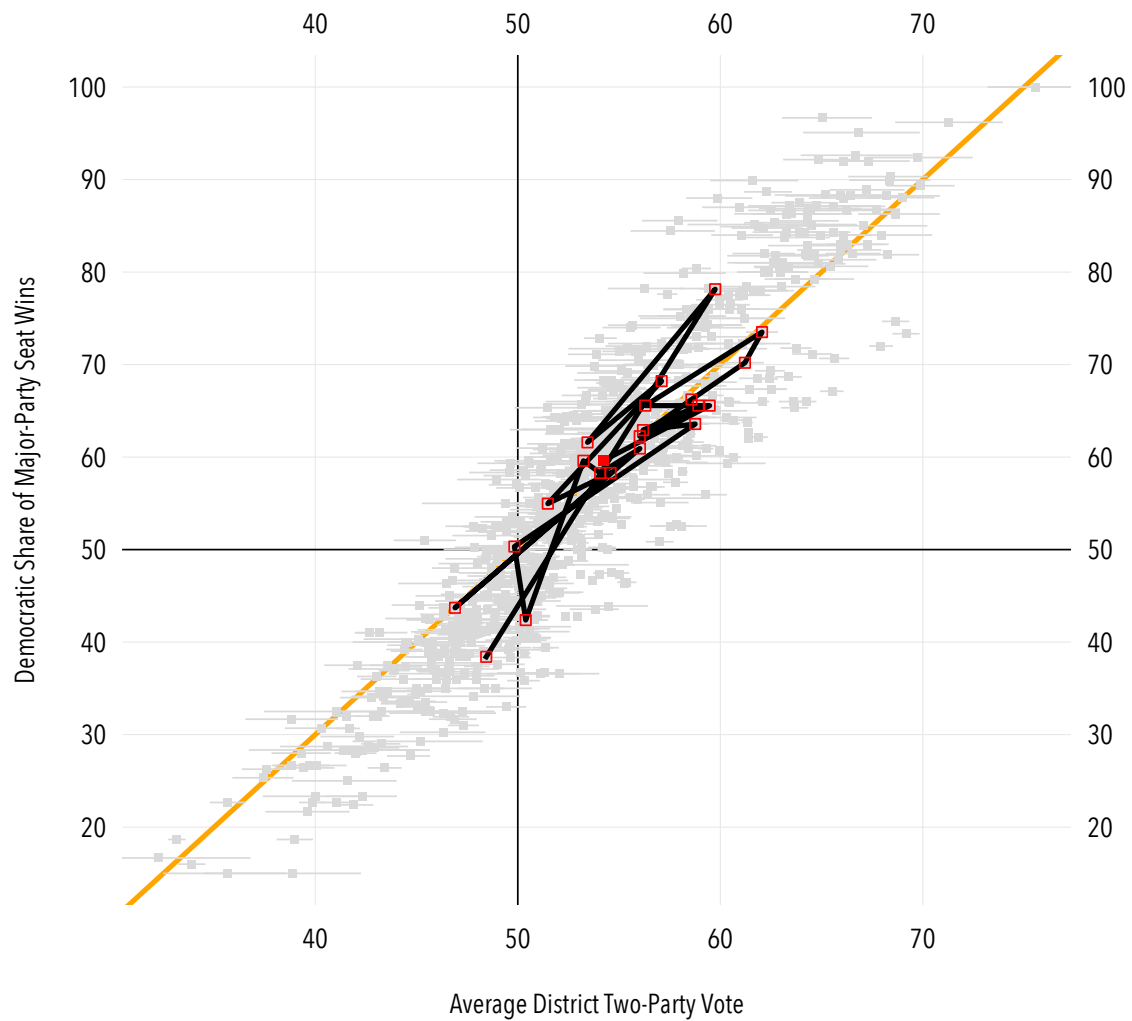


Figure 16: Connecticut, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

Democratic seat shares by vote shares, 1972-2014: Michigan in red, 2014 solid point

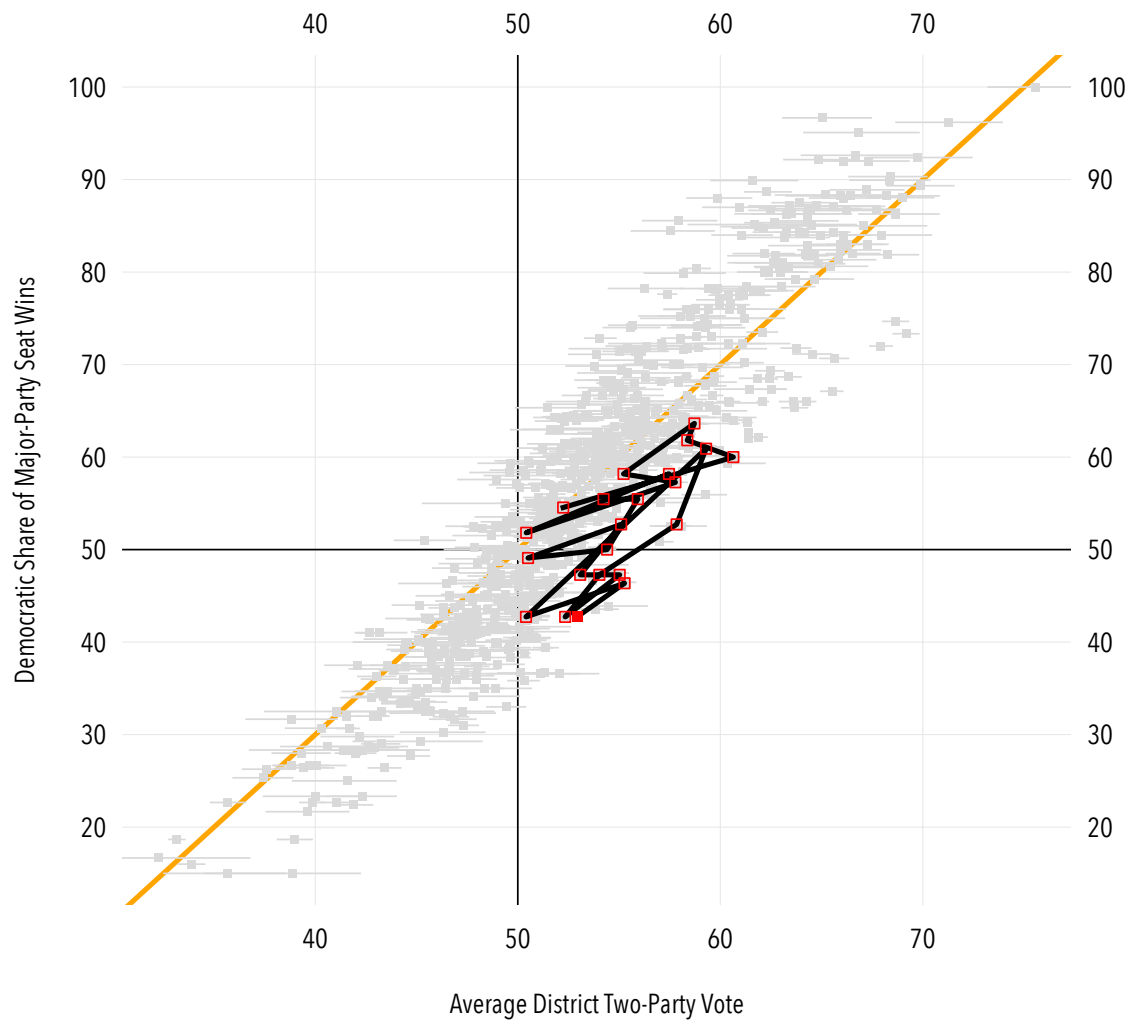


Figure 17: Michigan, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

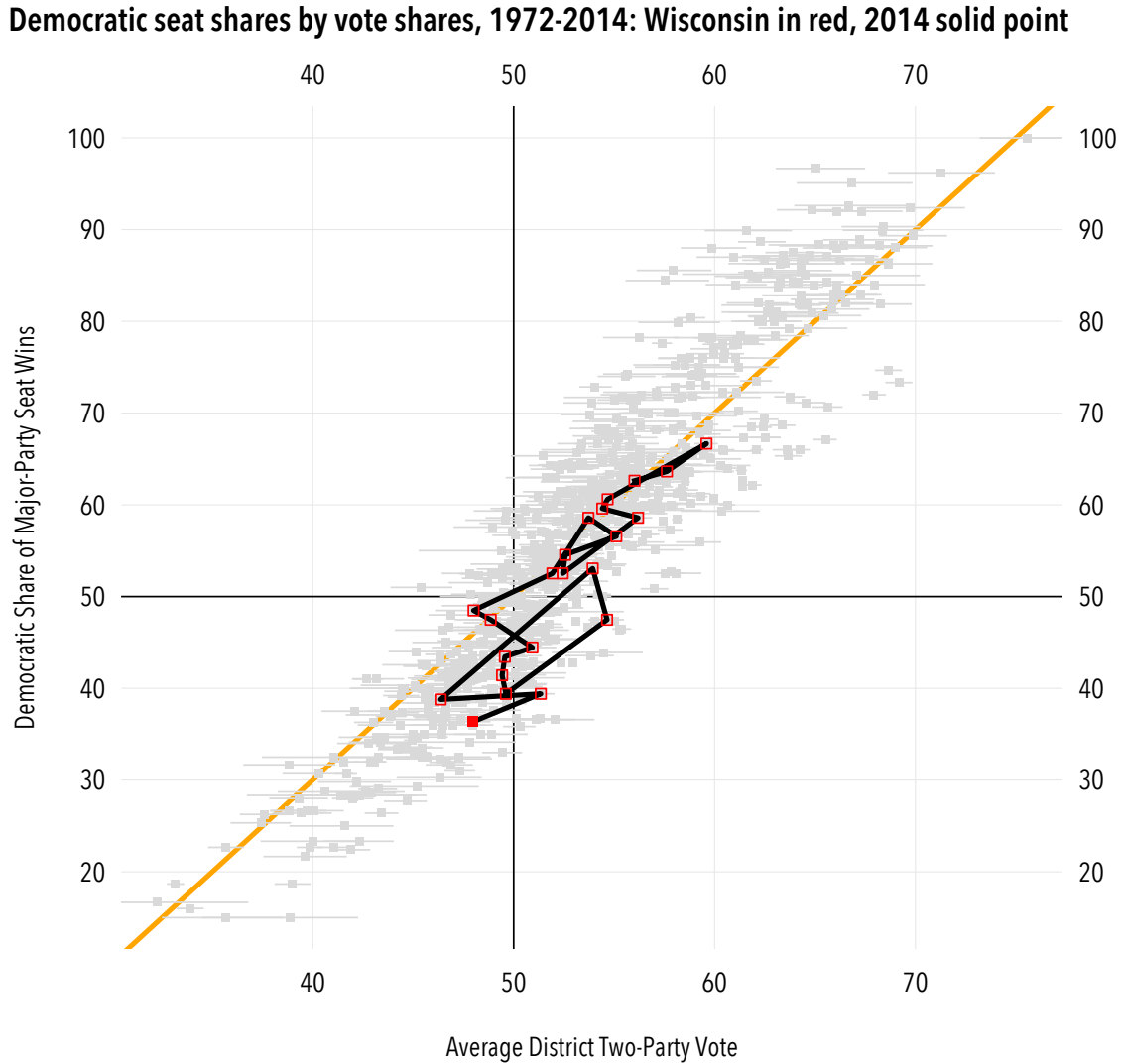


Figure 18: Wisconsin, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

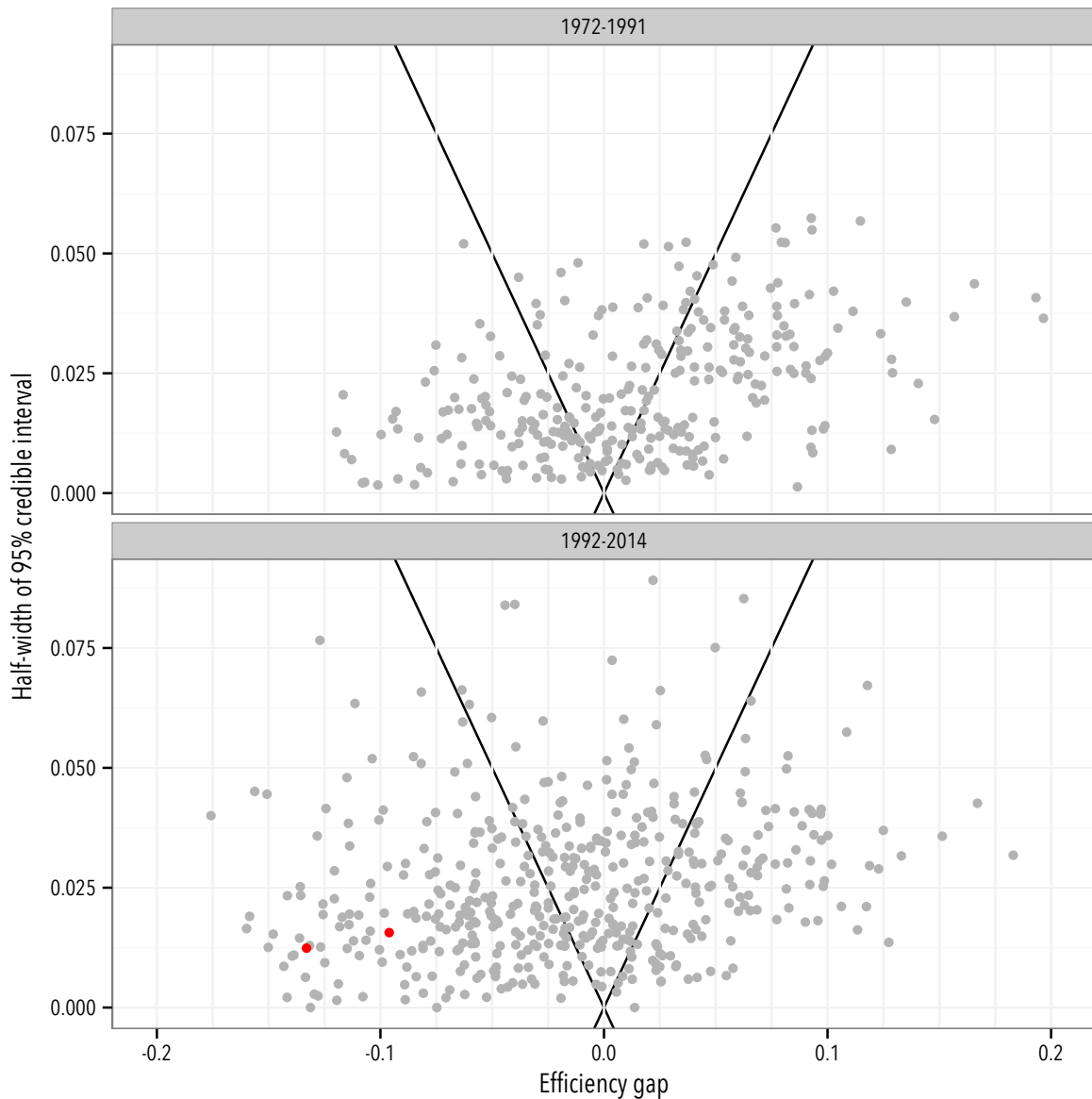


Figure 19: Uncertainty in the efficiency gap, against the *EG* estimate itself. The vertical axis is the half-width of the 95% credible interval for each *EG* estimate (plotted against the horizontal axis); points lying inside the cones have *EG* estimates that are small relative to their credible intervals, such that we would not distinguish them from zero at conventional levels of statistical significance. *EG* estimates from Wisconsin in 2012 and 2014 are shown as red points in the lower panel. Note the greater prevalence of large, negative and precisely estimated *EG* measures in recent decades.

9.2 Over-time change in the efficiency gap

Are large values of the efficiency gap less likely to be observed in recent decades? This is relevant to any discussion of a standard by which to assess redistricting plans. If recent decades have generally seen smaller values of the efficiency gap relative to past decades, then this might be informative as to how we should assess contemporary districting plans and their corresponding values of the *EG*.

Figure 20 plots *EG* estimates over time, overlaying estimates of the smoothed, weighted quantiles (25th, 50th and 75th) of the *EG* measures (the weights capture the uncertainty accompanying each estimate of the *EG*). The distribution of *EG* measures in the 1970s and 1980s appeared to slightly favor Democrats; about two-thirds of all *EG* measures in this period were positive. The distribution of *EG* measures trends in a pro-Republican direction through the 1990s, such that by the 2000s, *EG* measures were more likely to be negative (Republican efficiency advantage over Democrats); see Figure 21.

There is some evidence that the 2010 round of redistricting has generated an increase in the magnitude of the efficiency gap in state legislative elections. For most of the period under study, there seems to be no distinct trend in the magnitudes of the efficiency gap over time; see Figure 22. The median, absolute value of the efficiency gap has stayed around 0.04 over much of the period spanned by this analysis; elections since 2010 are producing higher levels of *EG* in magnitude.

It is also interesting to note that the estimate of the 75th percentile of the distribution of *EG* magnitudes jumps markedly after 2010, suggesting that districting plans enacted after the 2010 census are systematically more gerrymandered than in previous decades. Of the almost 800 *EG* estimates in the analysis, spanning 42 years of elections, the largest, negative estimates (an efficiency gap disadvantaging Democrats) are more likely to be recorded in the short series of elections after 2010. These include Alabama in 2014 (-.18), Florida in 2012 (-.16), Virginia in 2013 (-.16), North Carolina in 2012 (-.15) and Michigan in 2012 (-.14); these five elections are among the 10 least favorable to Democrats we observe in the entire set of elections. Among the 10 most pro-Democratic *EG* scores, *none* were recorded after 2000. The most favorable election to Democrats in terms of *EG* since 2010 is the 2014 election in Rhode Island ($EG = .12$), which is only the 20th largest (pro-Democratic) *EG* in the entire analysis.

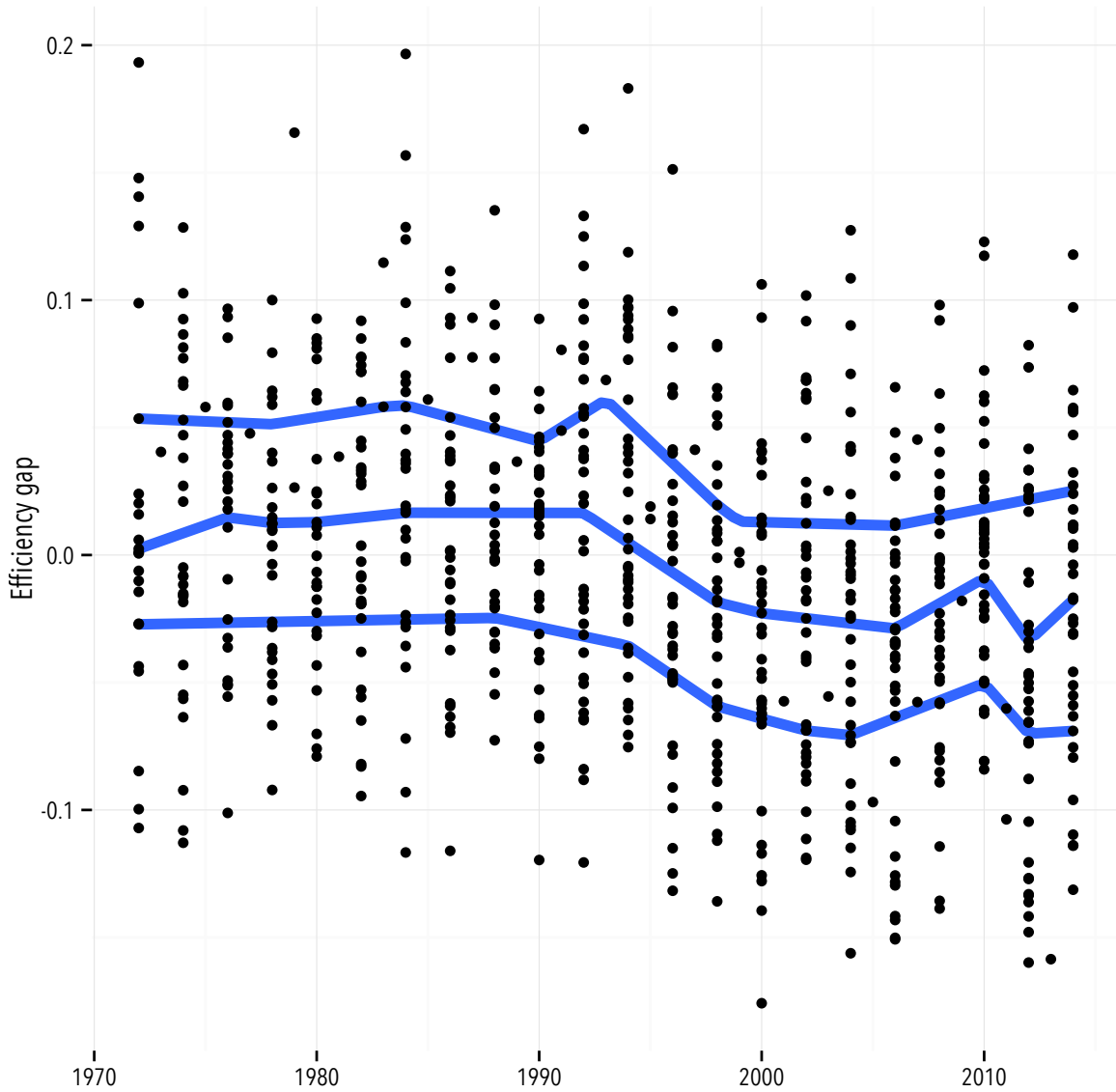


Figure 20: Efficiency gap estimates, over time. The lines are smoothed estimates of the 25th, 50th and 75th quantiles of the efficiency gap measures, weighted by the precision of each *EG* measure.

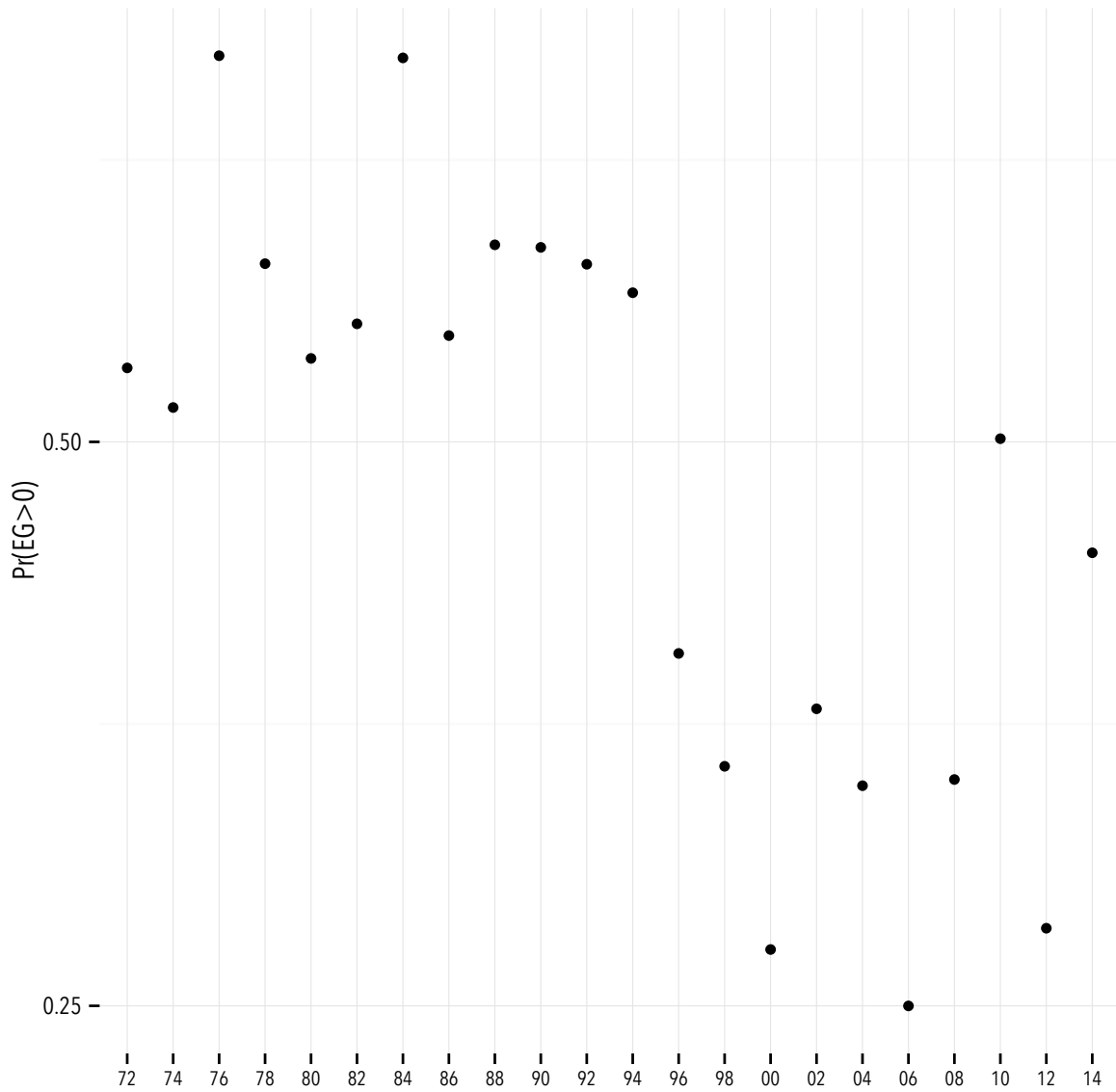


Figure 21: Proportion of efficiency gap measures that are positive, by two year intervals.

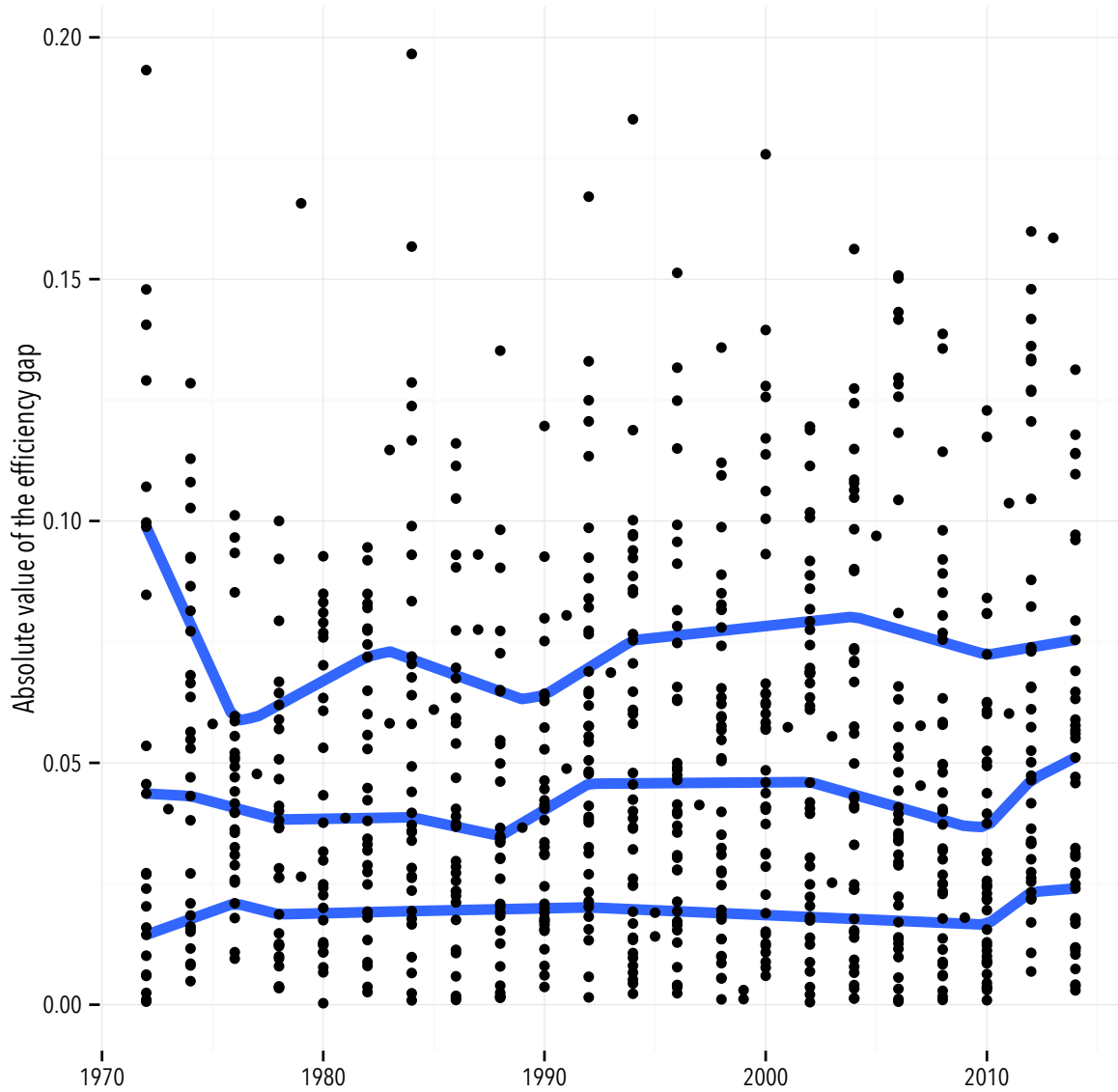


Figure 22: Absolute value of efficiency gap measures, over time. The lines are smoothed estimates of the 25th, 50th and 75th quantiles of the absolute value of the efficiency gap measure, weighted by the precision of each *EG* measure.

9.3 Within-plan variation in the efficiency gap

The efficiency gap is measured at each election, with a given districting plan typically generating up to five elections and hence five efficiency gap measures. Efficiency gap measures will change from election to election as the distribution of district-level vote shares varies over elections. Some of this variation is to be expected. Even with the same districting plan in place, districts will display “demographic drift,” gradually changing the political complexion of those districts. Incumbents lose, retire or die in office; sometimes incumbents face major opposition, sometimes they don’t. Variation in turnout — most prominently, from on-year to off-year — will also cause the distribution of vote shares to vary from election to election, even with the districting plan unchanged. All these election-specific factors will contribute to election-to-election variation in the efficiency gap.

Precisely because we expect a reasonable degree of election-to-election variation in the efficiency gap, we assess the magnitude of this “within-plan” variability in the measure. If a plan is a partisan gerrymander — with a systematic advantage for one party over the other — then the “between-plan” variation in *EG* should be relatively large relative to the “within-plan” variation in *EG*.

About 76% of the variation in the *EG* estimates is between-plan variation. The *EG* measure does vary election-to-election, but there is a moderate to strong “plan-specific” component to variation in the *EG* scores. We conclude that the efficiency gap *is* measuring an enduring feature of a districting plan.

We examine some particular districting plans. The 786 elections in this analysis span 150 districting plans. For plans with more than one election, we compute the standard deviation of the sequence of election-specific *EG* measures observed under the plan. These standard deviations range from .011 (Kentucky’s plan in place for just two elections in 1992 and 1994, or Indiana’s plan 1992-2000) to .079 (Delaware’s plan between 2002 and 2010).

A highly variable plan: Delaware 2002-2010. Figure 23 shows the seats, votes and *EG* estimates produced under the Delaware 2002-2010 plan. This is among the most variable plans we observe with respect to the *EG* measure. An efficiency gap running against the Democrats for 2002, 2004 and 2006 (the latter election saw Democrats win only 18 seats out of 41 with 54.5% of the state wide vote) falls to a small gap in 2008 ($V = 0.584, S = 25/41 = .61, EG = -0.058$) and

Delaware ends the decade with a positive efficiency gap in 2010. The Democratic district-average two-party vote share fell to $V = 0.561$ in 2010, but translated into $S = 26/41 = .63$, $EG = 0.012$.

A plan with moderate variability in the EG. The median, within-plan standard deviation of the EG is about .03. This roughly corresponds to the within-plan standard deviation of the EG observed under the plan in place for five Wisconsin state legislative elections 1992-2000, presented in Figure 24. This was a plan that generated relatively small values of EG that alternated sign over the life of the plan: negative in 1992, positive in 1994 and 1996, and negative in 1998 and 2000.

A low variance plan, Indiana 1992-2000. See Figure 25. The EG measures recorded under this plan are all relatively small and positive, ranging from 0.008 to 0.041 and correspond to an interesting period in Indiana state politics. Democrats won 55 of the 100 seats in the Indiana state house in the 1992 election with what I estimate to be just over 50% of the district-average vote (29 of 100 seats were uncontested). Democratic vote share fell to about 45% in the 1994 election (38 uncontested seats), and Democrats lost control of the legislature. The 1996 election resulted in a 50-50 split in the legislature. Democrats won legislative majorities in the 1998 and 2000 elections, while the last election might have been won by Democrats with just less than 50% of the district-vote; I estimate $V = 0.495 \pm .012$ and $EG = 0.041$.

Highlighting Delaware plan 4

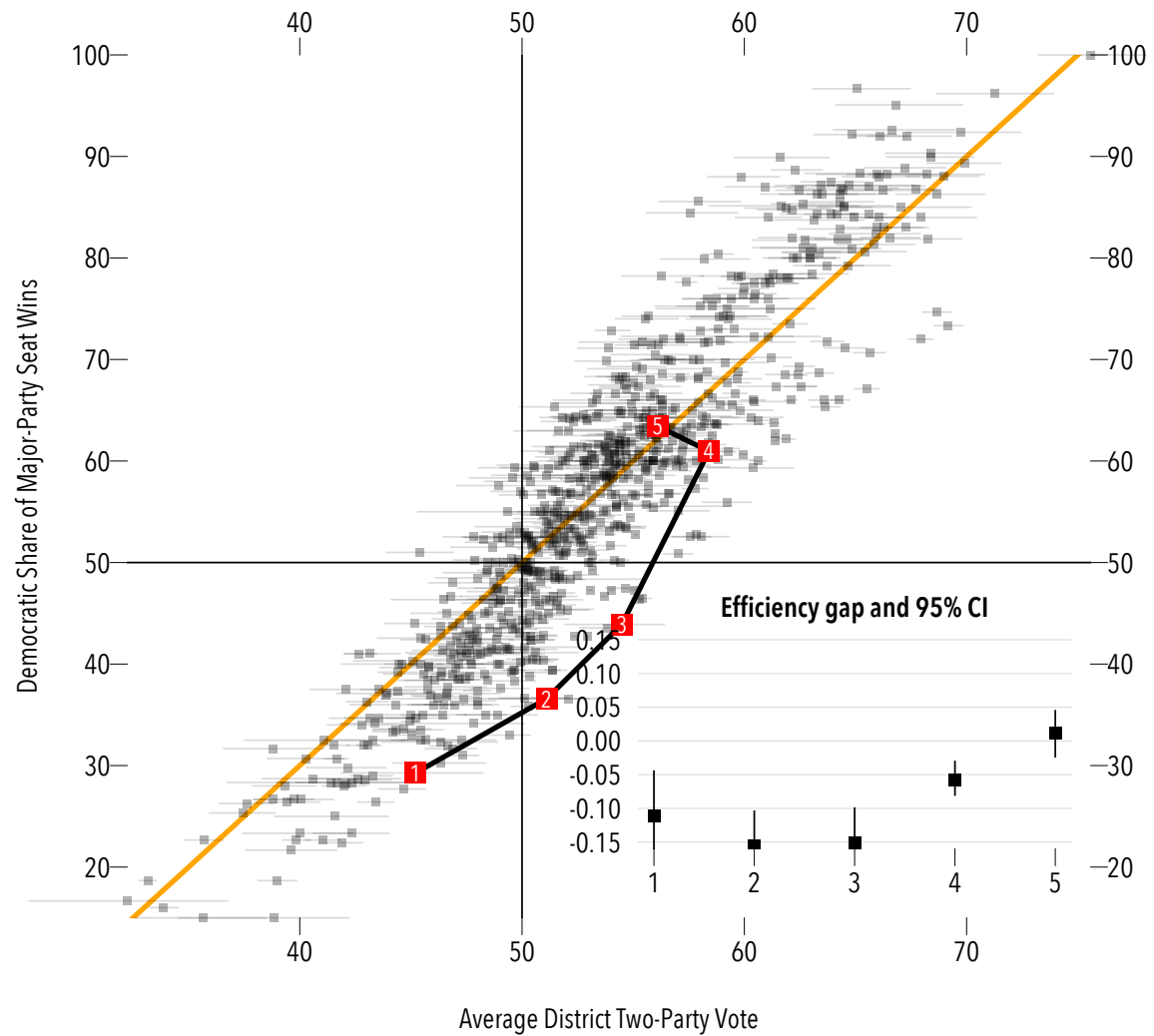


Figure 23: Seats, votes and the efficiency gap recorded under the Delaware plan, 2002-2010. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts. The inset in the lower right shows the sequence of efficiency gap measures recorded under the plan; vertical lines are 95% credible intervals.

Highlighting Wisconsin plan 3

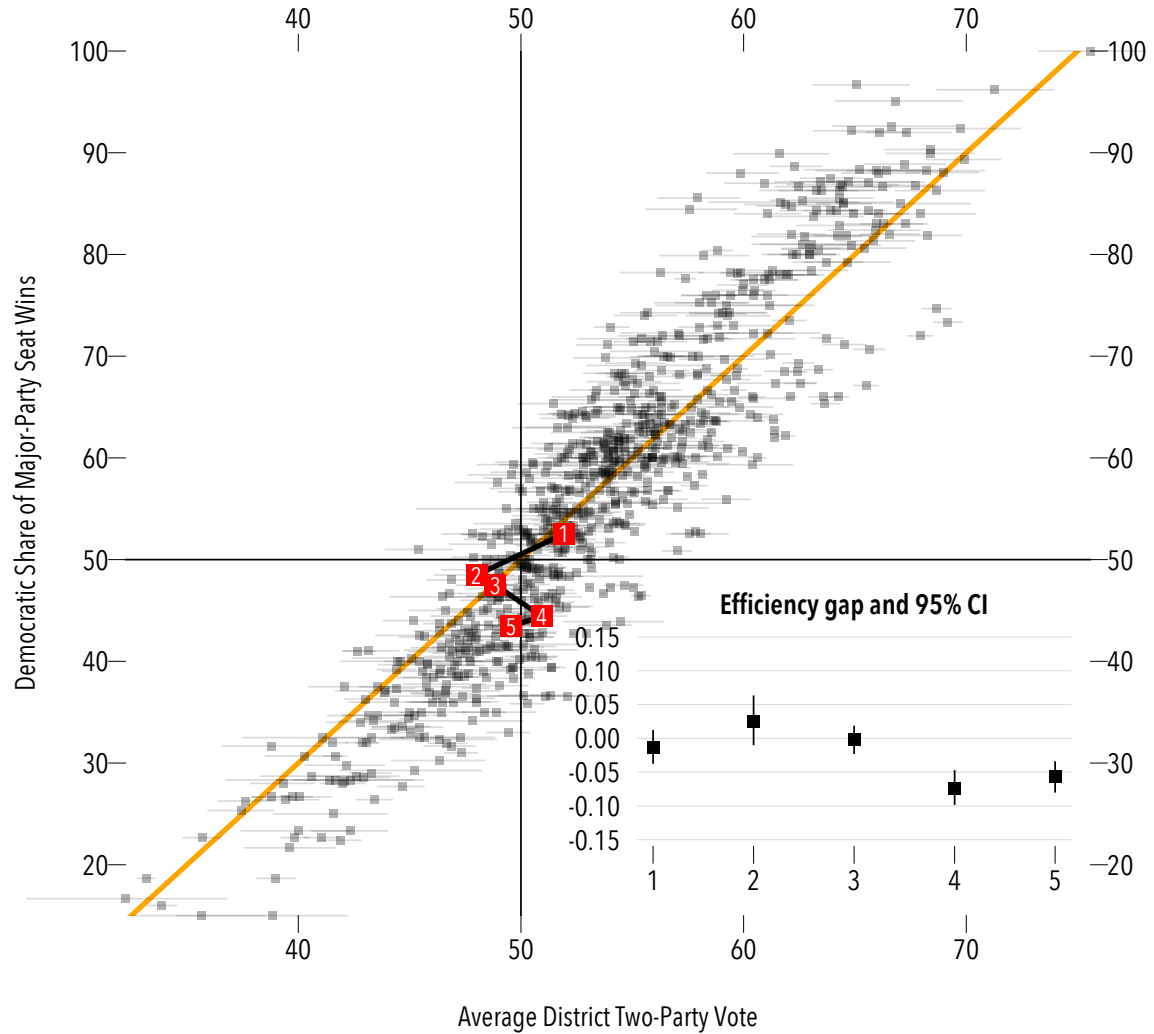


Figure 24: Seats, votes and the efficiency gap recorded under the Wisconsin plan, 1992-2000. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts. The inset in the lower right shows the sequence of efficiency gap measures recorded under the plan; vertical lines are 95% credible intervals.

Highlighting Indiana plan 3

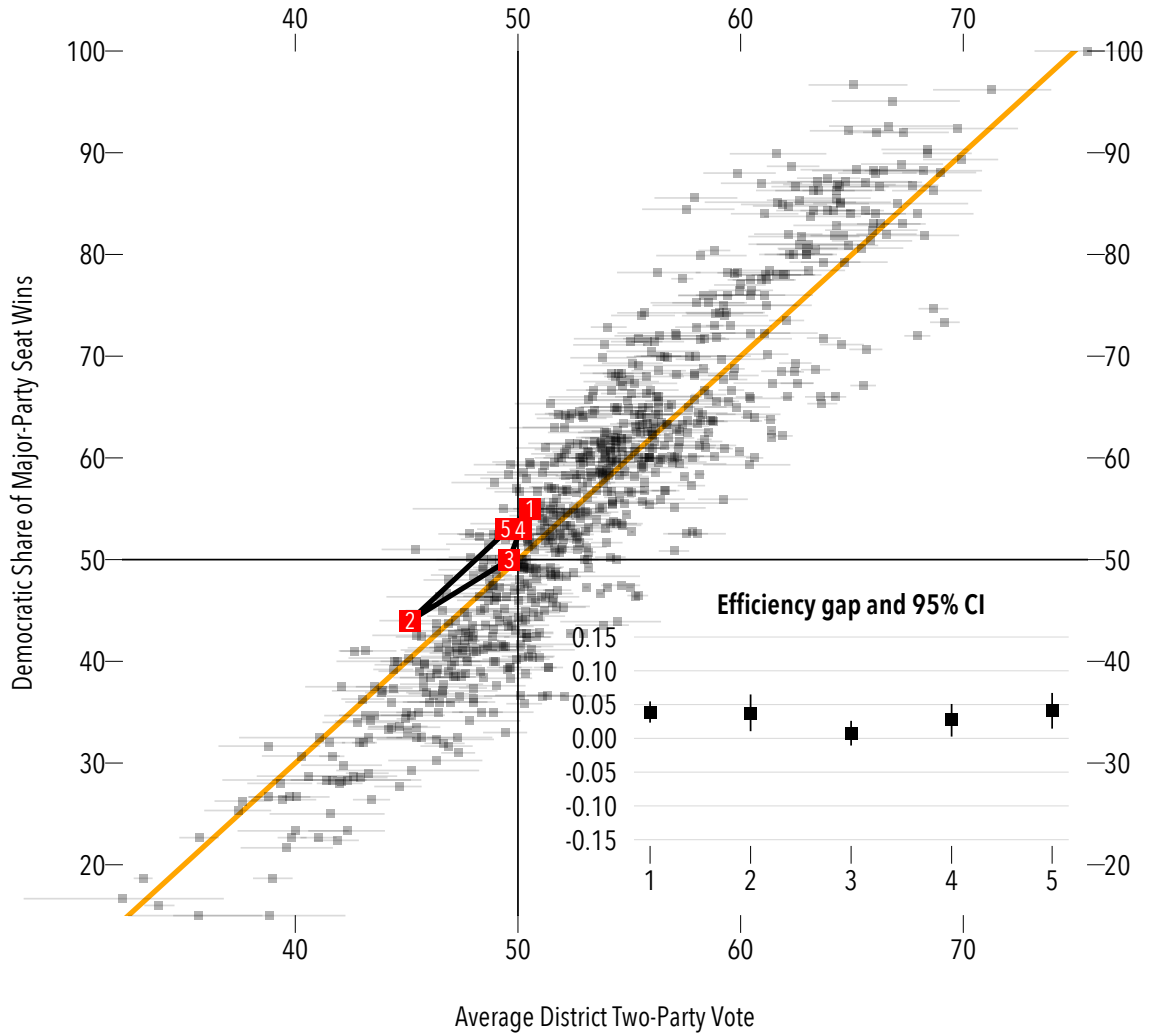


Figure 25: Seats, votes and the efficiency gap recorded under the Indiana plan, 1992-2000. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts. The inset in the lower right shows the sequence of efficiency gap measures recorded under the plan; vertical lines are 95% credible intervals.

9.4 How often does the efficiency gap change sign?

Having observed a particular value of EG , how confident are we that:

- the EG measure is distinguishable from zero at conventional levels of statistical significance? That is, how sure are we as to the sign of any particular EG estimate? We addressed this question in section 9.1.
- it will be followed by one or more estimates of EG that are of the same sign?
- over the life of a districting plan, EG remains on one side of zero or the other?

The latter two questions are key. It is especially important that we assess the *durability* of the sign of the EG measure under a districting plan, if we seek to assert that a districting plan is a partisan gerrymander. We will see that *magnitude* and *durability* of the efficiency gap go together: *large* values of the efficiency gap don't seem to be capricious, but likely to be repeated over the life of a districting plan, consistent with partisan disadvantage being a systematic feature of the plan.

We begin this part of the analysis by considering temporally adjacent *pairs* of EG estimates. Can we be confident that these have the same sign? In general, yes. Of the full set of 786 elections for which we compute an efficiency gap estimate, 580 are temporally adjacent, within state and districting plan. Figure 26 shows that we usually see efficiency gap measures with the same sign; this probability exceeds 90% for almost half of the temporally adjacent pairs of efficiency gap measures. Averaged over all pairs, this “same sign” probability is 74%. While the efficiency gap does vary election to election, these fluctuations are not so large that the *sign* of the efficiency gap is likely to change election to election.

What about over the life of an entire redistricting plan? How likely is it that the efficiency gap retains the same sign over, say, three to five elections in a given state, taking into account election-to-election variation *and* uncertainty arising from the imputation procedures used for uncontested districts?

We have 141 plans that supply three or more elections with estimate of the efficiency gap. Of these, 17 plans are *utterly unambiguous* with respect to the sign of the efficiency gap estimates recorded over the life of the plan: for each of these plans we estimate the probability that the EG has the same sign over the life of the plan to be 100%. These plans are listed below in Table 1.

Probabilities that efficiency gap has the same sign as in previous election

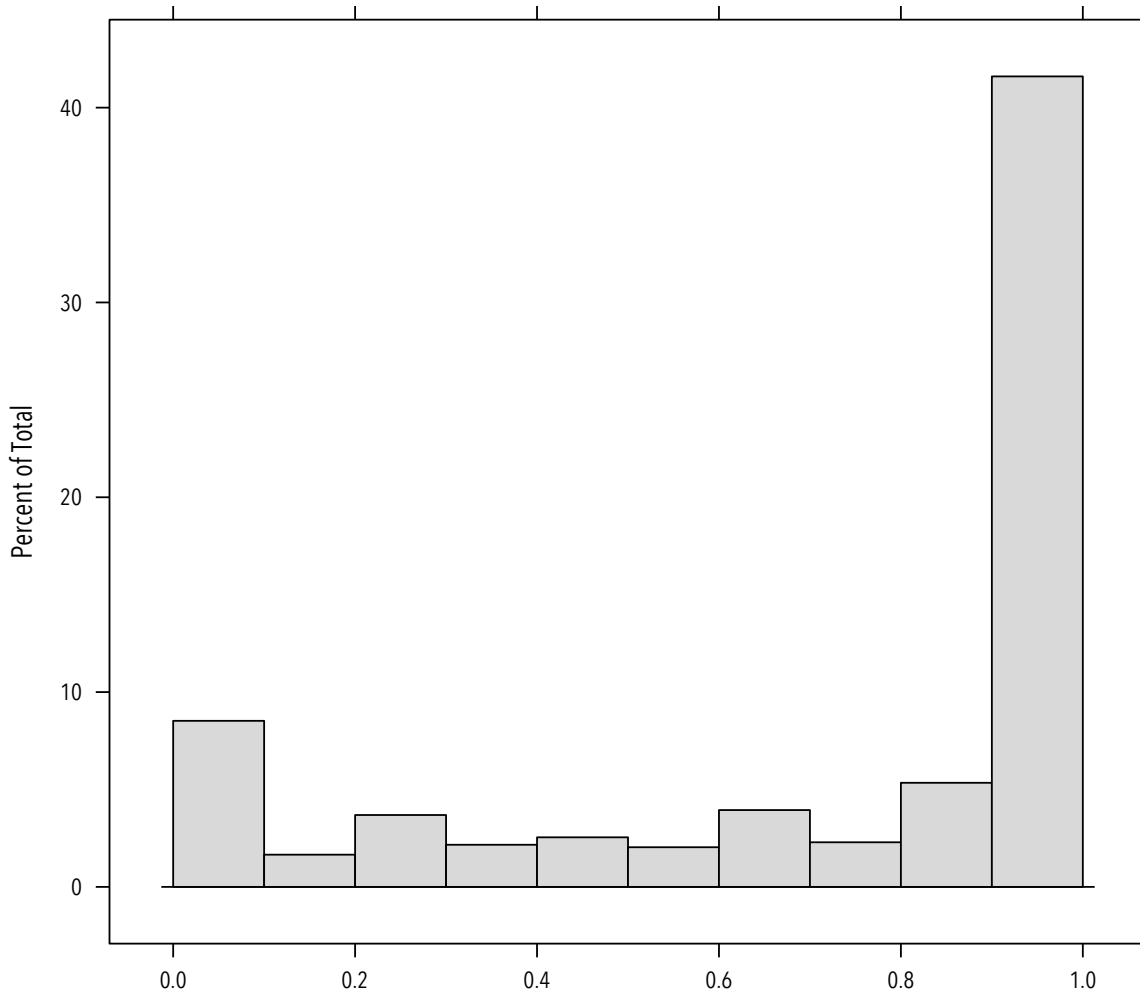


Figure 26: Stability in 580 successive pairs of efficiency gap measures

State	Plan	Start	End	EG avg	EG min	EG max
Florida	4	2002	2010	-0.112	-0.136	-0.084
New York	4	2002	2010	-0.111	-0.150	-0.078
Illinois	3	1992	2000	-0.103	-0.136	-0.058
Michigan	4	2002	2010	-0.103	-0.130	-0.077
New York	3	1992	2000	-0.098	-0.139	-0.048
New York	1	1972	1980	-0.097	-0.108	-0.079
Missouri	4	2002	2010	-0.091	-0.142	-0.061
Ohio	4	2002	2010	-0.090	-0.143	-0.049
New York	2	1982	1990	-0.084	-0.120	-0.028
Ohio	3	1994	2000	-0.083	-0.109	-0.025
Michigan	3	1992	2000	-0.080	-0.128	-0.019
Wisconsin	4	2002	2010	-0.076	-0.118	-0.039
Colorado	2	1982	1990	-0.075	-0.117	-0.055
Colorado	1	1972	1980	-0.041	-0.067	-0.018
California	3	1992	2000	-0.041	-0.057	-0.018
Pennsylvania	2	1982	1990	-0.033	-0.056	-0.020
Florida	1	1972	1980	0.070	0.052	0.099

Table 1: Plans with no doubt as to the sign of the efficiency gap over the life of the plan (3+ elections).

Interestingly, these plans with an utterly unambiguous history of one-sided *EG* measures are almost all plans with efficiency gaps that are disadvantageous to Democrats. Michigan’s 2002-2010 plan is on this list, as is the plan in place in Wisconsin 2002-2010 (average *EG* of -.076).

We examine this probability of “3+ consecutive *EG* measures with the same sign” for all of the plans with 3 or more elections in this analysis. 35% of 141 plans with 3 or more elections have at least a 95% probability of recording plans with *EG* measures with the same sign. If we relax this threshold to 75%, then 46% of plans with 3 or more elections exhibit *EG* measures with the same sign. Again, there is a reasonable amount of within-plan movement in *EG*, but in a large proportion of plans the efficiency gap appears to be a stable attribute of the plan.

10 A threshold for the efficiency gap

We now turn to the question of what might determine a threshold for determining if the EG is a *large and enduring* characteristic of a plan. We pose the problem as follows:

for a given threshold $EG^* > 0$, what is the probability that having observed a value of $EG \geq EG^*$ we then see $EG < 0$ in the remainder of the plan?

To answer this we compute

- if (and optionally, when) a plan has $EG \geq EG^*$;
- conditional on seeing $EG \geq EG^*$, do we also observe $EG < 0$ (a sign flip) in the same districting plan?

For $EG < 0$, the computations are reversed: conditional on seeing $EG < EG^*$, do we also see $EG > 0$ under the same plan?

Figure 27 displays two proportions, plotted against a series of potential thresholds on the horizontal axis. The two plotted proportions are:

- the proportion of plans in which we observe an EG more extreme than the specified threshold EG^* (on the horizontal axis);
- among the plans that trip the specified threshold, the proportion in which we see a EG in the same plan with a different sign to EG^* .

Plans with at least one election with $|EG| > .07$ are reasonably common: over the entire set of plans analyzed here — and again, with the uncertainty in EG estimates taken into account — there is about a 20% chance that a plan will have at least one election with $|EG| < .07$.

Observing $EG > .07$ is not a particularly informative signal with respect to the other elections in the plan. Conditional on observing an election with $EG > .07$ (an efficiency gap favoring Democrats), there is an a 45% chance that *under the same plan* we will observe $EG < 0$. That is, making an inference about a plan on the basis of one election with $EG > .07$ would be quite risky. Estimates

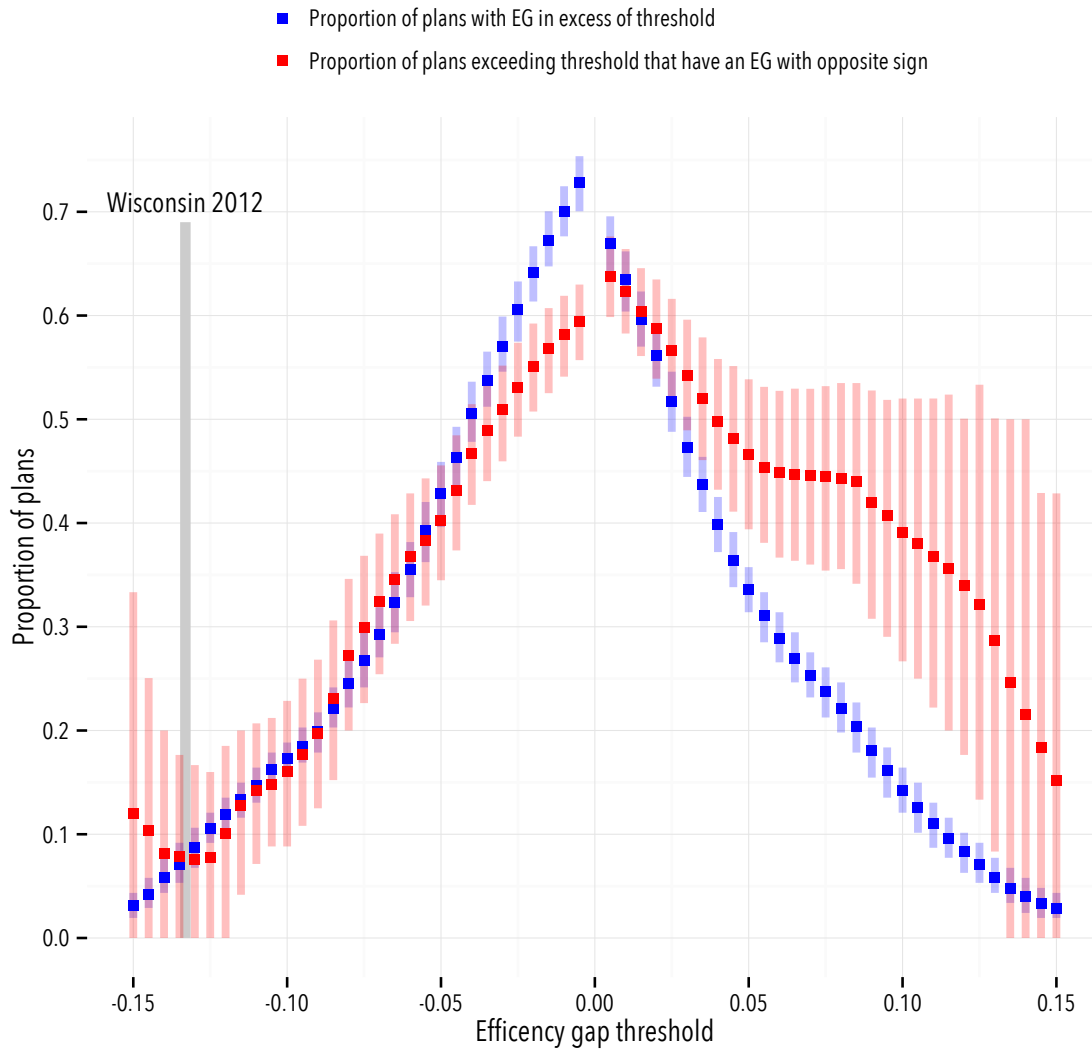


Figure 27: Proportion of plans that (a) record an efficiency gap measure at least as extreme as the value on the horizontal axis; and (b) conditional on at least one election with *EG* in excess of this threshold (not necessarily the first election), the proportion of plans where there is another election in the plan with an *EG* of the opposite sign.

of the “sign flip” rate conditional on a plan generating a relatively large, pro-Democratic EG estimates are quite unreliable because there are so few plans generating large, pro-Democratic EG estimates to begin with; note the confidence intervals on the “sign flip” rate get very wide as the data become more scarce on the right hand side of the graph.

This finding is not symmetric. The “signal” $EG < -.07$ (an efficiency gap disadvantageous to Democrats) is much more informative about other elections in the plan than the opposite signal $EG > .10$ (a pro-Democratic efficiency gap). If any single election in the plan has $EG < -.07$ then the probability that *all* elections in the plan have $EG < 0$ is about .80. That is, there is a smaller degree of within-plan volatility in plans that disadvantage Democrats. Observing a relatively low value of the EG such as $EG < -.07$ is much more presumptive of a systematic and enduring feature of a redistricting plan than the opposite signal $EG > .07$. Efficiency gap measures that appear to indicate a disadvantage for Democrats are thus more reliable signals about the respective districting plan than efficiency gap measures indicating an advantage for Democrats.

We repeat this previous exercise, but restricting attention to more recent elections and plans, with the results displayed in Figure 28. Again we see that plans with pro-Democratic EG measures are quite likely to also generate an election with $EG < 0$; and again, note that estimates of the “sign flip” rate are quite unreliable because there are so few plans generating large, pro-Democratic EG estimates to begin with.

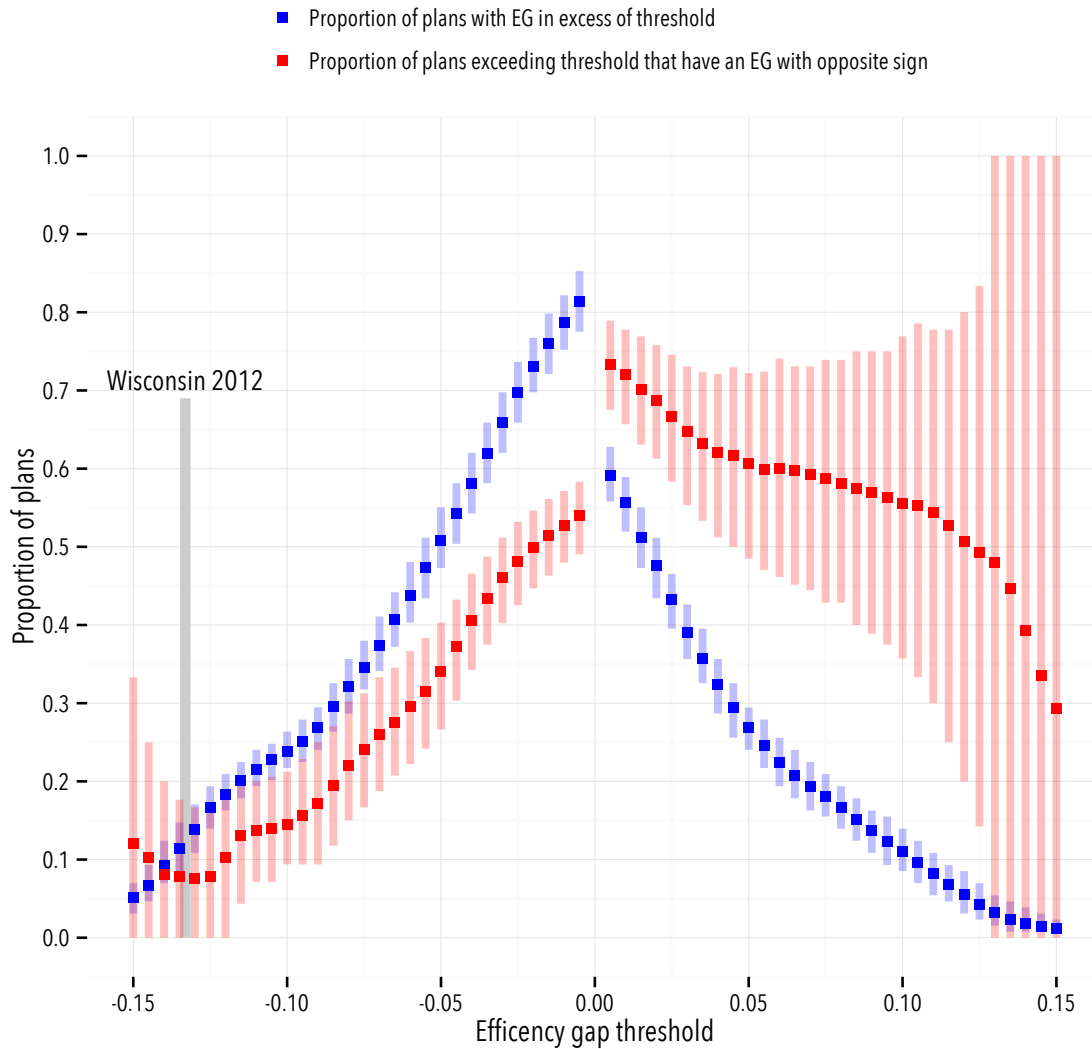


Figure 28: Proportion of plans in which (a) the efficiency gap measure is at least as extreme as the value on the horizontal axis; and (b) of these plans with at least one election with *EG* in excess of this threshold (not necessarily the first election), the proportion of plans in which there is another election in the plan with an *EG* of the *opposite* sign. Analysis of state legislative elections in 129 plans, 1991-present.

10.1 Conditioning on the first election in a districting plan

We also compute this probability of a sign flip in *EG* conditional on the magnitude of the *EG* observed with the *first* election under a districting plan. We perform this analysis twice: (1) for all elections in the data set and (2) for elections held under plans adopted in 1991 or later.

Figures 29 and 30 display the results of these analyses. First, over the full set of data (Figure 29) we observe a roughly symmetric set of *EG* scores in the first election under a plan. But we seldom see plans in the 1990s or later that commence with a large, pro-Democratic efficiency gap; the probability of a first election having $EG > .10$ is zero and the probability of a first election having $EG > .05$ (historically, not a large *EG*) is only about 11%. Negative efficiency gaps (not favoring Democrats) are much more likely under the first election in the post-1990 plans: almost 40% of plans open with $EG < -.05$ and about 20% of plans open with $EG < -.10$.

As noted earlier, pro-Democratic efficiency gaps seem much more fleeting than pro-Republican efficiency gaps. Conditional on a pro-Republican estimate of $EG > 0$ in the first election under a plan, the probability of seeing *EG* change sign over the life of the plan is almost always around 40% (1972-2014, Figure 29) or 50% (1991-present, Figure 30).

A very different conclusion holds if the first election observed under a plan indicates a sizeable efficiency gap working to disadvantage Democrats. In fact, the more negative the initial *EG* observed under a plan, the more confident we can be that we will continue to observe $EG < 0$ over the sequence of elections to follow under the plan. Conditional on a first election with $EG < -.10$, the probability of *all subsequent* efficiency gaps being negative is about 85%. Indeed, it is more likely than not that if the first election has $EG < 0$ (no matter how small), then so too will all subsequent elections (a 60% chance of this event).

Note that the Current Wisconsin Plan opens with $EG = -.13$ in the 2012 election. Analysis of efficiency gap measures in the post-1990 era (Figure 30) indicates that conditional on an *EG* measure of this size and sign, there is a 100% probability that *all subsequent elections* held under that plan will also have efficiency gaps disadvantageous to Democrats. That is, in the post-1990 era, if a plan's first election yields $EG \leq -.13$, we *never* see a subsequent election under that plan yielding a pro-Democratic efficiency gap. In short, a signal such as

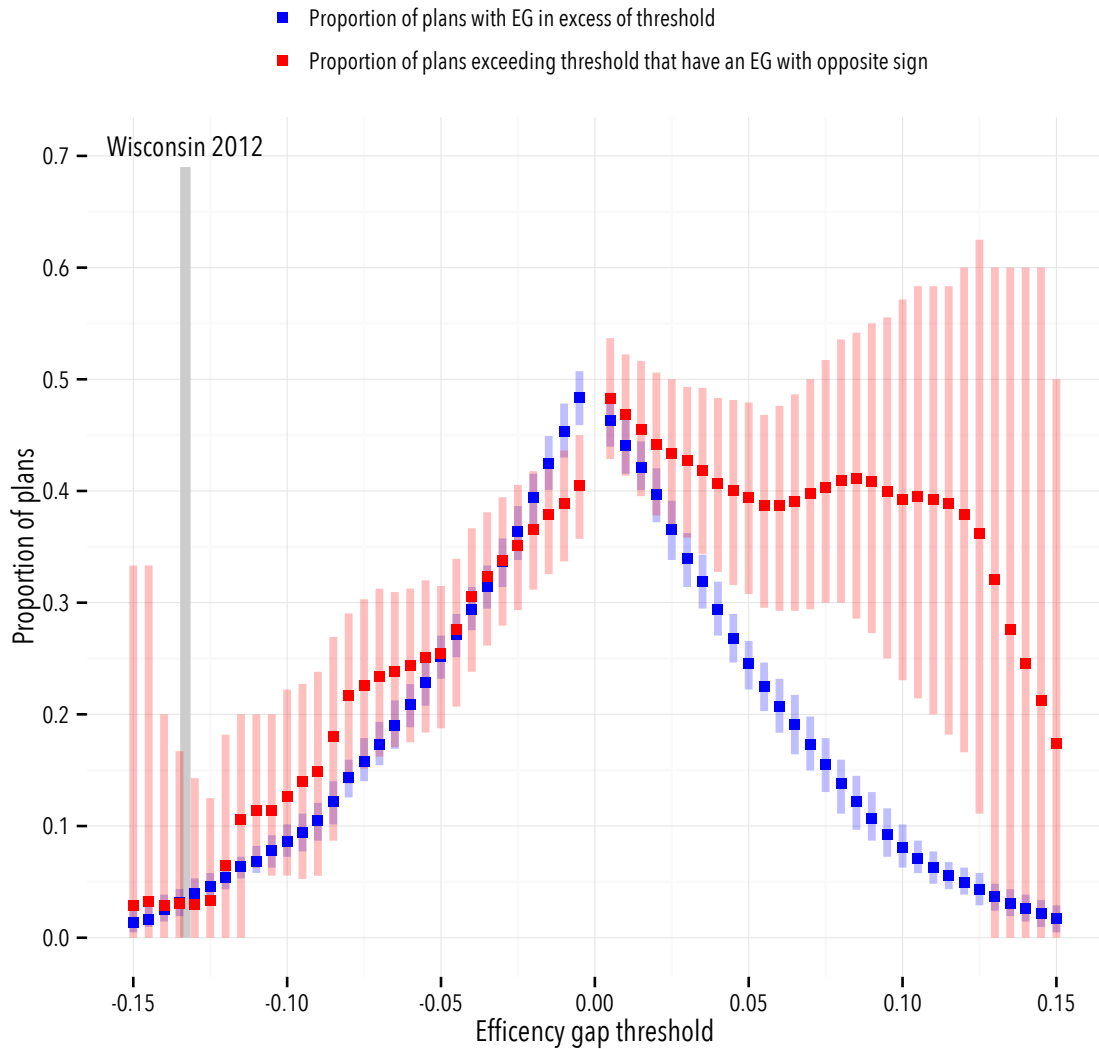


Figure 29: Proportion of plans in which the *first election* (a) has an efficiency gap measure at least as extreme as the value on the horizontal axis; and (b) conditional on the first election having an *EG* in excess of this threshold, the proportion of those plans in which a *subsequent election* has an *EG* of the *opposite* sign. Analysis of all state legislative elections in all plans with more than one election, 1972-present.

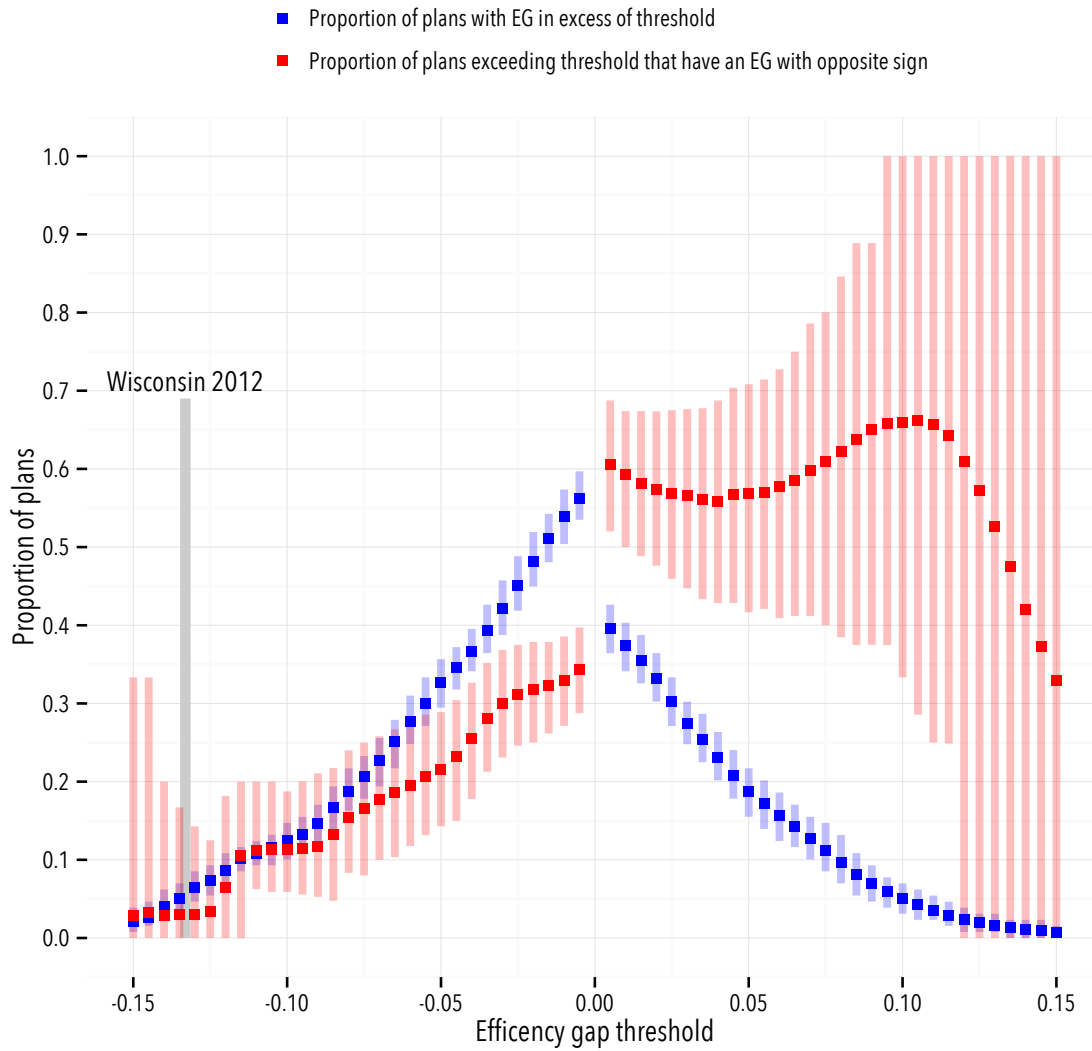


Figure 30: Proportion of plans in which the *first election* (a) has an efficiency gap measure at least as extreme as the value on the horizontal axis; (b) conditional on the first election having an *EG* in excess of this threshold, the proportion of those plans in which a *subsequent election* has an *EG* of the *opposite* sign. Analysis of state legislative elections in 129 plans, 1991-present.

$EG \leq -.13$ is extremely reliable with respect to the districting plan that generated it, at least given the post-1990 record.

10.2 Conditioning on the first two elections in a districting plan

The difficulty with conditioning on the first two elections of a districting plan is that the data start to thin out. In the entire data set there simply aren't many districting plans that equal or surpass the two, relatively large values of EG observed in Wisconsin in the first two elections of the current plan. Indeed, the only cases with a similar history of EG measures like Wisconsin's in 2012 and 2014 are contemporaneous cases: Florida, Michigan, and North Carolina in 2012 and 2014.

We relax the threshold of what counts as a similar case to encompass plans whose first two efficiency gap measures are within 75% of the magnitude of Wisconsin's 2012 and 2014 EG measures; we now pick up 11 roughly comparable cases, 4 of which date from earlier decades. Again, this is testament to how recent decades have seen an increase in the prevalence of larger, negative values of the efficiency gap.

For the four prior cases we plot the sequence of EG estimates in Figure 31. With the exception of the last election in the highly unusual Delaware sequence (among the most volatile observed in the data set; see section 9.3), the other proximate cases all go on to record efficiency gap measures that are below zero over the balance of the plan. We stress that four cases doesn't provide much basis for comparison, but this only speaks to the fact that the sequence of two large, negative values of the efficiency gap in Wisconsin in 2012 and 2014 are virtually without historical precedent. We have little guidance from the historical record as to what to expect given an opening sequence of EG measures like the ones observed in Wisconsin. But the little evidence we do have suggests that a stream of similarly sized, negative values of the efficiency gap are quite likely over the balance of the districting plan.

10.3 An actionable EG threshold?

We now consider a more general question: what is an actionable threshold for the efficiency gap?

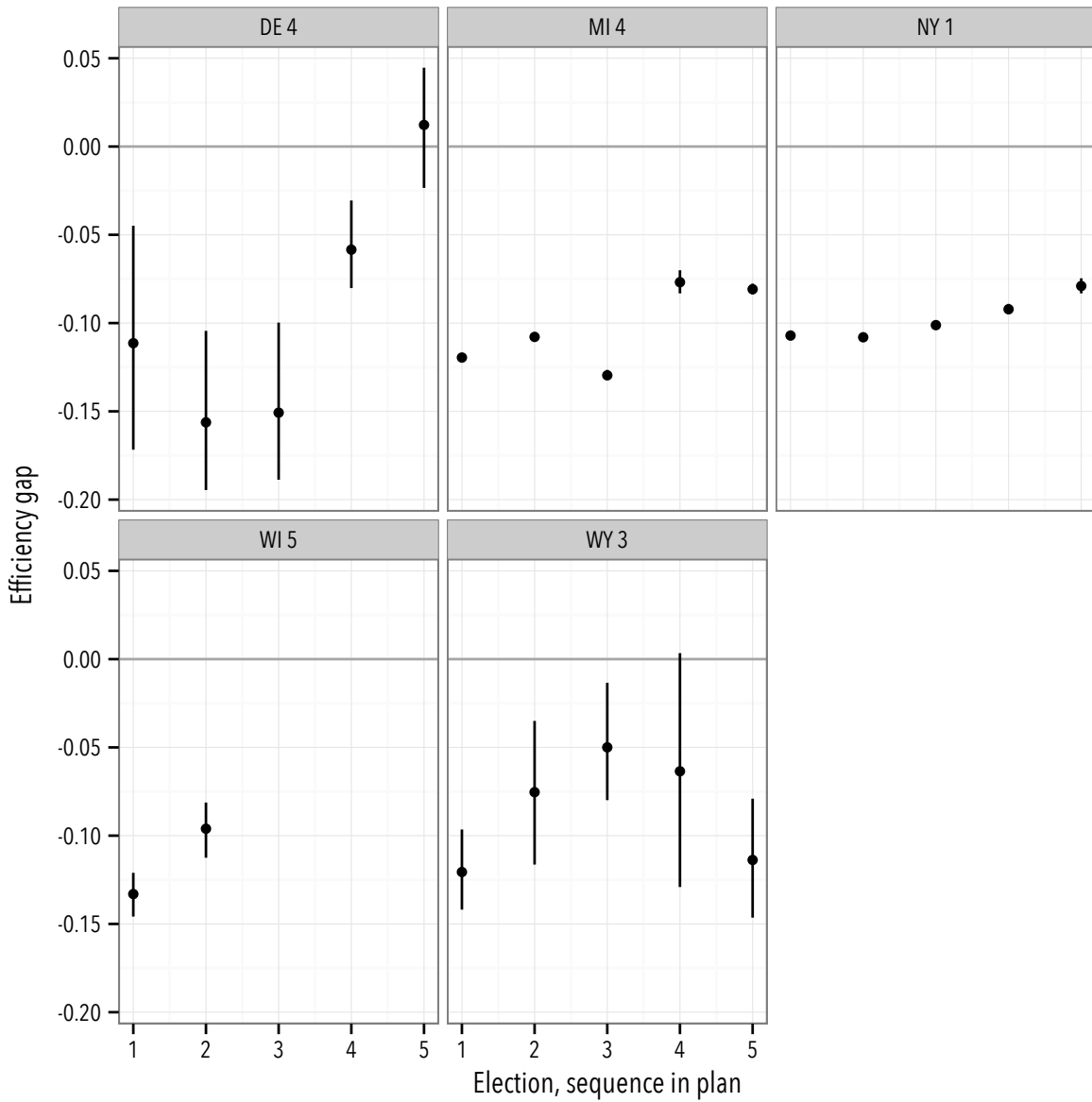


Figure 31: Sequence of *EG* estimates observed over the life of districting plans, for pre-2010 plans with first two *EG* scores within 75% of the magnitude of the *EG* scores observed in Wisconsin in 2012 and 2014.

First, recall that relatively small *EG* estimates are likely to be swamped by their estimation uncertainty, depending on the proportion of uncontested districts in the given election and the statistical procedures. In every instance though, this is an empirical question; at least in the approach I present here, each *EG* estimate I generate is accompanied with uncertainty bounds, letting us assess the *probability* that a given estimate is positive or negative. Figure 19 provides a summary of the relationship between the size of the *EG* estimate and the “statistical significance” of the estimate (in the sense that the 95% credible interval for each estimate does not overlap zero).

Second, the distribution of *EG* statistics in the 1972-2014 period is roughly symmetric around zero. Reference to this empirical distribution might also be helpful in setting actionable thresholds, and answering the question “is the *EG* measure at issues large relative to those observed in the previous 40 years of state legislative elections?” Double digit *EG* measures (-.10 or below; .10 or above) are pushing out into the extremes of the observed distribution of *EG* estimates: *EG* estimates of this magnitude are comfortably past the question of “statistical significance.” Just 15% of the 786 *EG* measures generated in this analysis are below -.07; fewer than 12% are greater than .07.

We do need to be careful when making these kinds of *relative* assessments about the magnitude of the efficiency gap. If pro-Republican gerrymandering is widespread, then it will be less unusual to see a large, negative *EG* estimate, at least contemporaneously; in fact this appears to be the case in the post-2010 set of elections, where the longer-term distinctiveness of the Wisconsin numbers is matched and in some cases exceeded by other states also recording unusually large, negative *EG* estimates (e.g., Florida, Michigan, Virginia and North Carolina). This speaks to the utility of the longer-term, historical analysis in both [Stephanopolous and McGhee \(2015\)](#) and in this report. It is important to remember that $EG = 0$ corresponds to a partisan symmetry in wasted vote rates; we should be wary of arguments that would lead us to tolerate small to moderate levels of the efficiency gap because they appear to be the norm in some period of time, or in some set of jurisdictions.

In any litigation, much will turn on the question of *durability* in the efficiency gap, and this concern motivates much of the preceding analysis. We cannot wait until three, four, or more elections have transpired under a plan in order to

assess its properties. Courts will be asked to assess a plan based on only one *EG* estimate, or two. Analysis of the sort I provide here will be informative in these cases, assessing whether the estimate is so large that the historical record suggests that the first election's *EG* estimate is a reliable indicator as an enduring feature of the plan, and not an election-specific aberration.

10.4 Confidence in a given threshold

Figures 32 and 33 present my estimate of a “confidence rate” associated with a range of possible “actionable thresholds” for the efficiency gap. These figures essentially re-package the information shown in Figures 29 and 30. Suppose a court rejects or amends every plan with a first election *EG* more extreme (further away from zero) than the proposed threshold shown on the horizontal axis of these graphs. A certain number of plans fail to trip this threshold, and so are upheld by the courts if they are challenged. Of those that do trip the threshold and are rejected by a court, what is our confidence that the plan, if left undisturbed, would go on to produce a sequence of *EG* measures that lie on the same side of zero as the threshold? Combining these two proportions gives us an overall confidence measure associated with a particular threshold.

This analysis points to a benchmark of about $-.06$ or $-.07$ as the actionable threshold given a first election with $EG < 0$ (Democratic disadvantage) or $.08$ or $.09$ when we observe $EG > 0$ in the first election under a redistricting plan (Democratic advantage); the asymmetry here reflects the fact that districting plans evincing apparent Democratic advantages are not as durable or as common (in recent decades) as plans presenting evidence of pro-Republican gerrymanders. At these proposed benchmarks the overall confidence rates are estimated to be 95%, with this confidence rate corresponding to a benchmark used widely in statistical decision-making in many fields of science.

Figures 32 and 33 also highlight that $EG < -.07$ or $EG > .07$ would be an extremely conservative threshold. On the pro-Democratic side, $EG > .07$ is a rare event. Districting plans unfavorable to Democrats, with $EG < -.07$ are not unusual; about 10% of post-1990 plans generate *EG* measures below $-.07$; the proportion of these plans that then record a sign flip is only about 10%; see Figure 30. If the presumption was that any plan with a first election showing

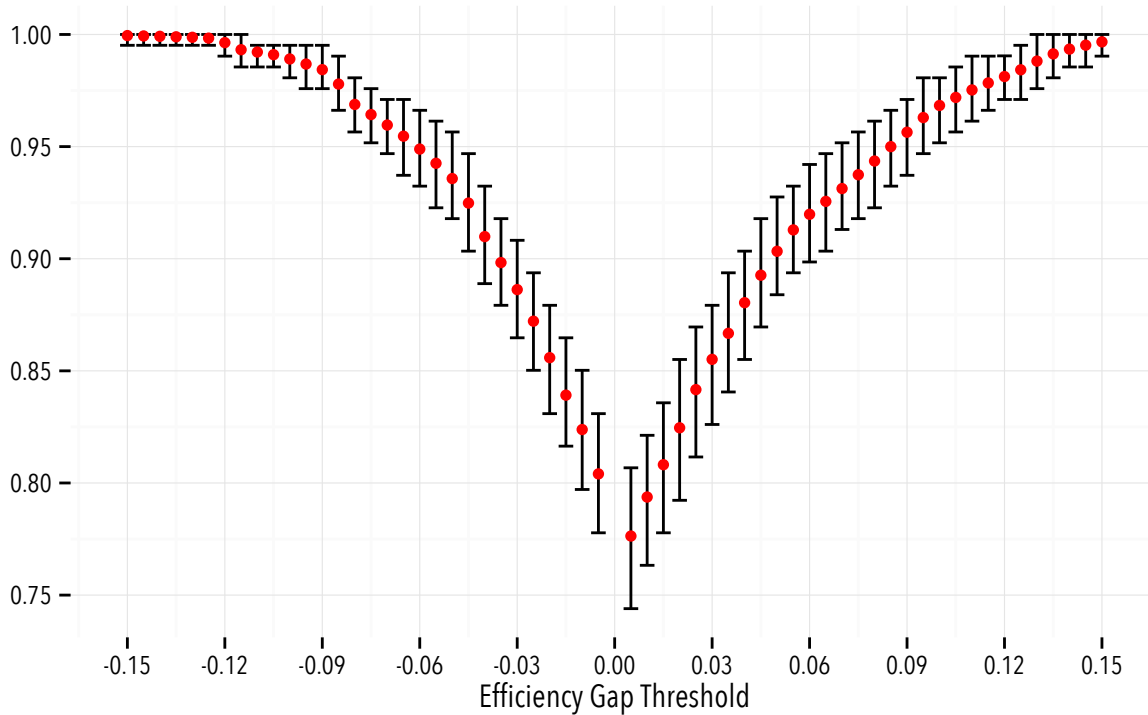


Figure 32: Proportion of plans being either (a) undisturbed or (b) if left undisturbed, would continue to produce one-sided partisan advantage (no sign change in subsequent *EG* measures), as a function of the proposed “first election,” efficiency gap threshold (horizontal axis), based on analysis of all multi-election districting plans, 1972-2014. The proportion on the vertical axis is thus interpretable as the “confidence level” associated with intervention at a given first election, *EG* threshold. Vertical lines indicate 95% credible intervals.

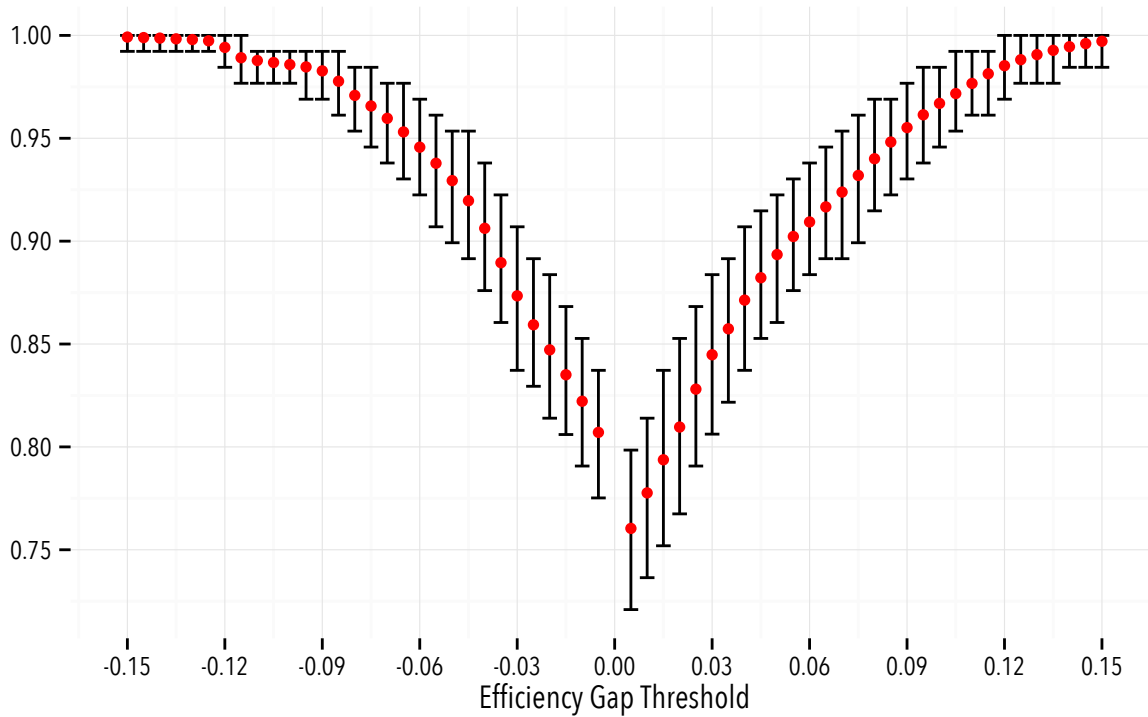


Figure 33: Proportion of plans being either (a) undisturbed or (b) if left undisturbed, would continue to produce one-sided partisan advantage (no sign change in subsequent *EG* measures), as a function of the efficiency gap threshold (horizontal axis), based on analysis of post-1990 plans and elections. The proportion on the vertical axis is thus interpretable as the “confidence level” associated with intervention at a given first election, *EG* threshold. Vertical lines indicate 95% credible intervals.

$EG < -.07$ would be rejected, then we'd be “wrong” to do so in about 10% of those cases (in the sense that if left in place, the plan would go on to produce at least one election with $EG > 0$). The total error rate in this case would be 1% of all plans. Equivalently, 99% of all plans would be either left undisturbed or appropriately struck down or amended by a court, given the historical relationship between “first election” EG measures and the sequence of EG measures that follow.

11 Conclusion: the Wisconsin plan

Wisconsin has had two elections for its legislature under the plan currently in place, in 2012 and 2014. Both elections were subject to considerable rates of uncontestedness (27 of 99 seats in 2012 and 52 of 99 seats in 2014), but these rates are hardly unusual; Wisconsin's rates of uncontested districts in these two elections are low to moderate compared to other states. We use the relationship between state legislative election results and presidential election results in state legislative districts (and incumbency) to impute two-party vote shares in uncontested seats (see section 7.2). With a complete set of vote shares, we then compute average district-level Democratic two-party vote share (V) and note the share of seats (contested and uncontested) won by Democratic candidates (S).

In Wisconsin in 2012, and after imputations for uncontested seats, V is estimated to be 51.4% (± 0.6); recall that Obama won 53.5% of the two-party presidential vote in Wisconsin in 2012. Yet Democrats won only 39 seats in the 99 seat legislature ($S = 39.4\%$), making Wisconsin one of 7 states in 2012 where we estimate $V > 50\%$ but $S < 50\%$ and where Democrats failed to win a majority of legislative seats despite $V > 50$ (the other states are Florida, Iowa, Michigan, North Carolina and Pennsylvania). In 2014, V is estimated to be 48.0% (± 0.8) and Democrats won 36 of 99 seats ($S = 36.4\%$).

This provides the raw ingredients for computing the efficiency gap (EG) for these two elections (recalling equation 1). Repeating these calculations across a large set of state elections provides a basis for assessing whether the efficiency gap estimates for Wisconsin in 2012 and 2014 are noteworthy.

Wisconsin's efficiency gap measures in 2012 and 2014 are $-.13$ and $-.10$ (to two digits of precision). These negative estimates indicate the disparity between

Highlighting Wisconsin plan 5

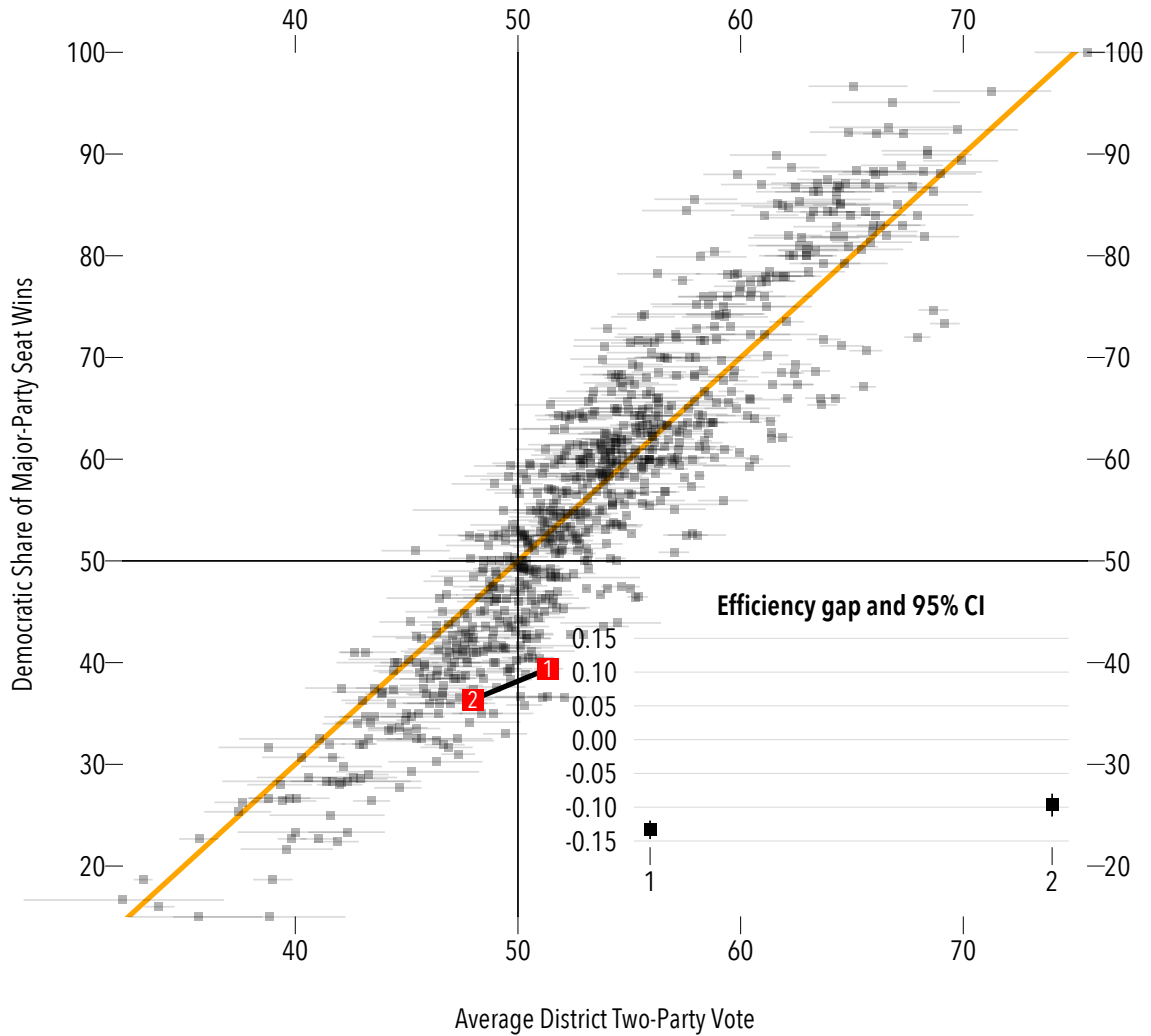


Figure 34: Seats, votes and the efficiency gap recorded under the Wisconsin plan, 2012 and 2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts. The inset in the lower right shows the sequence of efficiency gap measures recorded under the plan; vertical lines are 95% credible intervals.

vote shares and seat shares in these elections, which in turn, is consistent with partisan gerrymandering. The negative *EG* estimates generated in 2012 and 2014 are unusual relative to Wisconsin's political history (see Figure 35). The 2012 estimate is the largest *EG* estimate in Wisconsin over the 42 year period spanned by this analysis (1972-2014); the 2014 estimate is the fourth largest (behind 2012, 2006 and 2004, although it is essentially indistinguishable from the 2004 estimate). The jump from the *EG* values being recorded towards the end of the previous districting plan in Wisconsin (2002-2010) to the 2012 and 2014 values strongly suggests that the districting plan adopted in 2011 is a driver of the change, systematically degrading the efficiency with which Democratic votes translate into Democratic seats in the Wisconsin state legislature.

Wisconsin's 2012 and 2014 *EG* estimates are also large relative to the *EG* scores being generated contemporaneously in other state legislative elections. Figure 36 shows *EG* estimates recorded under plans in place since the post-2010 census round of redistricting; the *EG* estimates are grouped by state and ordered, with Wisconsin highlighted. We have 78 *EG* scores from elections held since the last round of redistricting. Among these 79 scores, Wisconsin's *EG* scores rank eighth (2012, 95% CI 3 to 12) and seventeenth (2014, 95% CI 13 to 20).

The historical analysis reported above supports the proposition that Wisconsin's *EG* scores are likely to endure over the course of the plan. Few states ever record *EG* scores as large as those observed in Wisconsin; indeed, there is virtually no precedent for the lop-sided, two election sequence of *EG* scores generated in Wisconsin in 2012 and 2014 in the data I analyze here (1972-2014). The closest historical analogs suggest that a districting plan that generates an opening, two-election sequence of *EG* scores like those from Wisconsin will continue to do so, generating seat shares for Democrats that are well below those we would expect from a neutral plan.

The Current Wisconsin Plan is generating estimates of the efficiency gap far in excess of the proposed, actionable threshold (see section 10). In 2012 elections to the Wisconsin state legislature, the efficiency gap is estimated to be -.13; in 2014, the efficiency gap is estimated to be -.10. Both measures are separately well beyond the conservative .07 threshold suggested by the analysis of efficiency gap measures observed from 1972 to the present.

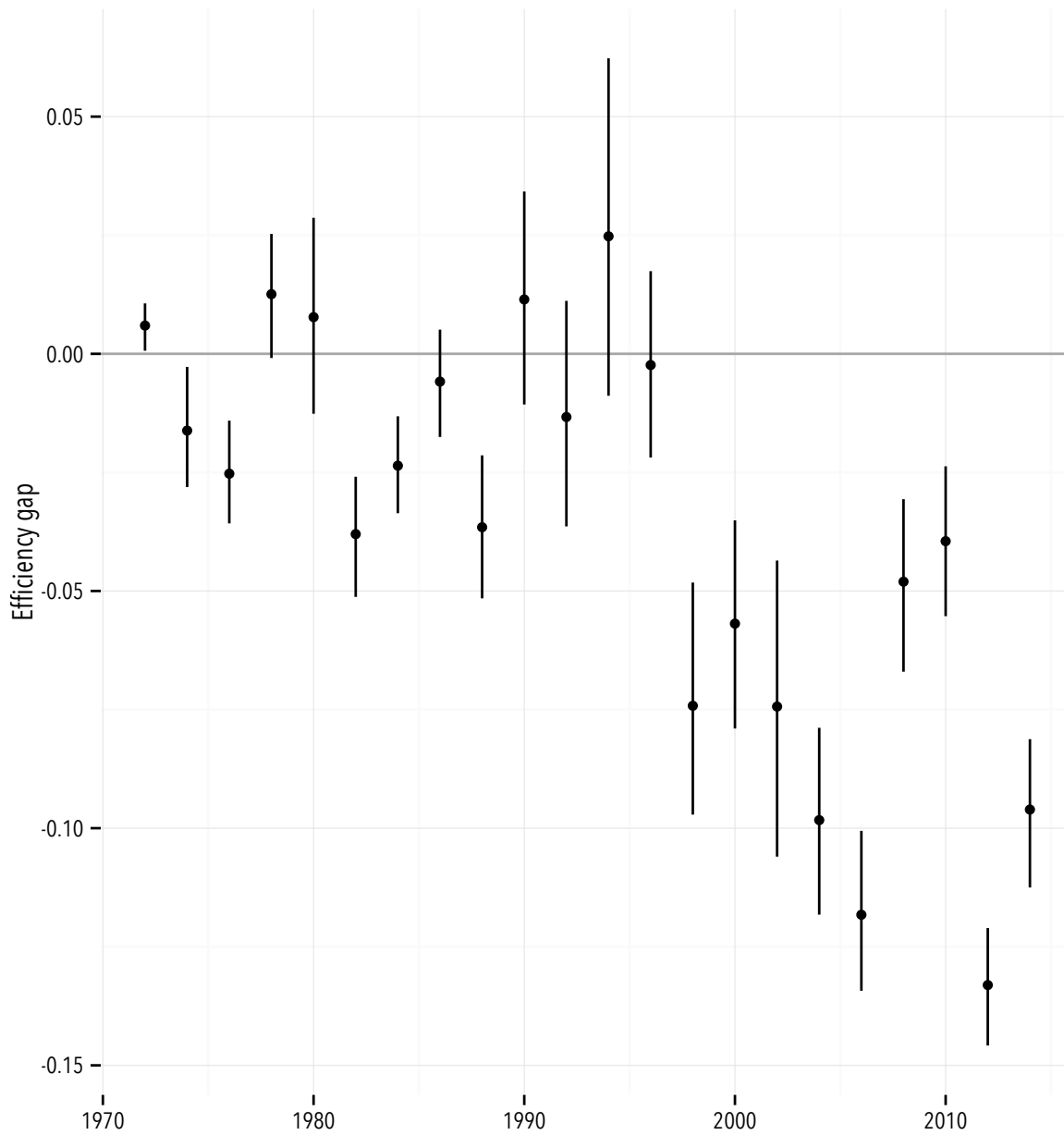


Figure 35: History of efficiency gap estimates in Wisconsin, 1972-2014. Vertical lines indicate 95% credible intervals.

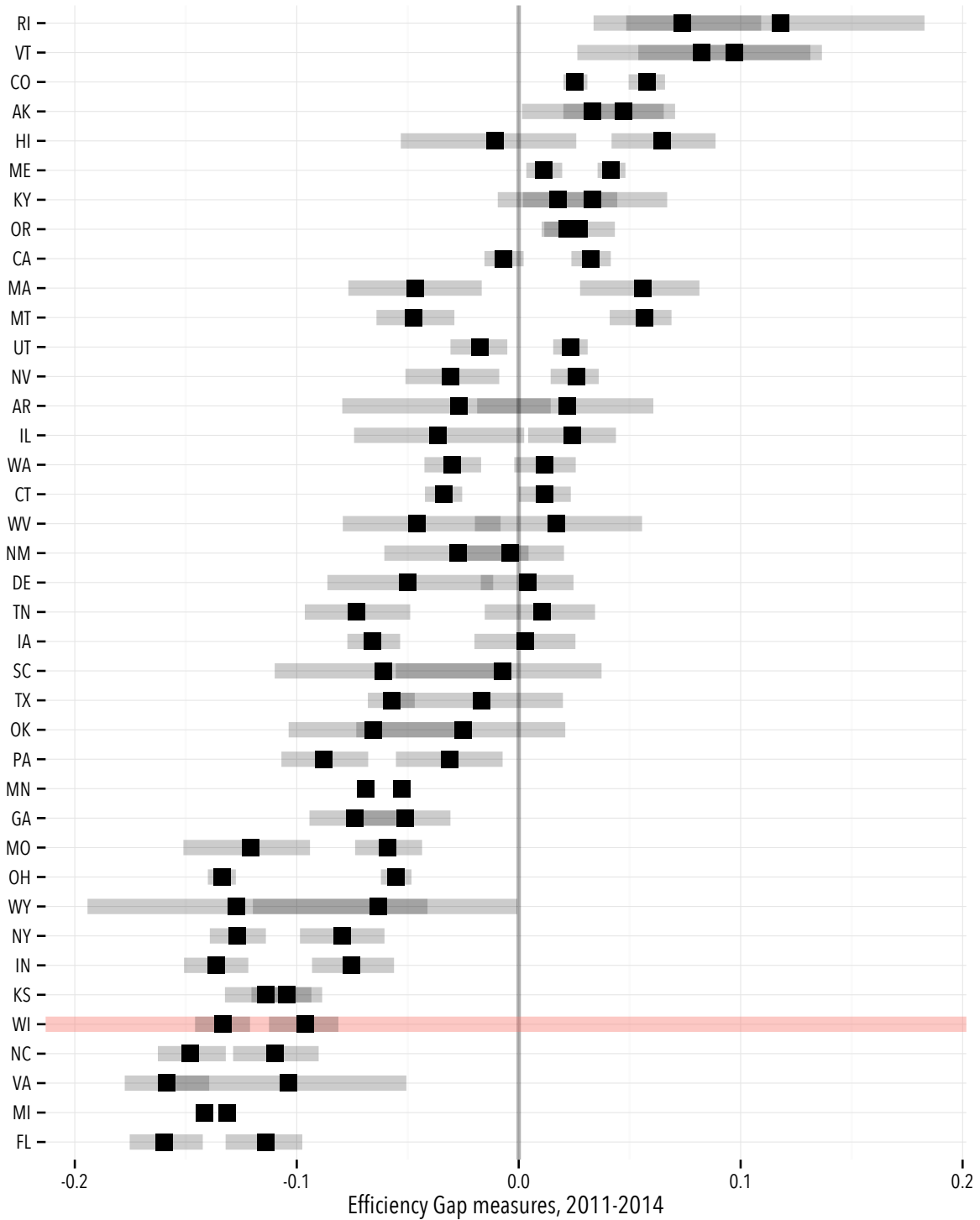


Figure 36: *EG* estimates in 2012 and 2014, grouped by state and ordered. Horizontal bars indicate 95% credible intervals.

References

- Butler, David and Donald Stokes. 1974. *Political change in Britain: the evolution of electoral choice*. London: Macmillian.
- Edgeworth, FY. 1898. "Applications of the Theory of Probability." *JR Statist. Soc* 51:534.
- Gelman, Andrew and Gary King. 1990. "Estimating the Consequences of Electoral Redistricting." *Journal of the American Statistical Association* 85:274–82.
- Kendall, M. G. and A. Stuart. 1950. "The Law of Cubic Proportion in Election Results." *The British Journal of Sociology* 1(3):183–196.
- McGhee, Eric. 2014. "Measuring Partisan Bias in Single-Member District Electoral Systems." *Legislative Studies Quarterly* 39(1):55–85.
- Niemi, Richard G. and Patrick Fett. 1986. "The Swing Ratio: An Explanation and an Assessment." *Legislative Studies Quarterly* 11:75–90.
- Niemi, Richard G. and Simon Jackman. 1991. "Bias and Responsiveness in State Legislative Districting." *Legislative Studies Quarterly* 16:183–202.
- Stephanopolous, Nicholas O. and Eric M. McGhee. 2015. "Partisan Gerrymandering and the Efficiency Gap." *University of Chicago Law Review* 82. Forthcoming.
- Tufte, Edward R. 1973. "The Relationship Between Seats and Votes in Two-Party Systems." *American Political Science Review* 67:540–554.