

# An Empirical Bayesian Framework for Assessing Partisan Bias in Redistricting Plans

Kevin Baas and Colin McAuliffe

June 26, 2017

## Abstract

We present a framework for assessing partisan bias in redistricting plans using an empirical Bayesian approach, as well as a new metric for measuring partisan bias called the specific asymmetry. Given at least one observation of the election results in a state, we estimate a probability model representing likely future election outcomes. The expected value of the specific asymmetry is then computed by sampling from the model using the Monte Carlo method. Redistricting plans which have an expected asymmetry of zero may be considered fair, since on average neither of the two major parties will be given an advantage. Conversely, redistricting plans which have a nonzero expected asymmetry may be considered to be unfair, since partisan bias is a persistent feature of the plan. This analysis technique is applied to the United States congressional elections from 1972-2016 to examine the total and net effects of partisan bias in recent history. We also examine the post 2010 census Wisconsin State Assembly, which is the subject of *Whitford v. Gil*.

## 1 Introduction

In the United States, the process of drawing districts for a state's legislative bodies and federal congressional delegation is usually in the hands of that state's legislature. Partisan state legislatures have used the redistricting process in pursuit of various goals such as incumbent protection, maximizing their share of seats, and even diluting the voting power of citizens on the basis of race. Opportunities to manipulate the redistricting process for partisan advantage abound, particularly with the advent of sophisticated algorithms for redistricting as well as a high degree of partisan and geographic polarization in the current political climate. However, the supreme court has yet to rule definitively on the justiciability of partisan gerrymandering, and recent rulings suggest that a manageable standard for measuring gerrymandering is a prerequisite for a ruling on justiciability CITE.

A ruling on the justiciability of partisan gerrymandering would have significant consequences for democracy in the United States. Partisan gerrymandering not only distorts the composition of legislative bodies, it represents a dilution of the voter's ability to attempt to influence policy by electing representatives of their choosing. The ability of state legislatures to use redistricting to influence the outcome of elections comes at the direct expense of voter's ability to use their ballot to influence the outcome of elections. However, in order to properly assess the harm caused by partisan gerrymandering, courts require mathematical tools for analyzing partisan bias in a redistricting plan.

Development of a standard that fits with existing rulings, accurately measures partisan bias under generic conditions, and is simple enough to be effectively communicated to a court is not an easy task. For example, a common sense standard might be to require that a party's share of seats is proportional to its share of votes, which is simple and applicable to swing states as well as more partisan states. However, even when redistricting is fair, single winner voting systems tend not to produce such proportional outcomes [1], and the supreme court has stated that proportionality is not acceptable CITE. **Do you recall which ruling this was?** A second example is the mean median difference test [2, 3, 4], which fits with existing rulings since it measures partisan bias without regard to proportionality and is simple enough to be calculated by a judge without an expert witness. However, it is only effective in measuring bias for states which are close to even in terms of partisanship. [3]

Several standards have been proposed in literature, and for clarity we describe a standard as consisting of two parts. The first part is a metric, which is some quantity that can be

calculated from a given election result and is a measure of the unfairness or bias. The second part is a model (or in the absence of a statistical model per se an analysis of the metric in historical elections), which allows one to assess whether or not the observed value of the metric can be regarded as the result of chance in an otherwise unbiased redistricting plan, or if bias is a persistent feature of the plan. For example, the plaintiffs in *Whitford v. Gil* used a metric called the efficiency gap [5], which measures the discrepancy in the so called wasted votes between the two major parties. An examination of historical election data determined that it is reasonable to assume that observing an efficiency gap of greater than 7% in favor of the redistricting party would typically be predictive of efficiency gaps favoring that party for the entire redistricting cycle. Browning, Grofman and King proposed a metric based on the deficit in seats at 50% votes in a seats votes curve [6, 7], which was used in a Bayesian model for evaluating redistricting plans in [8, 9]. Nagle has proposed several other metrics based on the seats votes curve [10], including the geometric area between the seats votes curve and its reflection, although no accompanying model was proposed.

A special group of standards are those where a particular metric is associated with an existing analytic statistical model. Among these are the lopsided wins test, the mean median difference test, and the chi square test proposed by Wang [2, 3]. These standards are attractive because they are based on statistics with properties that are well understood, that are used broadly for applications across many fields, and are relatively simple. However, these tests have some drawbacks. Namely, the models used for computing the significance level for those tests are valid asymptotically as the number of districts increases, meaning that the statistical power of these standards is diminished for states with just a few districts. Additionally, a null distribution must be specified for each of the test statistics in question. This is a relatively minor point, since the null distributions used in the t-test and chi squared test are well established and the effect of the null distribution on the mean median difference test is small in many cases. [11, 12] I would argue its a major point. And it's kind of subjective, anyways, whether its minor or major Lastly and most importantly, these models assume that samples are independent and identically distributed (IID), which is an assumption that is crucial for the tests to be analytic (and hence simple) but not necessarily realistic for elections. For instance, the IID assumption would imply that the probability of a deeply partisan Republican district being won by a Republican is the same as the probability of a deeply partisan Democratic district in the same state being won by a Republican. Violation of the IID assumption can have a noticeable effect on statistical tests, for an example in securities law see [13]. Use of this assumption may make this group of standards too restrictive or too permissive depending on the specific circumstances. Despite these objections, standards based on analytic models are potentially very powerful due to their relative simplicity and their ability to be solved exactly, and we are currently investigating the effect the IID assumption has on the robustness of such standards.

Another group of standards are those based on Monte Carlo simulation. A major advantage of Monte Carlo based standards are flexibility, since they do not require the metrics to be analytically integrable I know, might be too mathy for the audience , and thus can be used with a wide variety of metrics or even multiple metrics in combinations. Additionally, once implemented by a computer algorithm, it is quick and easy to look at the data in many different ways.

An example is test III proposed by Wang [2, 3], where the metric used was the proportion of seats won by a party, and the model was a Monte Carlo technique where 'fantasy' congressional delegations were sampled from national congressional voting results. This method can be used to establish whether or not a party's vote share wins them seats excessively relative to a national baseline. A second example of a Monte Carlo based standard is that of Chen et al [14, 15], where electoral maps satisfying basic redistricting criteria are generated randomly and compared against the actual map. The standard that we propose is also based on Monte Carlo simulation. The metric, which we refer to as the specific asymmetry (see section 2), is a measure of partisan bias that determines discrepancy in seats won by the losing party under a reversal in statewide partisanship. The model, described in section 3, uses a Monte Carlo technique based simulations of an empirically estimated probability model for district and statewide partisanship. The drawback of Monte Carlo based standards is complexity relative to analytic standards which do not require a computer to implement. Since Monte Carlo is a numerical integration method, as opposed to analytic, it takes a lot more computer processing power, and thus aren't as well suited for use in optimization algorithms, relative

to analytic methods. Nonetheless, Monte Carlo simulation techniques are highly conventional tools that are in widespread use across numerous disciplines [16], and are useful for examining bias in redistricting.

The empirical Bayesian framework that we propose is not only appropriate for assessing bias that has existed in past redistricting plans, but it is useful for understanding how bias could be introduced into future redistricting plans, even as strategies for gerrymandering change in response to the introduction of a hypothetical standard. In fact, the empirical Bayesian model that we propose is compatible with all of the metrics we have discussed thus far, and can be used as part of a standard itself, or can be used to test the robustness of standards that do not require Monte Carlo simulation. We therefore contend that regardless of whether or not the approach we present is an acceptable judicial standard itself, it remains a powerful tool for assessing partisan bias in redistricting plans that will prove useful for researchers and advocates of redistricting reform. For example, the Monte Carlo technique can be used to evaluate the properties of metrics which do not have well known analytic statistical properties such as the efficiency gap and the specific asymmetry metric proposed in section 2, or to evaluate the effect of the IID assumption for metrics which do have known analytic properties such as the mean median difference and the t-test. Additionally, the technique could be used to help develop an understanding of the conditions where different metrics may agree or disagree with one another. For fast reference, table 1 summarizes each of the standards discussed here.

Table 1: Summary of the metrics and models employed by various gerrymandering standards

Standard	Metric	Model
Grofman and King	Deficit in Seats at 50% of the vote	None proposed
Nagle	Various metrics of seats votes curve asymmetry	None proposed
Stephanopolous and McGhee	Efficiency gap	None proposed, historical analysis used in <i>Whitford v. Gil</i>
Gelman et al	Deficit in Seats at 50% of the vote, various others	Bayesian simulation Bayesian simulation
Mean Median Difference	Mean median difference	asymptotic distribution (unit normal) assuming IID samples
Chi square test	Party in state and national winning vote share variances	asymptotic distribution (chi squared) assuming IID samples
Lopsided wins test	Difference in party mean winning vote share	asymptotic distribution (student's t) assuming IID samples
Test III [3] or 'excess seats test'	Party seat share	MC sampling of national district voting results
Chen and Rodden	Party seat share	MC sampling of randomly generated districting plans
Present study	Specific Asymmetry	MC sampling of the empirical Bayesian model

## 2 Measuring Bias with the Specific Asymmetry

Measuring partisan bias accurately presents several challenges from a mathematical and legal perspective. Redistricting will always be an artificial process, and therefore we must establish what characteristics of a redistricting plan can be regarded as normative before attempting to measure deviations from the normative standard. We propose the following definition: *a fair redistricting plan is one that, on average, will not be biased in favor of either of the two major parties* With this definition in mind, we turn our attention to several challenges in devising a metric for bias. First, single winner elections tend not to produce proportional

outcomes CITE, and we therefore can not use proportionality or the lack thereof as a measure of bias. Bias should therefore be a measure of symmetry, which permits disproportionate outcomes but requires that any advantage that might arise from the disproportionate nature of the single winner election system be equally available to either party. In other words, we say that the result is unbiased if and only if any advantage gained by a party, above what is proportional, is a consequence of the disproportionate nature of the single winner election system, and not an additional advantage piled on top of that.

A few measures of bias that follow this definition exist in the literature, such as Grofman and King, and the mean median difference CITE. However, these measures of bias tend to apply only to states which are close to 50-50 partisanship CITE. For more partisan states, these measures become more difficult to interpret except as the bias which could have occurred had the partisanship in the state in question been closer to even. Another approach is to consider the bias at levels of partisanship by using the geometric area between the seats votes curve and its reflection CITE. Useful information may be gleaned from these approaches, but Justice Kennedy expressed a dislike for measures based on a hypothetical state of affairs. Additionally, measurement of bias at different levels of hypothetical partisanship may lead to significantly different results depending on the partisanship level considered. For example, the seats votes curve for the Wisconsin state assembly shows substantial bias favoring Republicans at close to even partisanship, but the complete reverse at high levels of Democratic partisanship. This means that bias measures which do not consider the actual partisanship in a state could result in false positives and false negatives. In light of this, a measure of bias that is applicable to any level of partisanship and that directly considers the actual level of partisanship in a given state seems appropriate. To address these challenges, we propose a measure of bias which we call the specific asymmetry, which is the deviation of the seats votes curve from the symmetry line measured at the statewide popular vote.

Bias by any measure may occur by chance in an otherwise fair redistricting plan, it is even possible for an antimajoritarian outcome to result from chance. It is also important not only to consider the harm a plan has caused to voters, but to also seek to assess the likelihood that a plan will cause harm to voters in future elections. For example, the plaintiffs in *Whitford v. Gil* argued that based on an examination of a large body of election data, an observed efficiency gap of greater than 7% in a given election was strongly suggestive that bias in favor of the redistricting party would persist. In other words, the observation of a large bias was not only indicative of harm to voters that had already occurred, but it was also a strong predictor of likely future harm to voters. Analysis of historical data is not the only way to assess likely future harm due to bias, and we would like to use a method that takes advantage of as much specific information about the redistricting plan in question as possible. Therefore we employ an empirical statistical model to determine whether a particular redistricting plan will be unbiased *on average*, or if bias is a persistent feature of the plan. This allows us to distinguish between chance occurrences of bias on an otherwise fair map, and to assess the magnitude of likely future harm a map can be expected to cause. Full details of the particular model employed here are presented in the following section.

### 3 Modeling the Likelihood of Bias in Future Elections

### 4 Partisan Bias in the U.S. Congress 1972-2016

We present an examination of partisan bias in the United States congress for elections from 1972 to 2016. First, to illustrate how the effect of asymmetry on the composition of congress, we compute the specific asymmetry for each state and election, and compare the Democratic popular vote share to the actual Democratic seat share as well as the Democratic seat share that would have resulted if each election had zero asymmetry. This illustrates a few of the features of the specific asymmetry as a metric, but we note that this is not exactly representative of how the composition of congress would have been different had the proposed standard been implemented. The reason for this is that the proposed standard would enforce an expected asymmetry of zero, meaning on average no asymmetry would occur but in some elections asymmetry could occur by chance. This result is shown in figure 1, where it is evident that despite the natural multiplying effect of majoritarian elections, removing specific partisan asymmetry produces congressional representation closer to proportional with the popular vote, regardless of which party the net partisan asymmetry favors. Thus

while proportionality of seats and votes is not enforced, gerrymandering tends to exacerbate and amplify the natural tendency of single winner elections to lead to disproportionate representation in favor of the majority party, and therefore removal of asymmetry caused by redistricting tends to lead to results that are closer to proportional. If one were to use a metric that assumes votes should be proportional to seats, such as the efficiency gap, the 1970s would still be counted as gerrymandered, despite all partisan asymmetry being removed. That is, efficiency gap would mistake the unavoidable multiplying effect of the sigmoidal seats-votes curve for gerrymandering. I'm not sure about this, the eg does assume proportionality but it does not assume that the constant of proportionality between seats and votes is 1 Specific partisan asymmetry, however, never makes this error, while still leading to a substantial increase in the representativeness of congress as a whole.

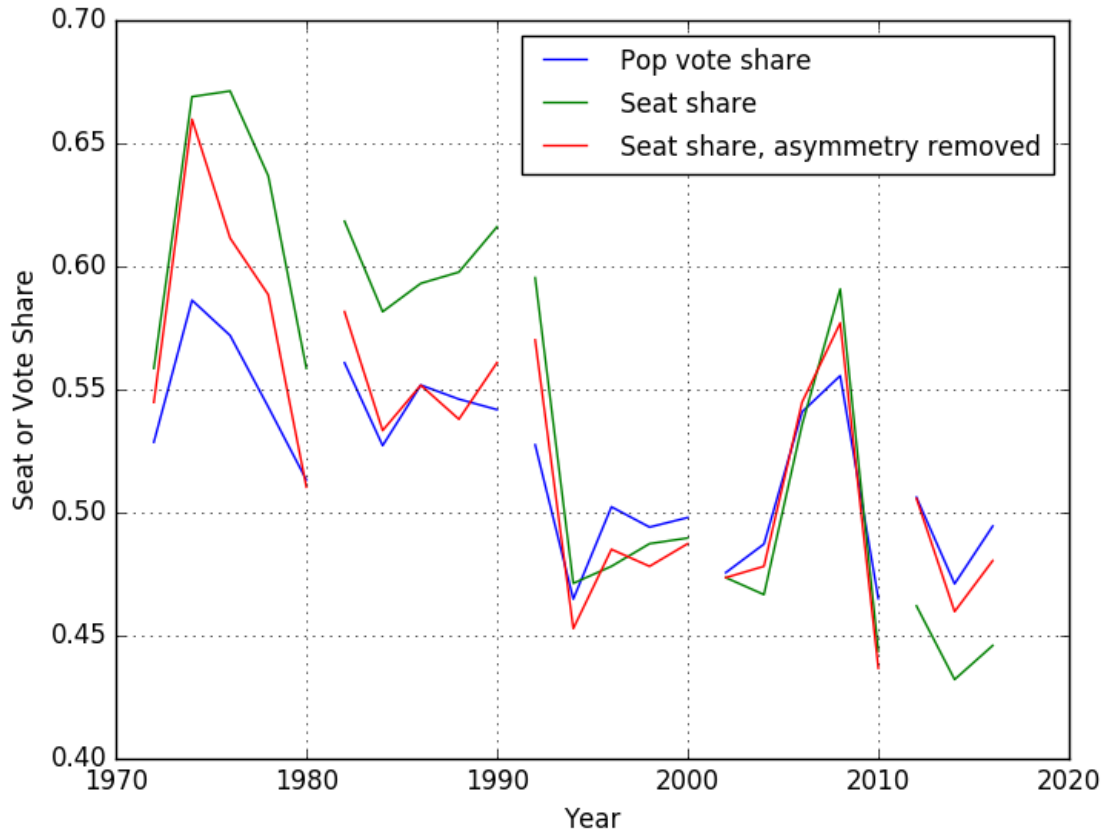


Figure 1: Democratic vote share and seat share with asymmetry removed 1972-2016

Next we account for random variability in the results by computing expected specific asymmetry for each state and redistricting cycle using 10,000 Monte Carlo samples. First, we consider the net asymmetry, which is indicative of how gerrymandering may have affected the partisan composition of congress. In figure 2, the net observed asymmetry is shown for each election. The expected net asymmetry for each redistricting cycle is also shown with 50% and 90% confidence intervals. A progression is observed where the net asymmetry strongly favored Democrats in the 70's and 80's, mildly favored Democrats in the 90's, was roughly zero net asymmetry in the 2000's, and strongly favored Republicans in the 2010 cycle. Histograms for the simulations can be found in figure 3, which shows the expectation and variance of the net specific asymmetry. In the 80's for example, Democrats could expect about 20 seats on average due to redistricting bias, with a negligibly small probability that the net national asymmetry would favor Republicans. This situation has reversed entirely by the 2010 cycle, where Republicans can expect 18 seats on average due to redistricting bias. Interestingly the 2000's show zero net bias on average, but this does not mean that the country was free of gerrymandering. As we will show later in this section, there was significant gerrymandering by Republicans and Democrats in their respective territories, but the net effect of these gerrymanders tended to roughly cancel.

The results for the 2010s showing significant Republican gerrymandering are consistent with the well known REDMAP program, where Republicans lead a highly successful effort to gain control of several state houses in advance of the 2010 redistricting. For earlier cycles, we observe qualitative agreement with similar studies CITE in terms of which party

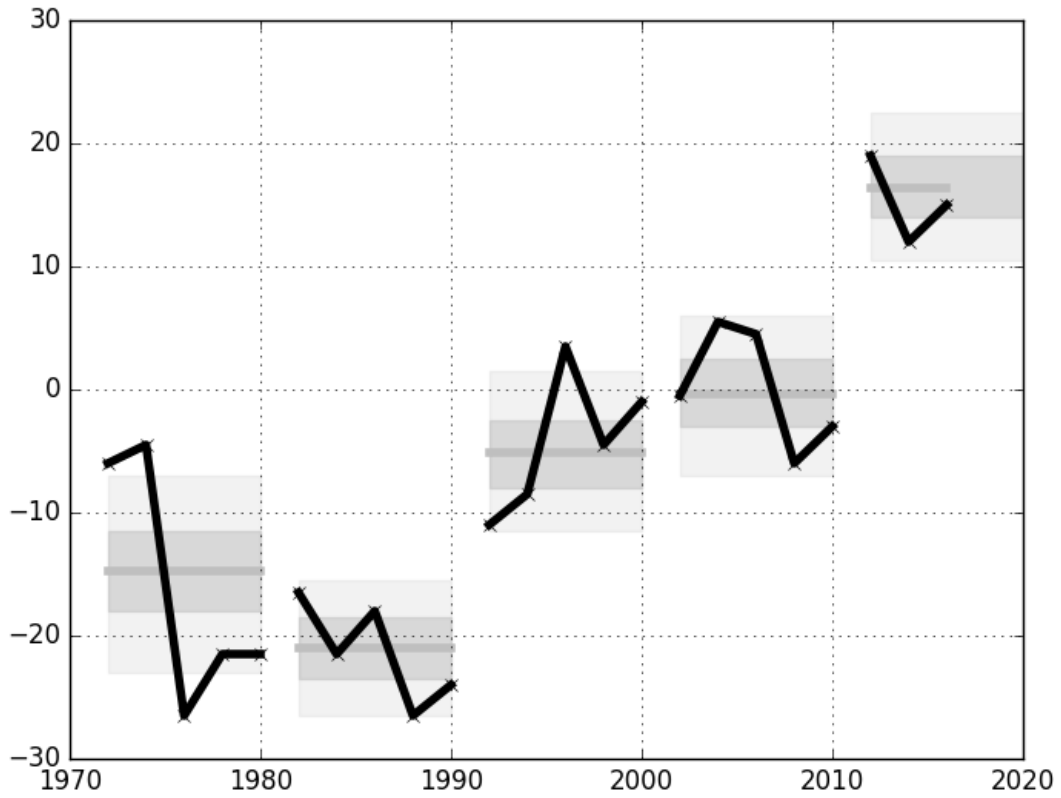


Figure 2: Net national asymmetry based on 10,000 simulations

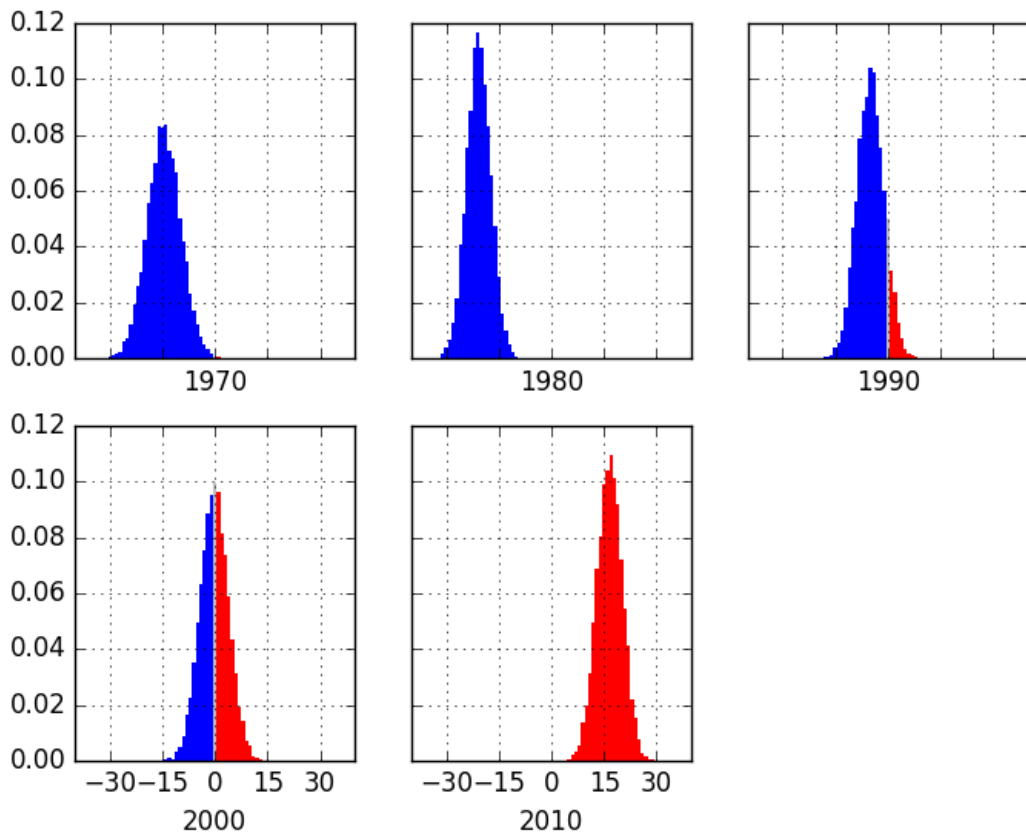


Figure 3: Net national asymmetry based on 10,000 simulations

benefited on net from gerrymandering. Quantitatively, we observe some disagreement in terms of the magnitude of gerrymandering, with the present approach attributing more seats to redistricting bias. This can be explained by differences in the analysis methodology. The simulation technique employed in CITE draws a 'fantasy' delegation from national voting

results in order to determine if a party is winning seats excessively relative to its vote share. National voting results are used as a baseline, but this baseline itself may contain substantial bias, which means only large gerrymanders will be successfully identified. In contrast, the present approach does not use a national baseline, and instead detects asymmetry in each state on a case by case basis.

Differences in methodology aside, the result that the expected asymmetry favoring Democrats in the 80's is larger than that favoring Republicans in the 2010 cycle is surprising. In the 2010 cycle, Republicans had the advantages of sophisticated redistricting algorithms, a high degree of geographic polarization, and low variability in voter preferences. These factors enable effective gerrymandering and would not have been available to redistricters in the 80's. On the other hand, Democrats controlled a significant number of states in the 80's and the national level of Democratic partisanship in congressional elections was fairly high, while it is closer to even in the 2010 cycle. Compared to Republicans in the 2010 cycle, Democrats in the 80's obtained seats due to small levels of bias in a number of states, and additionally obtained significant numbers of seats in larger states like California and Texas. Figure 4 compares the simulated asymmetry distributions for California and Texas in the 80's and for Pennsylvania and North Carolina in the 2010's. We can see that the variability of the asymmetry in Texas and California was higher than that of Pennsylvania and North Carolina. It is also worth noting that the number of seats available in these states in the cycles under consideration. In the 80's Texas has 24 seats and California had 43, while in the in the 2010's, Pennsylvania had 18 seats and North Carolina had 13. More district lines offers more opportunities for parties to manipulate redistricting to obtain a partisan advantage, and as a fraction of available seats, the gerrymanders in Pennsylvania and North Carolina are truly remarkable. This is consistent with the fact that conditions and technology to produce highly stable gerrymanders would exist to a greater degree in the 2010s than in the 1980s.

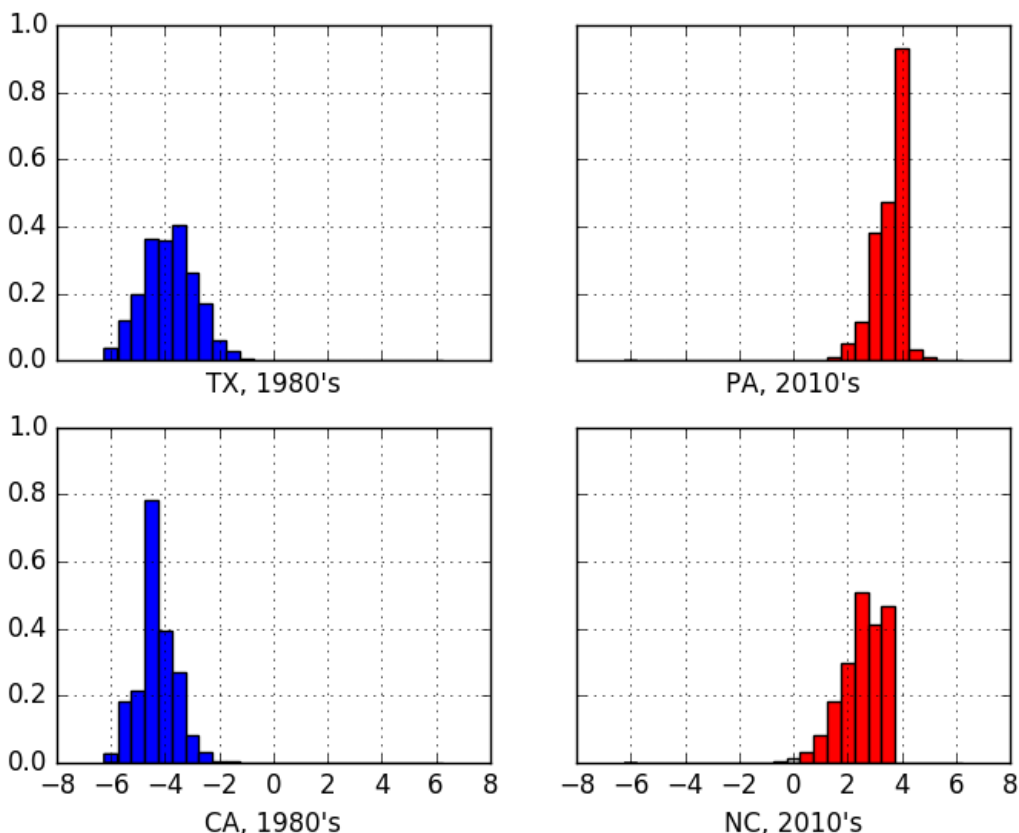


Figure 4: Histograms of the simulated asymmetries in Texas and California in the 1980's, and Pennsylvania and North Carolina in the 2010s

Net zero gerrymandering is preferable to a distorted congress, but net zero gerrymandering may still result in the congressional delegations of individual states being highly distorted. We would therefore also like to examine the total asymmetry, or the sum of asymmetries in all states without regard to which party the asymmetry favors. This tells us to what extent redistricting is being used to control election outcomes nationally. This is plotted for each

election cycle in 5, where we observe that the total asymmetry has held fairly constant over each cycle. **GOOD PLACE TO MENTION THAT ELECTIONS ARE A ZERO-SUM GAME - THAT TO THE EXTENT THAT THE REDISTRICTERS DETERMINED THE OUTCOME - THE VOTERS DID NOT.** **agree will add** Since there is a fixed number of seats available, every seat of partisan bias that redistrictors secure by gerrymandering is an equal amount of representation taken away from voters. In other words, the relationship between voter representation and partisan bias resulting from redistricting is that of a zero-sum game: what one side gains, the other side loses. Regardless of which party benefited from bias on balance, bias in redistricting appears to have affected the representation of millions of voters, for every congressional election since 1972.

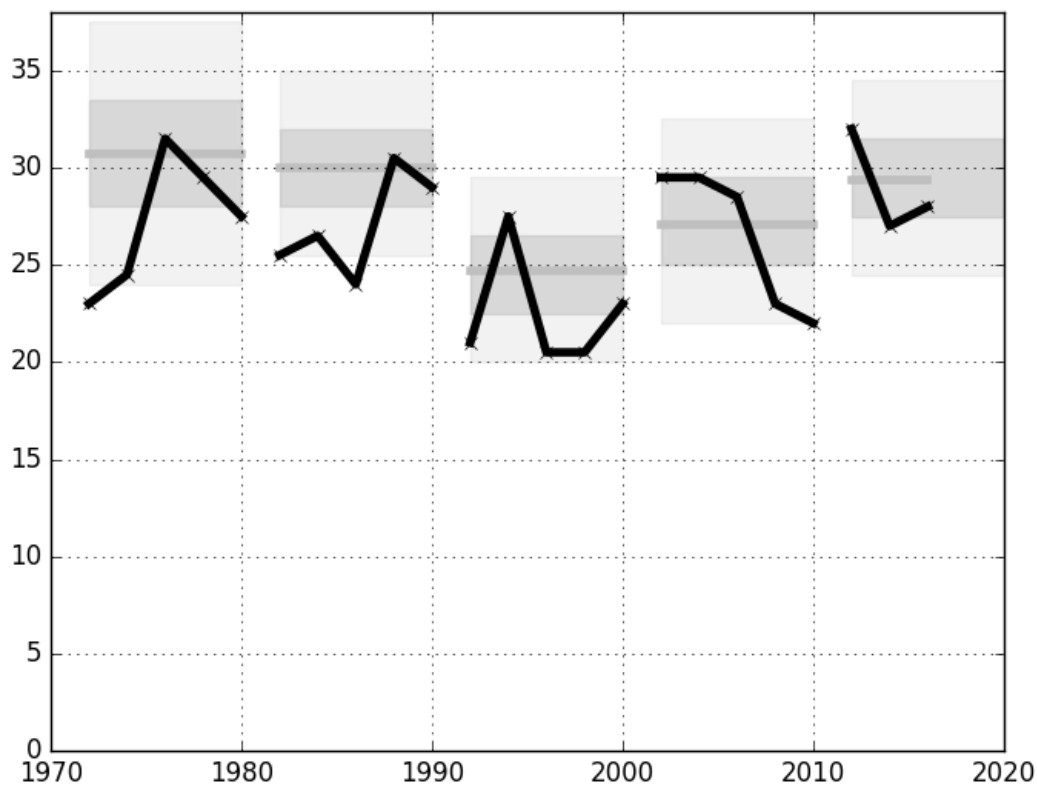


Figure 5: Total national asymmetry based on 10,000 simulations

While the total national asymmetry has held roughly constant over time, that is not to say that gerrymandering practices have held constant. To the contrary, some aspects of the 2010 redistricting cycle are quite unprecedented. Simultaneously, asymmetry has reduced in some states (at least in part due to redistricting reforms enacted in some of those states), while asymmetry has exploded in others. States known to have been targeted by the REDMAP program are gerrymandered very aggressively in terms of the number of seats available and the level of partisanship in those states. We can visualize how extreme the most aggressive gerrymanders are in each cycle are by sorting the expected asymmetry (as a percentage of available seats) in descending order for each state. States with fewer than five seats are excluded since small asymmetries result as very large percentages. This is shown in figure 6, where the 2010 cycle stands out as having three states with expected asymmetries that exceeded anything seen previously: Pennsylvania (19.7%), North Carolina (19.5%), and Georgia (16.7%) Following these are Virginia (13.3%), Michigan (12.4%), Ohio (11.6%), and Wisconsin (10.8%), which each have large asymmetries by historical standards. All of these states showed a sudden increase in asymmetry in favor of Republicans after the 2010 redistricting, see table 2 and figure 7. In contrast, California and New York had substantial asymmetries favoring Democrats in the 2000 cycle, and while they still show asymmetries favoring Democrats in the 2010 cycle, the magnitude of the asymmetry has decreased. **I believe CA implemented an independent commission starting in 2010** Previous cycles may have had moderate asymmetries in a larger number of states, but in terms of extreme gerrymandering, there has been no time since the Supreme Court ruled in *Reynold*



*v. Sims* (1964) that congressional districts must have roughly equal population, that is comparable to the 2010 cycle.

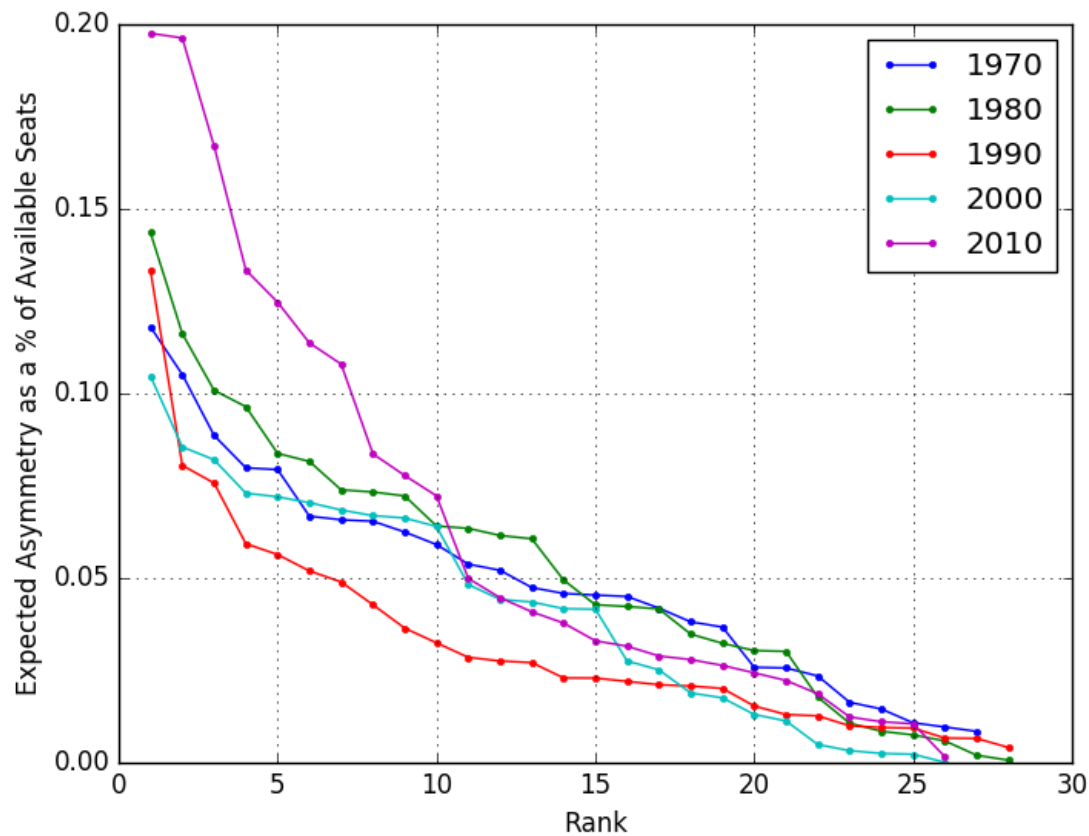


Figure 6: Expected asymmetry as a percentage of seats ranked in descending order for each redistricting cycle

Table 2: Change in Expected Specific Asymmetry from 2000 to 2010 for states with large asymmetries in 2010

State	Expected Specific Asymmetry, 2000 Cycle	Expected Specific Asymmetry, 2010 Cycle	Net Change
Pennsylvania	R 6.63%	R 19.76%	R 13.13
North Carolina	D 4.83%	R 19.63%	R 24.46
Georgia	R 4.16%	R 16.70%	R 12.55
Virginia	R 7.30%	R 13.35%	R 6.04
Michigan	R 6.69%	R 12.47%	R 5.78
Ohio	R 8.21%	R 11.37%	R 3.16
Wisconsin	D 1.89%	R 10.79%	R 12.68
California	D 8.55%	D 2.43%	R 6.12
New York	D 7.20%	D 3.15%	R 4.05
Illinois	R 1.30%	D 2.79%	D 4.10

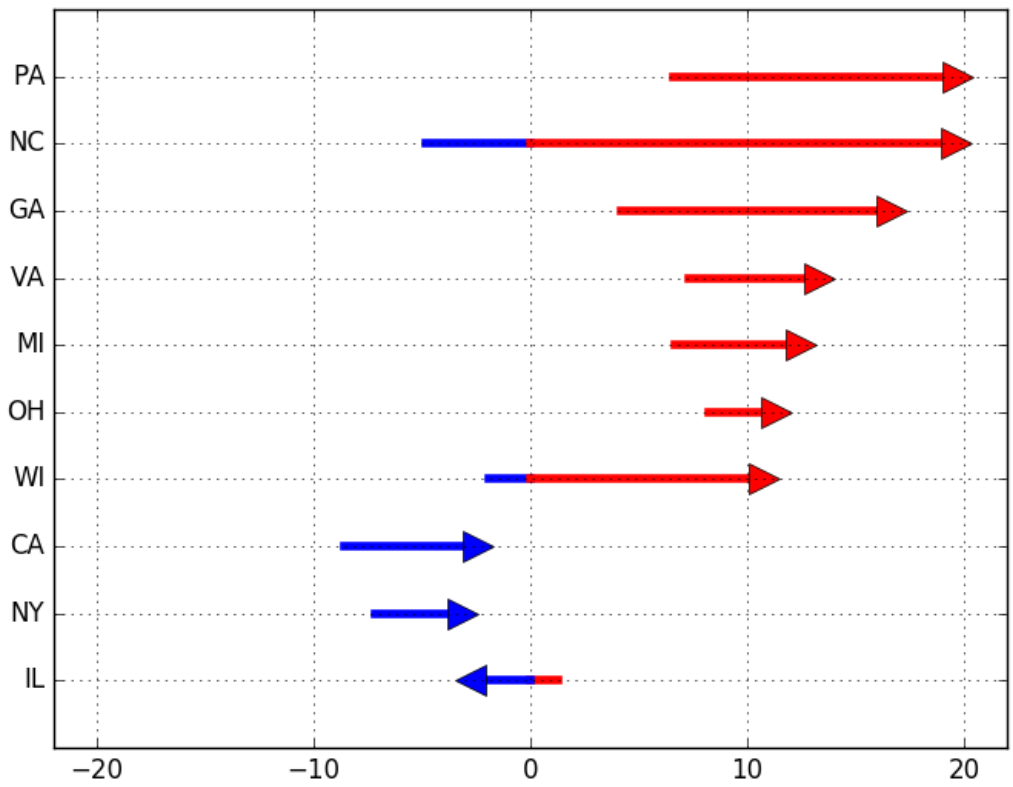


Figure 7: Changes in the expected specific asymmetry from the 2000 redistricting cycle to the 2010 cycle. The arrow indicates the direction of the change in asymmetry.

## 5 Partisan Bias in the Wisconsin State Assembly, post 2010 census

copied from 2000v2010

### 5.1 Method

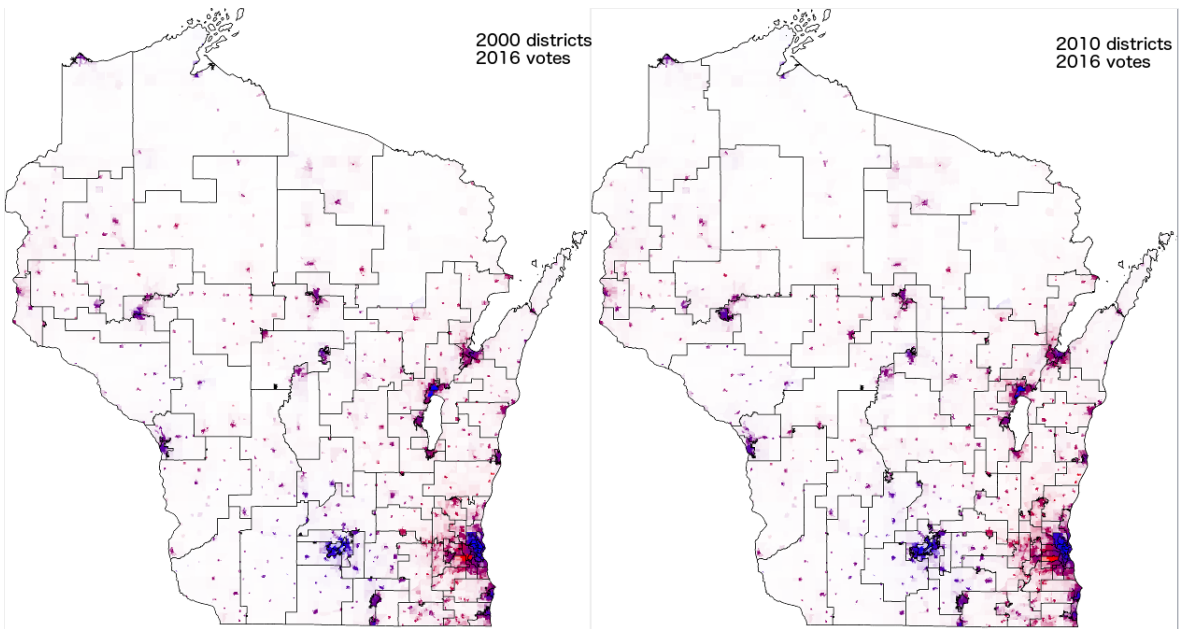


Figure 8: 2016 Assembly election results by precinct, projected onto Assembly districts before and after Act 43. Left: 2000 cycle districts, Right: 2010 cycle districts

To compare how much of the observed partisan asymmetry in 2010-cycle Wisconsin State Assembly districts was the result of changes in district designs versus changes in voter sentiment, we varied the first while holding the second constant. In other words, we projected actual election results onto both the current Assembly districts, and Assembly districts from the previous redistricting cycle, as shown in figure ??.

For this analysis we used official districting plans for the 2000 and 2010 census cycles, and actual vote counts from the 2006, 2008, 2010, 2012, 2014, and 2016 assembly elections, at precinct-level resolution, imputing uncontested elections with partisan swing matched presidential election results when available, and partisan swing matched federal congress election results when not. By "partisan swing matched" we mean that the Republican vote counts in all districts are multiplied by a constant factor so that the total vote count matches that for the assembly elections, for the subset of districts in which both elections were contested, and the same is done for the Democratic vote counts. Then election data for 2006, 2008, and 2010 was cross-aggregated at census block resolution to the 2010-cycle districts, and election data for 2012, 2014, and 2016 was cross-aggregated at census block resolution to the 2000-cycle districts.

Since the same exact elections are used in both the 2000 districts analysis and 2010 districts analysis, *all* differences in the results are the consequence of the redistricting scheme used, and conversely, changes in voter sentiment have exactly *zero* contribution to the difference.

Furthermore, if the 2000 census cycle districts show a partisan asymmetry much lower than the 2010 cycle, then we know that natural "political geography" – the tendency of Democratic voters to be clustered in cities – can be immediately ruled out as a significant factor in the partisan asymmetry inherent in 2010 cycle Wisconsin Assembly districts, since any contributions from political geography would be present in *both* districting schemes. However if they are not substantially different, this would not demonstrate that political geography was a significant contribution, because they both could be artificially gerrymandered on top of any unintentional gerrymandering due to political geography. In this case, further analysis would be required to separate the effects of unintentional bias caused by political geography and deliberate bias caused by gerrymandering.

## 5.2 Measurement

To measure partisan bias, we measured the asymmetry in the seats-votes curve at the actual vote count that occurred in the election. In other words, we take the number of seats that Republicans won, and subtract from that the number of seats that Democrats would have won if the popular vote percentage was reversed ( $1-x$ ), and divide by 2. This represents the number of seats Republicans won in excess of what they would have won if the seats-votes curve was symmetric. If the number is negative, that represents an asymmetry in favor of Democrats. We call this "specific partisan asymmetry".

Before applying this metric to concrete examples, we'd like to highlight some of the advantages of it.

- It doesn't assume proportionality in results - The seats-votes curve for single-winner elections naturally take on a cumulative Beta-binomial distribution, as opposed to a diagonal line representing seats proportional to votes. This metric doesn't assume a diagonal curve representing proportionality. It tests a looser (less strict) criteria: asymmetry, which still retains the essential feature of measuring disenfranchisement based on political belief. SCOTUS has ruled that seats proportional to votes can not be a legal standard (Thornburg v. Gingles 1986). But they remain open to a legal standard based on the idea of partisan symmetry (LULAC v. Perry 2004).
- It distinguishes between artificial partisan bias and the natural multiplying effect of single-winner elections - Single-winner aka "Majoritarian" elections naturally over-favor the majority party, giving them a larger fraction of the seats than they get of the popular vote. While some measures might mistakenly identify this as gerrymandering, specific asymmetry explicitly takes this into account by calculating and then subtracting this natural multiplying effect. In other words, it does not mistake disproportionality due to the sigmoidal shape of majoritarian election seats-votes curves with gerrymandering.

- It works for states with partisanship far from 50/50 - By sampling the asymmetry at the actual vote counts rather than e.g. 50/50, this metric maintains full relevance even for the most extremely partisan states.
- It represents deviation from what is practically achievable - It is trivial for a computer algorithm to design districts that make the seats-votes curve perfectly symmetric, while satisfying all constitutional requirements. Thus, to convert this absolute score to one relative to what can be accomplished by a redesign that meets all traditional redistricting principles, one simply subtracts zero. This "new" relative score gives you the number of voters whose votes were effectively stolen by the map drawers with the current map, and whose right to representation can be returned by a remedy map that can be trivially designed by a computer employing a multi-objective heuristic optimization algorithm.
- The only possible improvements to it are shift, scale, and shape - Since it is well ordered and monotonic with respect to actual partisan advantage, any "better" measure can only be better in the sense of having a more appropriate "center" point, a more appropriate scale, or a more appropriate "shape" (change of slope as you go up and down the scale). The scale varies from -1 to 1, or alternatively from negative the number of seats available to the positive of that. The "center", zero, is a non-arbitrary and very reasonable choice, representing a perfectly symmetric situation. The "shape" is the number (or percentage) of congressional seats affected.
- It gives a result in an empirically meaningful unit: number of seats, which can be directly converted into number of ballots/voters affected by multiplying by the number of ballots cast and dividing by the number of districts.

### 5.3 Observed partisan asymmetries in past elections

After imputing uncontested elections and projecting onto the two different districting schemes, we counted up the wins for each election on each map. We also counted up the wins for the other party if the popular vote was reversed, subtracted this from the first, and divided by 2, to get the specific partisan asymmetry. The seat counts are shown in figure ??, and the specific asymmetry is shown in figure ??. They are also tabulated in table ??. The horizontal lines and shaded regions represent the statistical expectation and 50% confidence intervals, respectively. Their calculation is described in the next subsection.

The charts show a clear and stable separation between both seat counts and partisan asymmetry between 2000 districts and 2010 districts. They both go up and down largely in parallel, responding about the same to changes in voter sentiment.

This tells us that in the 2012, 2014, and 2016 elections, the new districts led to an increase in Republican representation over what the old districts would have produced. Furthermore, were the new districts used in the 2000 redistricting cycle, that would have produced about the same increase in Republican representation for the 2006, 2008, and 2010 elections.

In figure ??, the difference in seats won by Republicans is plotted along with the difference in the expected asymmetries between the two maps. The increase in seats gained by Republicans in the 2010 map is entirely explained by the increase in asymmetry in the 2010 map. Therefore the 2010 map leads to greater Republican representation by *structurally disadvantaging Democratic voters*.

We can better visualize how the 2010 districts give Republicans more seats than the 2000 districts by coloring in the districts according to the victory margin, and comparing. Figure ?? shows such a comparison for 2012, 2014, and 2016 elections (left to right). The two Democratic districts above Lake Winnebago were combined into one, thus taking a seat away from Democrats. Democratic voters in Eau Claire once participated in two competitive districts, which would have given Democrats 2 seats with 2000 district lines, but the 2010 district lines wrapped them up in a single safe district, capping the number of elections those voters can swing at 1. In 2000 districts, Madison is surrounded by a bunch of competitive Democratic-leaning districts. But the 2010 lines packed Democratic voters surrounding Madison into a few safe Democratic districts, flipping the districts those voters came from, which would have been competitive, to Republican. In Milwaukee, the lines were redrawn to more tightly outline Democratic voters, thus creating a few more districts that Republicans

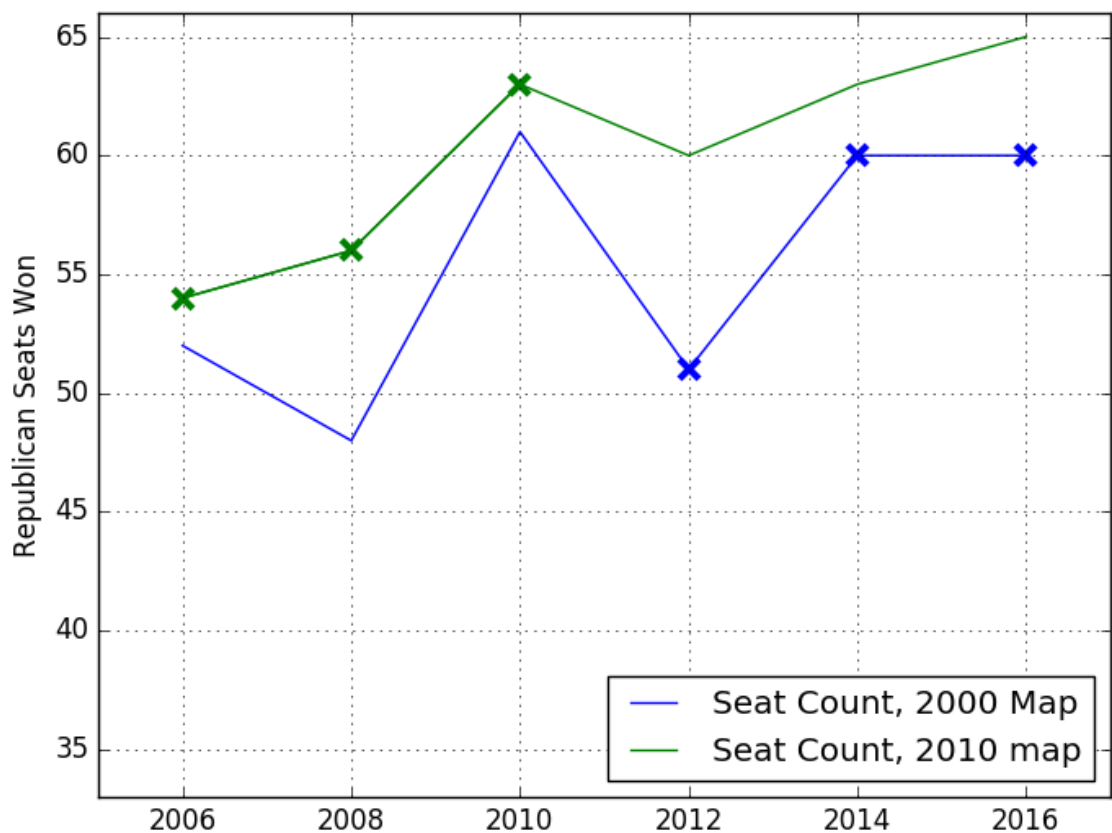


Figure 9: Republican Seat Counts in Assembly elections, using 2000 cycle districts vs 2010 cycle districts. The X's indicate that a result has been cross aggregated, meaning results were computed on a different district map than the one that actually existed at the time of the election.

Table 3: Summary of the differences in seat counts and asymmetries for elections on the two legislative maps

Election Year	Rep. Seats, 2000 Map	Rep. Seats, 2010 Map	Net Change in Seats
2006	52	54	R + 2
2008	48	56	R + 8
2010	61	63	R + 2
2012	51	60	R + 9
2014	60	63	R + 3
2016	60	65	R + 5
Election Year	Asymmetry, 2000 Map	Asymmetry, 2010 Map	Net Change in Asymmetry
2006	6.5	7.0	R + 0.5
2008	2.5	9.0	R + 6.5
2010	0.0	6.5	R + 6.5
2012	3.5	12.0	R + 8.5
2014	3.5	11.0	R + 7.5
2016	7.0	12.0	R + 5.0

can win in. Indeed, figure ?? shows that by comparing the districting schemes using 2016 votes, it is easy to identify where Republicans picked up 4 seats due to partisan vote packing.

Meanwhile, deeper red districts throughout the state were redrawn to give away some of their Republican voters to nearby competitive districts, thus reducing the likelihood that a Democratic representative was elected in the once-competitive districts, while still maintaining a safe edge in the originally deeper red district. You can see this by noticing the much "flatter" red color in the bottom maps versus the top maps. It will also be made clear in the next subsection when we examine the statistical properties of the maps more rigorously.

In sum, the changes in the district lines from 2000 districts to 2010 districts consist of

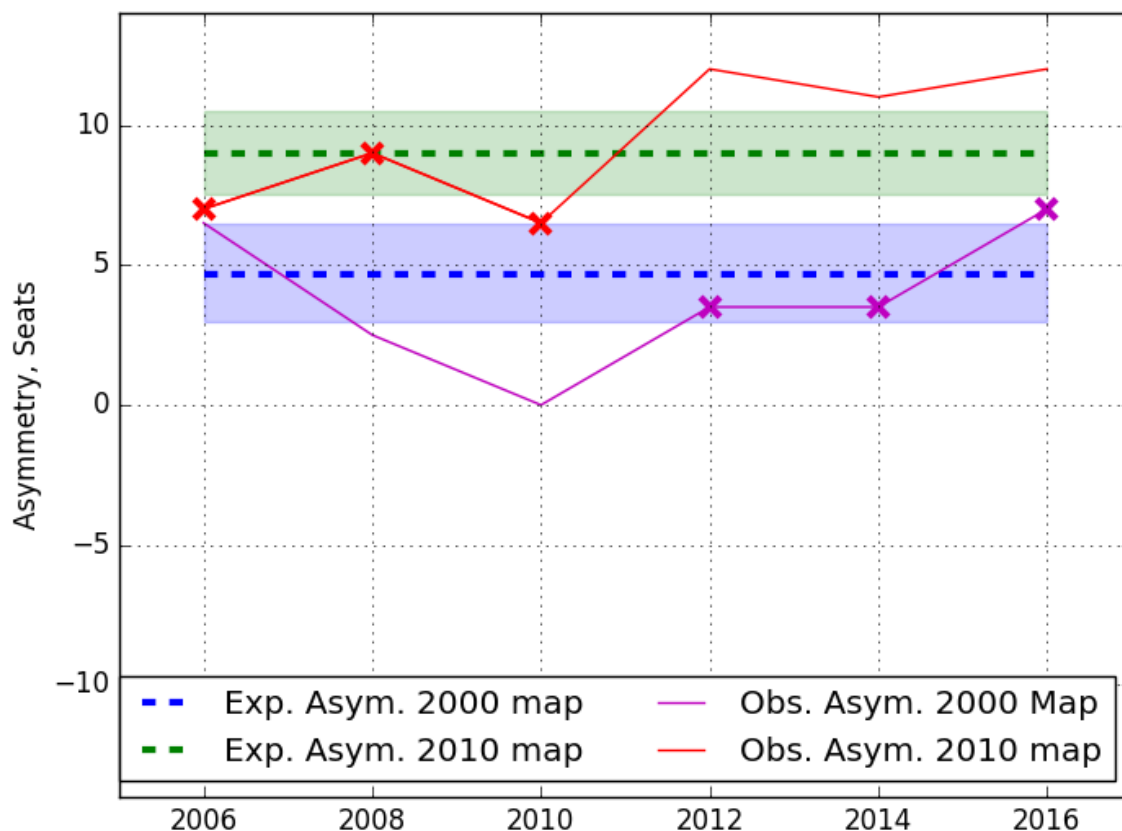


Figure 10: Historical asymmetry in Assembly elections, using 2000 cycle districts vs 2010 cycle districts. The X's indicate that a result has been cross aggregated, meaning results were computed on a different district map than the one that actually existed at the time of the election. The dashed lines represent the statistical expectations, and the shaded areas represent the 50% confidence intervals.

packing Democratic voters, thus decreasing the number of Democratic districts, and cracking Republican districts, thus increasing their number. Instances of the reverse – of packing Republican voters and cracking Democratic districts – are simply not present.

## 5.4 Expected partisan asymmetries in future elections

### 5.4.1 Probability model

To calculate the expected partisan asymmetries for future elections, we need to construct a probability model and compute the most likely parameters for it based on past observations. We choose our probability model by deduction from first principles.

An election between two parties is a number of repeated trials, each resulting in one of two outcomes. This is known in mathematical literature as a Bernoulli process. In our case, the rate of "success" (of getting the first of the two options) in the Bernoulli process – known as the "rate parameter" – is unknown. For reasons which we won't go into here, the standard accepted distribution for modeling an unknown rate parameter of a Bernoulli process, is a Beta distribution. Using a Beta distribution here amounts to making the fewest possible assumptions about the value of the rate parameter.

For our probability model we model at the resolution of districts vote percentages. We use a Beta distribution for the vote counts in each district, and one for the total popular vote. Thus we empirically estimate a probability model that represents the entire range of plausible outcomes for the election results under a particular redistricting plan. Sampling the model allows us to calculate a likelihood for every possible outcome, without relying on any prior assumptions. We can therefore not only determine the presence of partisan asymmetry in the outcome of one or more election results, but we can determine whether or not and to what degree partisan asymmetry is a persistent feature of a given plan.

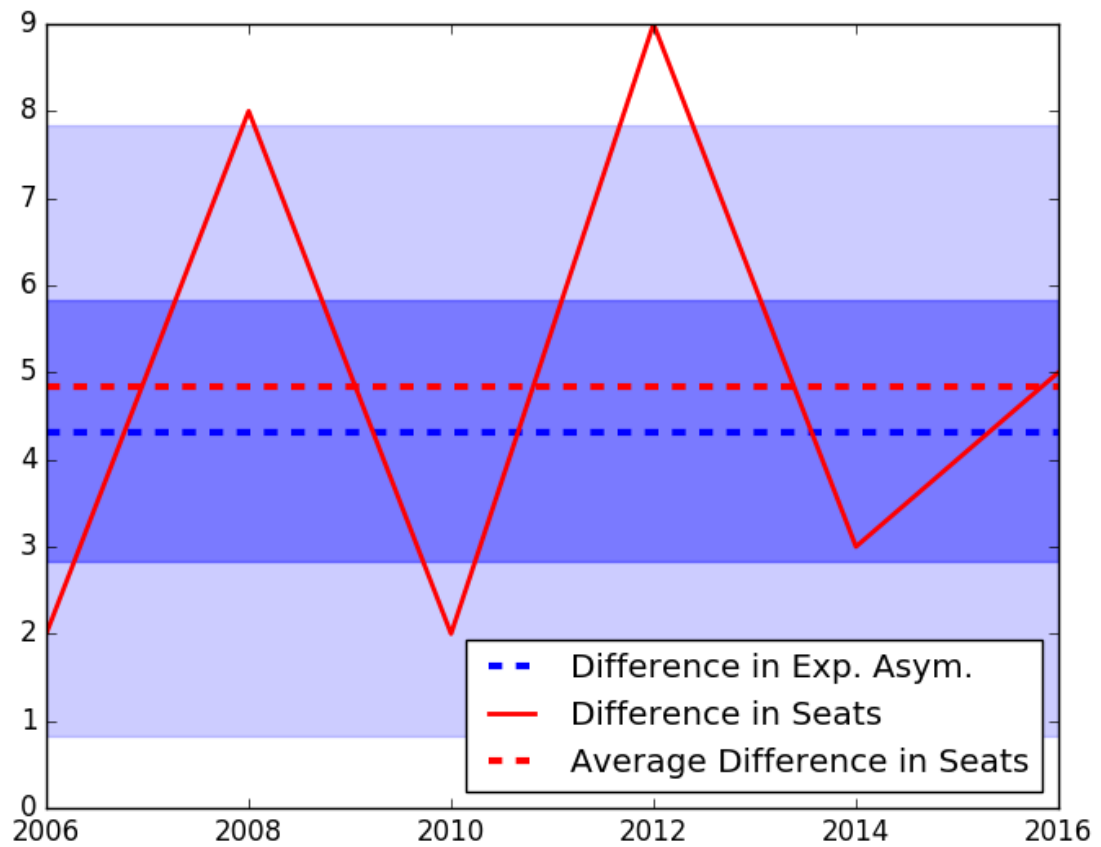


Figure 11: Difference in seats won by Republicans in the two maps. The dashed blue line shows the difference in expected asymmetry between the two maps, the blue and light blue shaded regions represent 50% and 90% confidence intervals

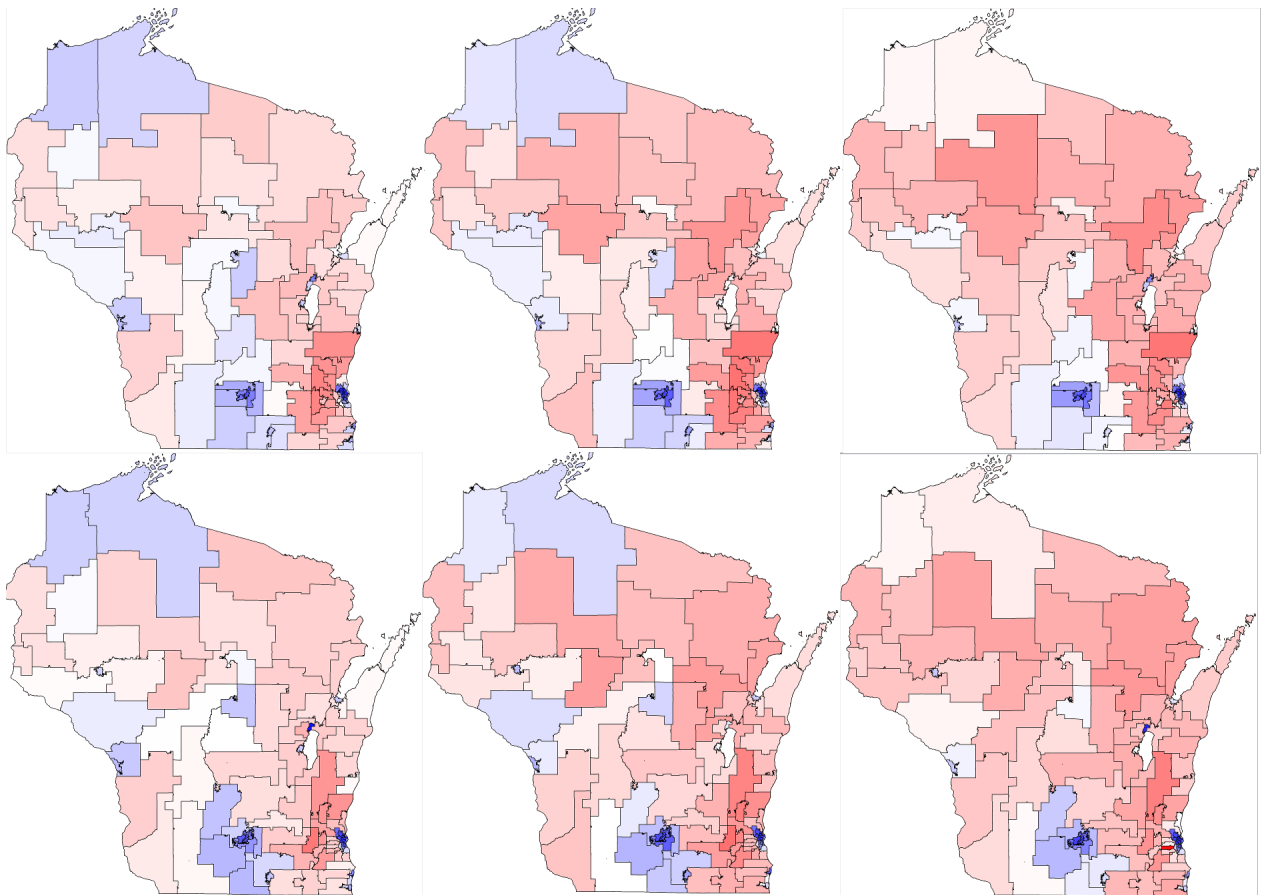


Figure 12: Partisan vote packing in Wisconsin Assembly elections. Top: 2000 cycle districts, Bottom: 2010 cycle districts, Left to right: 2012, 2014, and 2016 elections



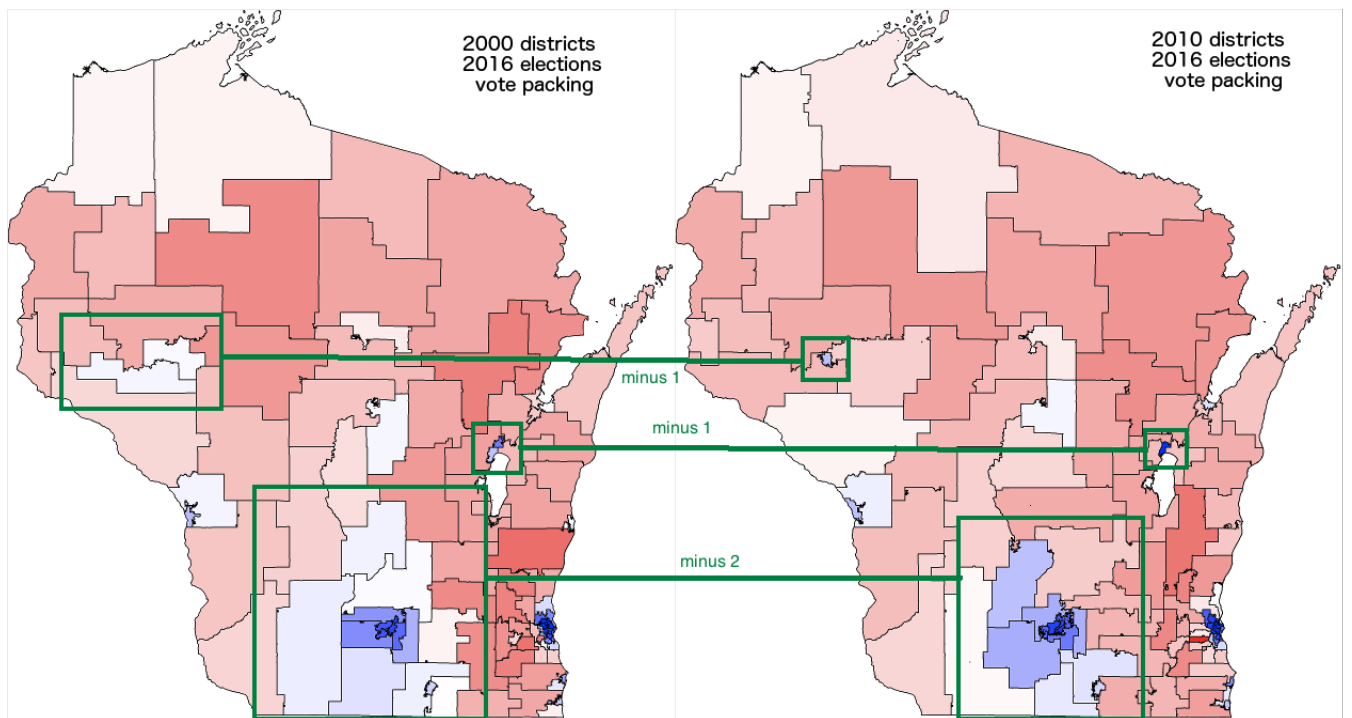


Figure 13: Partisan vote packing in Wisconsin Assembly elections, using 2016 election data. Outlined in green are areas where Republicans packed Democratic votes.

Before applying this model to a concrete examples, we'd like to highlight some of the advantages of this approach.

- It takes into account non-uniform swings - By looking at not only the average partisanship of each district, but how much the partisanship of each individual district changes from election to election, this metric is able to take into account the changes in voter sentiment that are not uniform throughout the state; that are perhaps concentrated in only certain districts.
- It takes into account all possibilities, weighted by likelihood - Every possible seats-votes curve and every possible popular vote ratio is taken into consideration, weighted by the combined likelihood of the two.
- It shows durability - By computing an entire likelihood function for specific partisan asymmetry, rather than a single point estimate, this approach enables quick and accurate assessment of how durable a gerrymander is; how much more harm it will cause in the future, including what the likelihood is that it will not cause harm.

#### 5.4.2 Future election result likelihoods

Shown in figure ?? is a graph of district partisanship, including dispersion. On top are 2000 cycle districts, and on bottom are 2010 cycle districts. To generate this graph, a Beta distribution for each district was calculated using the method of moments. Additionally a Beta distribution for the total popular vote was calculated using the method of moments. Then the probability density function was plotted out for each one of them (The black curve in the middle is the total popular vote.) Results that lead to a Democrat winning the seat are colored in blue, and results that lead to a Republican winning the seat are colored in red. Note, in addition to the average of the past 6 elections, the graph also shows the statistical dispersion, from which the durability of this effect for future elections can be assessed.

Both the 2000 and 2010 distributions show clear signs of pro-Republican gerrymandering: many Republican-leaning districts are tightly clustered at levels of partisanship that allow Republican candidates to win safely without wasting too many votes, while many Democrat-leaning districts are heavily packed and not competitive. However, in the 2010 distributions Republicans-leaning districts are more concentrated around a single point slightly but safely short of competitive, while Democratic-leaning districts that were competitive in the 2000 distributions have been shifted away from the center. Consequently, in the 2010 distributions, compared to the 2000 distributions, Republicans are expected to gain more seats by smaller margins, while Democrats are expected to gain fewer seats by larger margins.

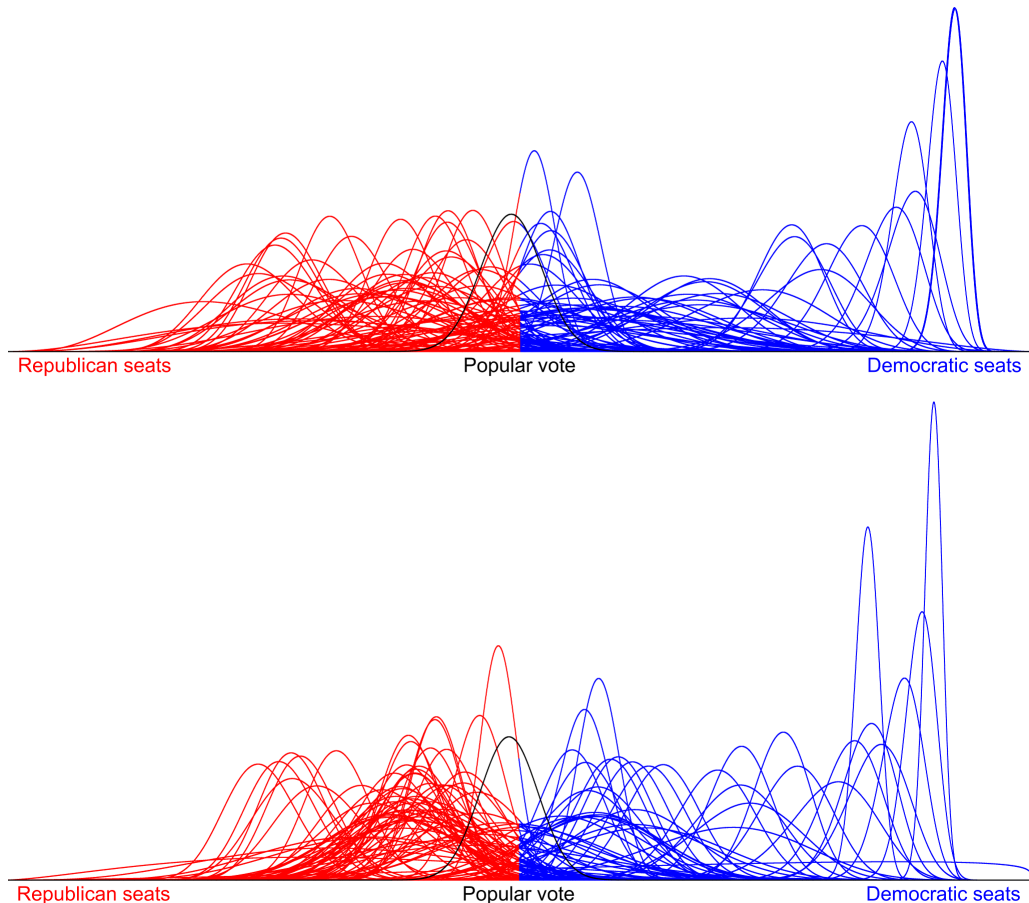


Figure 14: Beta distributions for Wisconsin 2010-cycle Assembly districts, Top: 2000 cycle districts, Bottom: 2010 cycle districts

Shown in figures ?? and ?? are twelve seats-votes curves drawn from these distributions (solid curve), along with their 2-axis reflection (dashed curve), and a popular vote percentage drawn from the statewide Beta distribution (vertical dotted gray line). While a small asymmetry in favor of Republicans is sometimes apparent in the 2000 curves, a large and consistent asymmetry in favor of Republicans is immediately evident in the 2010 curves.

#### 5.4.3 Numerical integration of likelihood functions

#### 5.4.4 Specific partisan asymmetry likelihoods

To get the likelihood function for partisan asymmetry, we use Monte-Carlo numerical integration. In each realization, we subtract the seat counts Democrats would get with that share of the vote from the number Republicans would get under the same share of the vote, and then divide by 2. The results are shown in figure ??, as a fraction of total available seats. These results were also used to compute the expectations and confidence intervals shown in figures ?? and ??.

For the 2000 districts, the partisan asymmetry favors Republicans, giving them an average seat boost of 10% of the available seats beyond what a symmetric seats-votes curve would produce. However, there is still a decent likelihood that voter sentiments would change enough that Democrats would get a small boost from seats-votes curve asymmetry.

For the 2010 districts, the partisan asymmetry favors Republicans by twice the amount of the 2000 districts, giving them an average seat boost of 20% of the available seats beyond what a symmetric seats-votes curve would produce. The most likely result in the Wisconsin state assembly might be that Republicans get 6/10ths of the legislative seats, whereas if the popular vote count were reversed, Democrats would get only 4/10th of the legislative seats. Remedying this would give Democrats 10% more seats. This means that roughly 10% of the entire Democratic-voting population of the state of Wisconsin were effectively disenfranchised, in violation of the one person, one vote principle. Conversely, 10% of the Republican-voting population effectively got an extra vote. That's approximately 306,000 felonies.

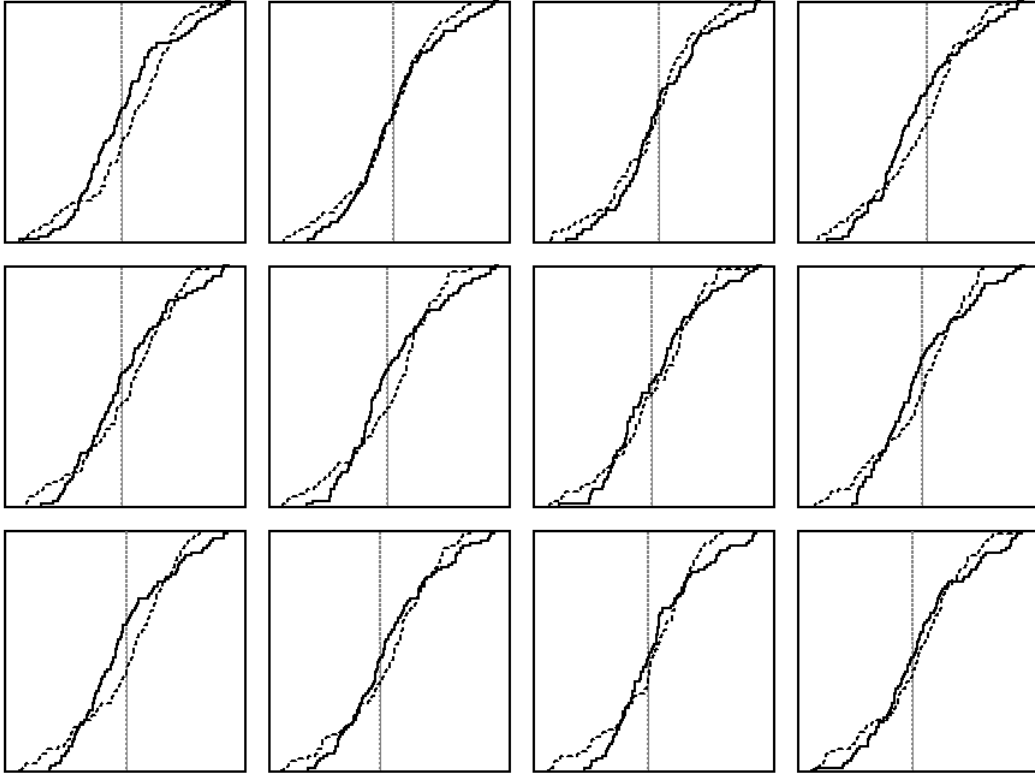


Figure 15: Some election outcome seats-votes curves generated from the probability model for Wisconsin 2000 Assembly districts

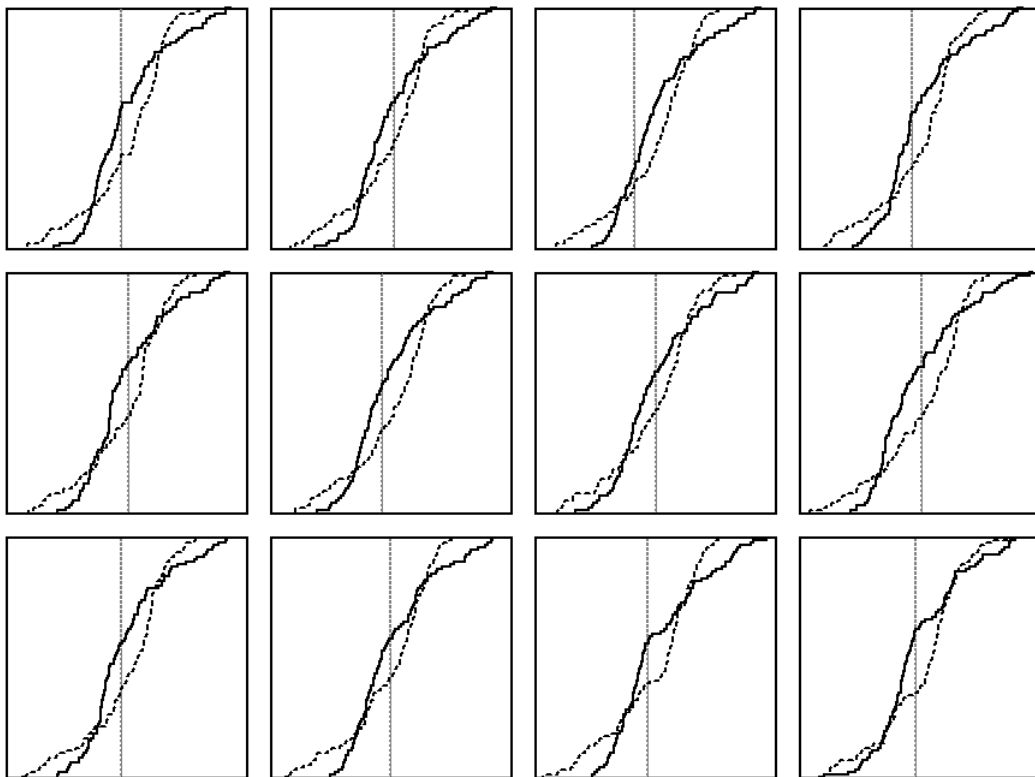


Figure 16: Some election outcome seats-votes curves generated from the probability model for Wisconsin 2010 Assembly districts

To put this figure in context, we conducted a study of asymmetry in the US federal congress from 1972 to 2016, where we found that asymmetries exceeding 10% of the available seats were rare and appeared only in the most aggressively gerrymandered states. The two largest gerrymanders in federal congress in this time period were North Carolina and Pennsylvania in the 2010 cycle, where the asymmetry was about 19% of the available seats for each of these states.

Unlike the 2000 districts, out of 100,000 samples drawn, *none* of them resulted in specific partisan asymmetry favoring Democrats. Since an assembly district election occurs every 2 years, this shows that, without significant changes in geo-spatial demographics, the asym-

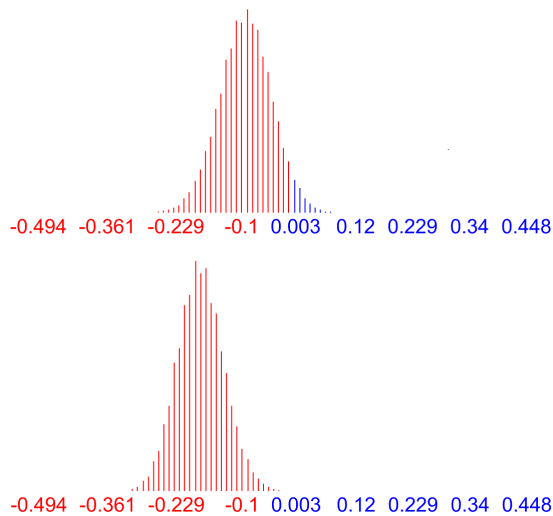


Figure 17: Specific asymmetry likelihoods, Top: 2000 cycle districts, Bottom: 2010 cycle districts

metric pro-Republican partisan advantage inherent in this redistricting will persist for the next 200,000 years.

#### 5.4.5 District partisanship histogram

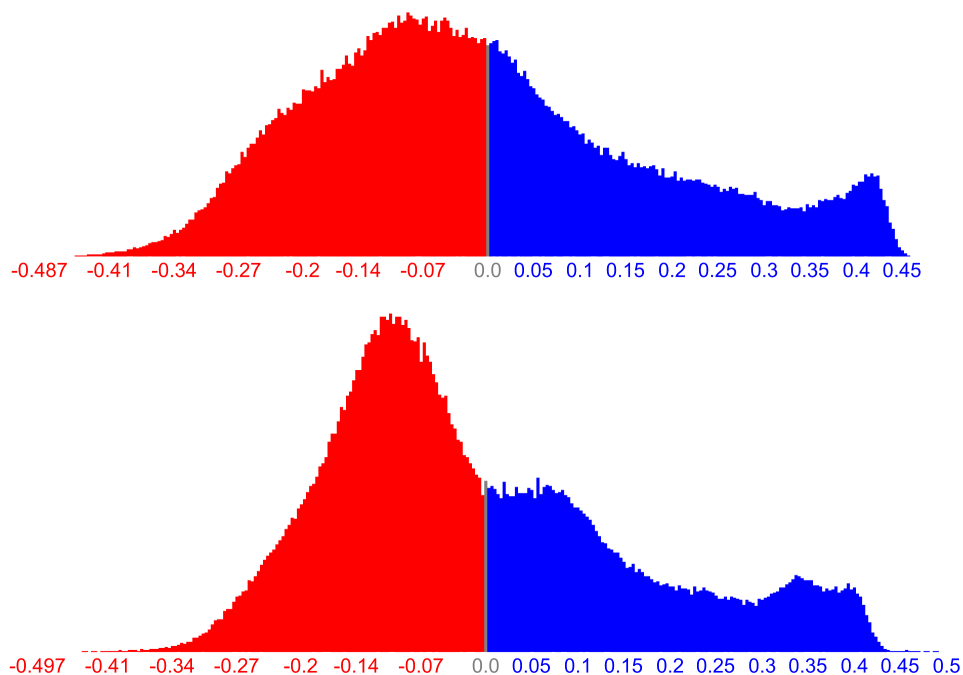


Figure 18: District partisanship likelihoods, Top: 2000 cycle districts, Bottom: 2010 cycle districts

Another way we can see this is by looking at the histogram of district partisanship.

We used Monte Carlo integration to construct the likelihood function for district partisanship, over all districts, and all possible election outcomes. In other words, if you were to pick a random district in a random election, this curve shows how likely the vote in that district is to be 25% Republican vote, a 50-50 split, etc. This is essentially what you'd see if you took the lines in figure ??, and stacked then on top of each other instead of behind each other.

The curve for the 2000 districts show a smooth peak near 50/50, with a slight skew towards Republicans and a long tail on the Democratic side with a brief peak at the end. This shows that Democratic districts were packed, while Republican districts were cracked.

On the other hand the curve for the 2010 districts show a much sharper distinction. It has a high red peak close to the center but still safely to the side, that already drops off sharply by the time it hits 50/50, and then on the blue side, drops slowly all the way

until it gets to a larger bump at the extremely packed end of the scale. This clearly reveals that Republican candidates in Republican leaning districts have a high likelihood of winning safely but without wasting too many votes, while Democratic candidates in Democratic leaning districts have a high likelihood winning by margins that waste a lot of votes. You can also see this effect visually on the maps in figure ??: the red colors are much "flatter" in 2010 districts (bottom) than 2000 districts (top).

#### 5.4.6 Statewide seat count likelihoods

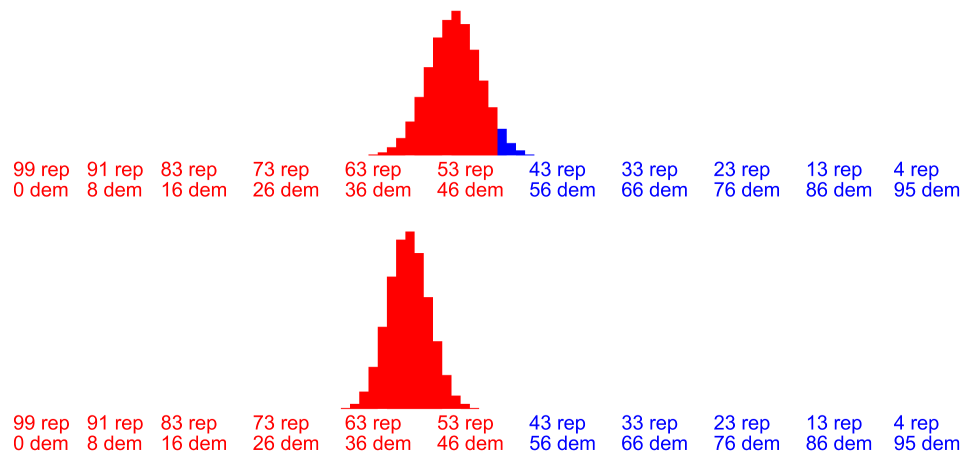


Figure 19: Seat count likelihoods, Top: 2000 cycle districts, Bottom: 2010 cycle districts

Using "draws" from the Beta distributions as shown in figures ?? and ??, we estimate the likelihood of every possible election outcome, district-by-district. We then accumulate these outcomes on a chart, with the seat count on one axis, and the likelihood on the other. The results are shown in figure ??.

In 2000 districts, the average outcome – also known as the "expected" outcome – is Republicans winning a 54-45 majority of seats, though Democrats still have a decent chance of winning the majority of seats.

In 2010 districts, Republican's expected lead grows by an additional 10 seats to 59-40, and despite being favored by roughly half the voters, Democrats have no chance at all to win a majority of seats, even in the most eccentric swings of voter sentiment. Indeed, out of 100,000 samples, Democrats rarely even get 45 of the seats – a threshold that, under the 2000 districting scheme, would be their average.

## 5.5 Conclusions

Measured by partisan asymmetry, both expected and observed, there was substantial gerrymandering in favor of Republicans in the 2000 cycle maps, leading to about 4 extra Republican seats beyond what would be allocated if the seats-votes curve was symmetric. The current method does not distinguish what part of this gerrymandering was deliberate and what part unintentional, though regardless of the proportion, it is trivial to generate a map without this asymmetry by using one of many well-known optimization algorithms.

If we disregard data that was not available when the 2010 cycle districts were drawn, and use only data from the 2006, 2008, and 2010 elections, we see that with the new redistricting lines, Republicans are expected to gain an additional 4 assembly seats, beyond what the previous (2000 cycle) redistricting lines would have given them. Note that for every seat Republicans gain, Democrats lose a seat. So in terms of seat *difference* gained, it's 8. Furthermore, this is in addition to the partisan asymmetry already present in the 2000 cycle maps, bringing the expected total seat difference gained over Democrats through partisan asymmetry to 18.

The fact that the 2010 cycle maps have double the partisan asymmetry as the 2000 maps, both expected and observed, demonstrates that regardless of what part of the gerrymandering in the 2000 cycle maps was deliberate, *at least* half of the gerrymandering in the 2010 cycle maps was deliberate. Regardless of what portion of the remainder is unintentional, a map optimized by a computer algorithm to maximize partisan symmetry could

remove almost all of the deliberate gerrymandering, and at least some of the unintentional gerrymandering.

From these comparisons, it is painfully obvious that the specific partisan asymmetry present in the 2010 cycle districts is *not* present in the 2000 cycle districts, even when using vote count data from the exact same elections. Therefore:

- The difference between the two schemes is a stable 10% gain in the Republican seat margin over Democrats. Since the same exact voter sentiments were used in both analyses, this difference in partisan asymmetry simply cannot be a consequence of changes in voter sentiment. Since the only thing that we changed was the districting scheme, all observed differences in partisan asymmetry are purely the result of changing the districting scheme and reflect the choices made by the redistricters and do not reflect changes in political geography.
- In order to claim that natural political geography was a significant factor in the 2010 cycle district, we would expect to see the exact same asymmetries in the 2000 districts. To the contrary, partisan asymmetry in 2010 districts is double that of 2000 districts, even after compensating for observed and potential changes in voter sentiment.
- Claims that the observed partisan bias in 2010 districts is simply a consequence of political geography are categorically false.
- Claims that the observed partisan bias in 2010 districts is simply a consequence of changes in voter sentiments are categorically false.
- Simply reverting back to the 2000 census cycle districts would remedy most of the present injustice, as measured in seats affected or ballots affected.

This comparison cannot estimate how much deliberate gerrymandering was present in both schemes, nor can it estimate how much of the gerrymandering present in the 2000 districts was deliberate. However it does prove with absolute certainty that the 2010 districts, under the same elections, give Republicans more seats than the 2000 districts, and that this difference will persist even in the most extreme swings of voter sentiment.

Currently Assembly Republicans have 64 seats, and Democrats have 35. If the district lines were not changed after the 2010 cycle, and everyone voted the same, Republicans would currently have 4 less seats in Wisconsin Assembly, and Democrats would have 4 more, making the count 60-39. If the district lines were instead optimized to favor neither party – to have an expected specific partisan asymmetry of zero – Republicans would currently have 9 less seats in Wisconsin Assembly, and Democrats would have 9 more, making the count 55-44. For comparison, the historical threshold for ruling in favor of plaintiffs on grounds of racial gerrymandering is 1 seat.

end copying from 2000v2010

## 6 Conclusions

# Acknowledgment

## References

- [1] M. G. Kendall and A. Stuart. The law of the cubic proportion in election results. *The British Journal of Sociology*, 1(3):183–196, 1950.
- [2] Samuel Wang. Three tests for practical evaluation of partisan gerrymandering. *Stanford Law Review*, 68(6):1263.
- [3] Samuel S.-H. Wang. Three practical tests for gerrymandering: Application to maryland and wisconsin. *Election Law Journal: Rules, Politics, and Policy*, 15(4):367–384, October 2016.
- [4] Michael D. McDonald and Robin E. Best. Unfair partisan gerrymanders in politics and law: A diagnostic applied to six cases. *Election Law Journal: Rules, Politics, and Policy*, 14(4):312–330, November 2015.
- [5] Nicholas Stephanopoulos and Eric McGhee. Partisan gerrymandering and the efficiency gap. SSRN Scholarly Paper ID 2457468, Social Science Research Network, Rochester, NY, October 2014.
- [6] Robert Browning and Gary King. Seats, votes, and gerrymandering: Measuring bias and representation in legislative redistricting. *Law and Policy*, 9:305–322, 1987.
- [7] Bernard Grofman and Gary King. The future of partisan symmetry as a judicial test for partisan gerrymandering after LULAC v. perry. *Election Law Journal*, 6(1):2–35, 2008.
- [8] Andrew Gelman and Gary King. A unified method of evaluating electoral systems and redistricting plans. *American Journal of Political Science*, 38:514–554, 1994.
- [9] Andrew Gelman and Gary King. Enhancing democracy through legislative redistricting. *American Political Science Review*, 88:541–559, 1994.
- [10] John F. Nagle. Measures of partisan bias for legislating fair elections. *Election Law Journal: Rules, Politics, and Policy*, 14(4):346–360, November 2015.
- [11] Paul Cabilio and Joe Masaro. A simple test of symmetry about an unknown median. *Canadian Journal of Statistics*, 24(3):349–361, September 1996.
- [12] Tian Zheng and Joseph L. Gastwirth. On bootstrap tests of symmetry about an unknown median. *Journal of data science : JDS*, 8(3):413–427, July 2010.
- [13] Yulia R. Gel, Weiwen Miao, and Joseph L. Gastwirth. The effect of dependence between observations on the proper interpretation of statistical evidence. *Law, Probability and Risk*, 8(1):25–38, March 2009.
- [14] Jowei Chen and Jonathan Rodden. Cutting through the thicket: Redistricting simulations and the detection of partisan gerrymanders. *Election Law Journal: Rules, Politics, and Policy*, 14(4):331–345, December 2015.
- [15] Jowei Chen and David Cottrell. Evaluating partisan gains from congressional gerrymandering: Using computer simulations to estimate the effect of gerrymandering in the u.s. house. *Electoral Studies*, 44:329–340, December 2016.
- [16] Dirk P. Kroese, Tim Brereton, Thomas Taimre, and Zdravko I. Botev. Why the monte carlo method is so important today. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(6):386–392, November 2014.