

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

Kevin Baas and Colin McAuliffe

April 12, 2018

## Abstract

There are several legal and technical challenges to the establishment of a standard for limiting partisan gerrymandering, and a few methods have been proposed thus far. All methods for examining gerrymandering use observations of the results of one or more elections to make inferences about the tendency of a given redistricting plan to lead to biased outcomes. Here we propose an empirical Bayesian framework which allows an analyst to make such inferences accurately by using all available election results for the redistricting plan in question in a robust and consistent statistical model. The framework can be used with any gerrymandering metric.

Additionally, we propose a new measure of gerrymandering called the specific asymmetry which we believe will stand up better to judicial and technical tests than any other measure proposed thus far. The specific asymmetry does not rely on proportionality of seats and votes, is applicable to any level of statewide partisanship, does not require national results as a baseline, and measures bias at the popular vote that actually occurred as opposed to some other hypothetical popular vote. All other available metrics fall short in at least one of these aspects, which leaves them vulnerable to criticism and manipulation by those seeking to gerrymander or to defend an existing gerrymander.

We analyze the magnitude and persistence of partisan bias under a particular districting plan by using the empirical Bayesian model to compute the expected value of the specific asymmetry. 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 Act 43 map for the Wisconsin State Assembly, which is the subject of *Whitford v. Gill*.

Keywords: Redistricting, gerrymander, Whitford v. Gill,

## 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 issue of partisan gerrymandering, and recent rulings suggest that a manageable standard for measuring gerrymandering is a prerequisite for a such a ruling (*LULAC v. Perry* 2004).

The establishment of some standards to reel in 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 ?, and the Supreme Court has stated that proportionality is not acceptable (Thornburg v. Gingles 1986). A second example is the mean median difference test ???, 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. ?

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. Gill used a metric called the efficiency gap ?, 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 ??, which was used in a Bayesian model for evaluating redistricting plans in ??. Nagle has proposed several other metrics based on the seats votes curve ?, 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 ??. 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. However, 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. ?? Lastly, 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 ?. 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 without sophisticated analysis. 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, 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 ??, 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 ??, 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 integration. The metric, which we refer to as the specific asymmetry (see section ??), 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 ??, uses a Monte Carlo technique based on draws from 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 take a lot more computer processing power, and thus isn't as well suited for use in optimization algorithms, relative to analytic methods. Nonetheless, Monte Carlo techniques are highly conventional tools that are in widespread use across numerous disciplines ?, 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 ??, 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 ?? 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. Gill</i>
Gelman et al	Deficit in Seats at 50% of the vote, various others	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 ? 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 ?, 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 ???. However, these measures of bias tend to apply only to states which are close to 50-50 partisanship. 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 ?. 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.

This could be accomplished with a modification to the approach proposed by Nagle where one computes the entire area between the seats votes curve and its reflection, but weights this by likelihood as follows

$$\int_0^1 \mathcal{L}(V) [S(V) - S(1 - V)] dV \quad (1)$$

Where  $\mathcal{L}(V)$  is a probability density function for statewide popular vote, which can be estimated empirically. The likelihood weighting means that asymmetries at vote levels that are unlikely to occur in a given state do not contribute much to the integral, while the vote levels that are most likely contribute the most. However, this integral form of the asymmetry requires data from multiple election years to estimate  $\mathcal{L}(V)$ , and we prefer to use a metric which can be calculated from a single election year in a completely self contained and deterministic fashion.

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, i.e.  $[S(V_{pop}) - S(1 - V_{pop})] / 2$ . This definition of partisan asymmetry simply requires the seat count to reverse when the state wide popular vote reverses. This has a certain intuitive appeal and is consistent with existing law since disproportionate results are allowed, as long as both parties are treated the same.

However, one may argue that a reversal of the popular vote is an unrealistic counterfactual scenario that is beyond the scope of any sensible analysis technique. This is true in the sense that we could not predict what the distribution of votes would be if such a partisan reversal would actually occur, but it is important to note that when we compute the discrepancy in seat counts under a uniform reversal in statewide partisan ship, we are computing the magnitude of structural disadvantage faced by the voters of one party. In other words, while the specific asymmetry may be phrased as a sort of hypothetical, it measures a *real* electoral disadvantage endured by one party resulting from that party winning seats less efficiently than the opposite party due to voter packing and cracking. Further, as the magnitude of

the structural disadvantage to one party increases due to a more aggressive gerrymander, the specific asymmetry increases as well.

The metric can be calculated deterministically from the results of a single election, and when combined with the sampling method in section ??, a quantity representing the distribution of likely asymmetries analogous to equation ?? can be computed. Sampling the specific asymmetry to compute likely asymmetries is preferable to using the integral in equation ??, since the specific asymmetry has units of seats. This is an intuitive unit of measure which makes it easier to communicate results with legal experts as well as the general public. Using seats as a unit of asymmetry also makes it easier to aggregate results across states to obtain the net effects of gerrymandering nationally.

There are several other advantages to using this metric, which we summarize below.

- It doesn't assume proportionality in results - The seats-votes curve for single-winner elections naturally takes 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). However 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 their fraction 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.
- 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.

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. Gill* 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 biased or unbiased *on average*. Full details of the particular model employed here are presented in the following section.

# Acknowledgment

## References

- 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.
- Samuel Wang. Three Tests for Practical Evaluation of Partisan Gerrymandering. *Stanford Law Review*, 68(6):1263.
- 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.
- 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.
- Nicholas Stephanopoulos and Eric McGhee. Partisan Gerrymandering and the Efficiency Gap. SSRN Scholarly Paper ID 2457468, Social Science Research Network, Rochester, NY, October 2014.
- Robert Browning and Gary King. Seats, Votes, and Gerrymandering: Measuring Bias and Representation in Legislative Redistricting. *Law and Policy*, 9:305–322, 1987.
- 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.
- Andrew Gelman and Gary King. A Unified Method of Evaluating Electoral Systems and Redistricting Plans. *American Journal of Political Science*, 38:514–554, 1994.
- Andrew Gelman and Gary King. Enhancing Democracy Through Legislative Redistricting. *American Political Science Review*, 88:541–559, 1994.
- John F. Nagle. Measures of Partisan Bias for Legislating Fair Elections. *Election Law Journal: Rules, Politics, and Policy*, 14(4):346–360, November 2015.
- Paul Cabilio and Joe Masaro. A simple test of symmetry about an unknown median. *Canadian Journal of Statistics*, 24(3):349–361, September 1996.
- 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.
- 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.
- 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.
- 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.
- 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.
- Philip Paolino. Maximum Likelihood Estimation of Models with Beta-Distributed Dependent Variables. *Political Analysis*, 9(4):325–346, 2001.

- Edward H. Kaplan and Arnold Barnett. A New Approach to Estimating the Probability of Winning the Presidency. *Operations Research*, 51(1):32–40, February 2003.
- Andreas Erwin Murr. The Party Leadership Model: An early forecast of the 2015 British General Election. *Research & Politics*, 2(2):2053168015583346, June 2015.
- Steven E. Rigdon, Sheldon H. Jacobson, Wendy K. Tam Cho, Edward C. Sewell, and Christopher J. Rigdon. A Bayesian Prediction Model for the U.S. Presidential Election. *American Politics Research*, 37(4):700–724, July 2009.
- Brian Remliger and Samuel Wang. Personal Communication. May 2017.
- Samuel Wang and Brian Remlinger. Can math stop partisan gerrymandering? *Los Angeles Times*, May 2017.
- Jowei Chen. The Impact of Political Geography on Wisconsin Redistricting: An Analysis of Wisconsin’s Act 43 Assembly Districting Plan. *Election Law Journal: Rules, Politics, and Policy*, 2017.