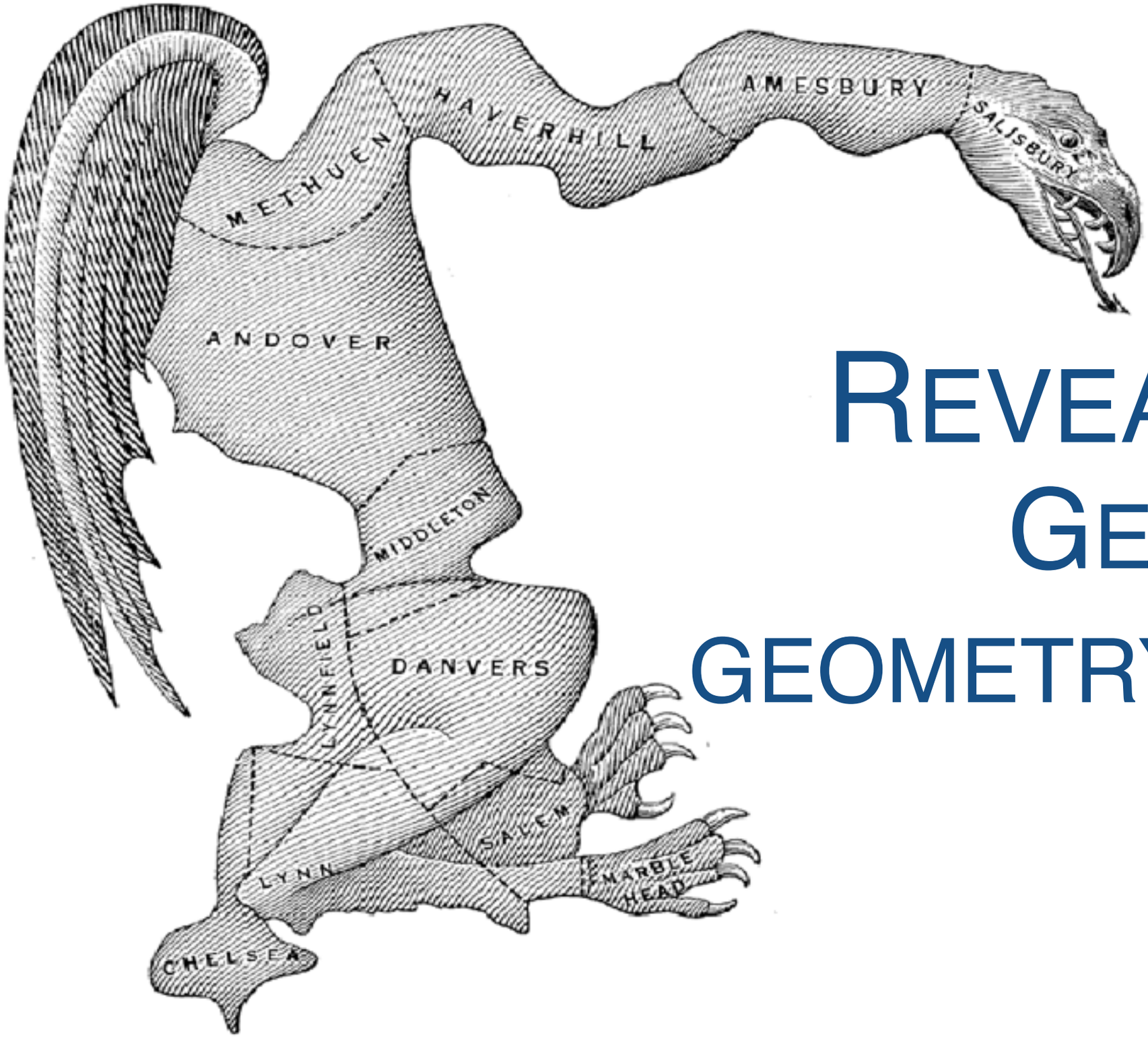
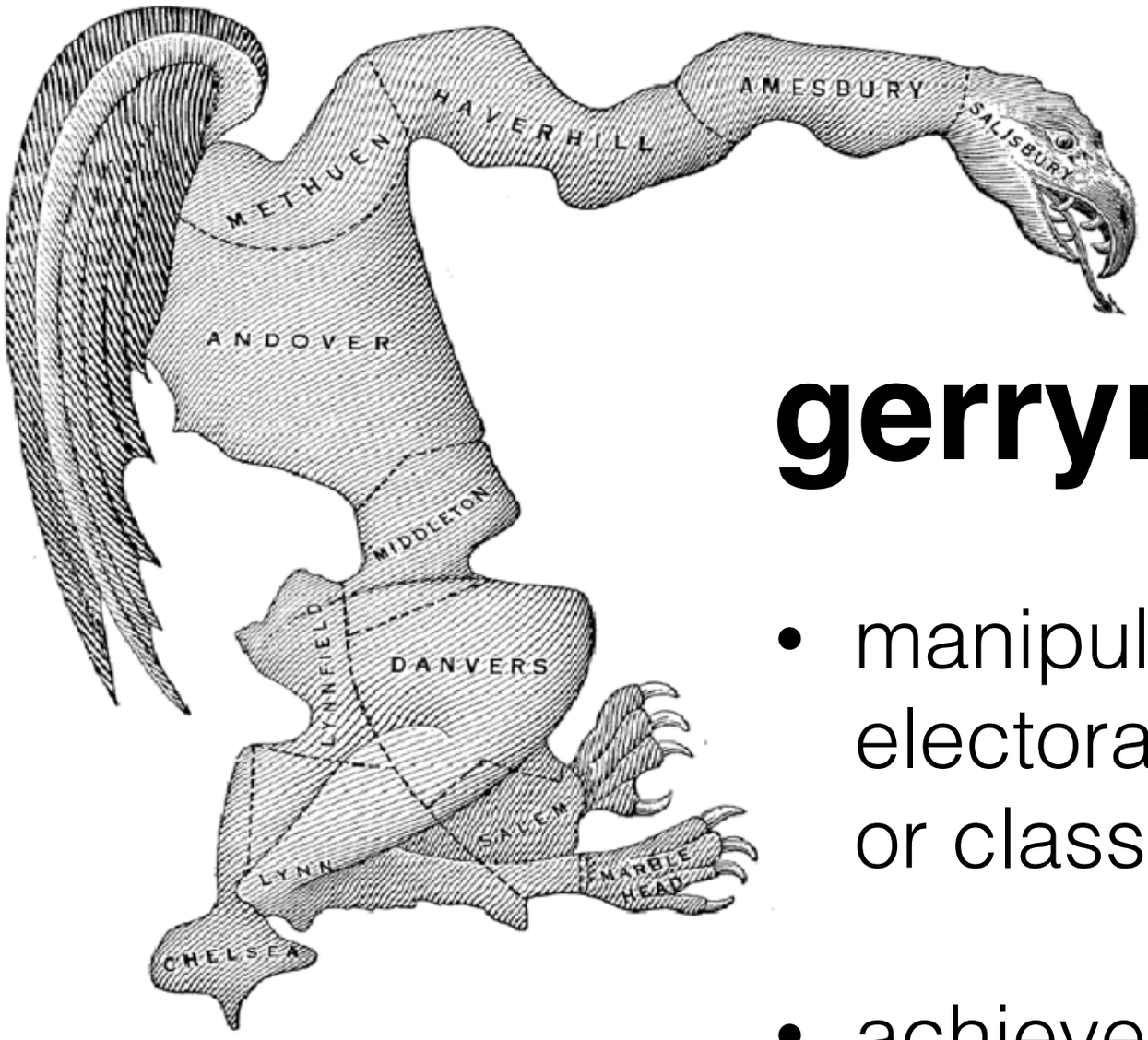


REVEALING THE GEOPOLITICAL GEOMETRY THROUGH SAMPLING

JONATHAN MATTINGLY (+ THE TEAM)
DUKE MATH



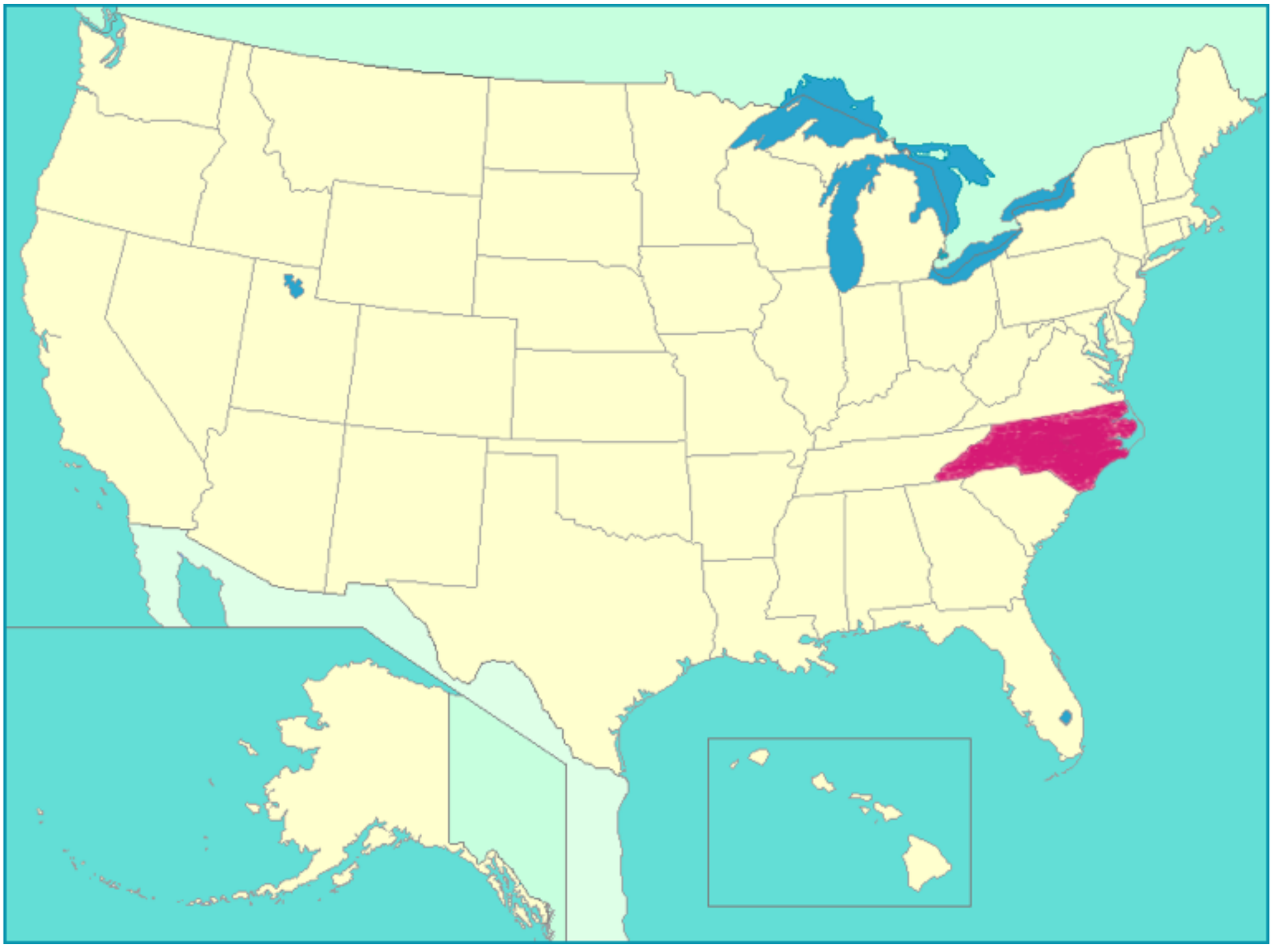
gerrymander

- manipulate the boundaries of an electoral constituency to favor one party or class.
- achieve (a result) by manipulating the boundaries of an electoral constituency.

"*a total freedom to gerrymander the results they want*"

***racial* vs *partisan* gerrymander**

North Carolina

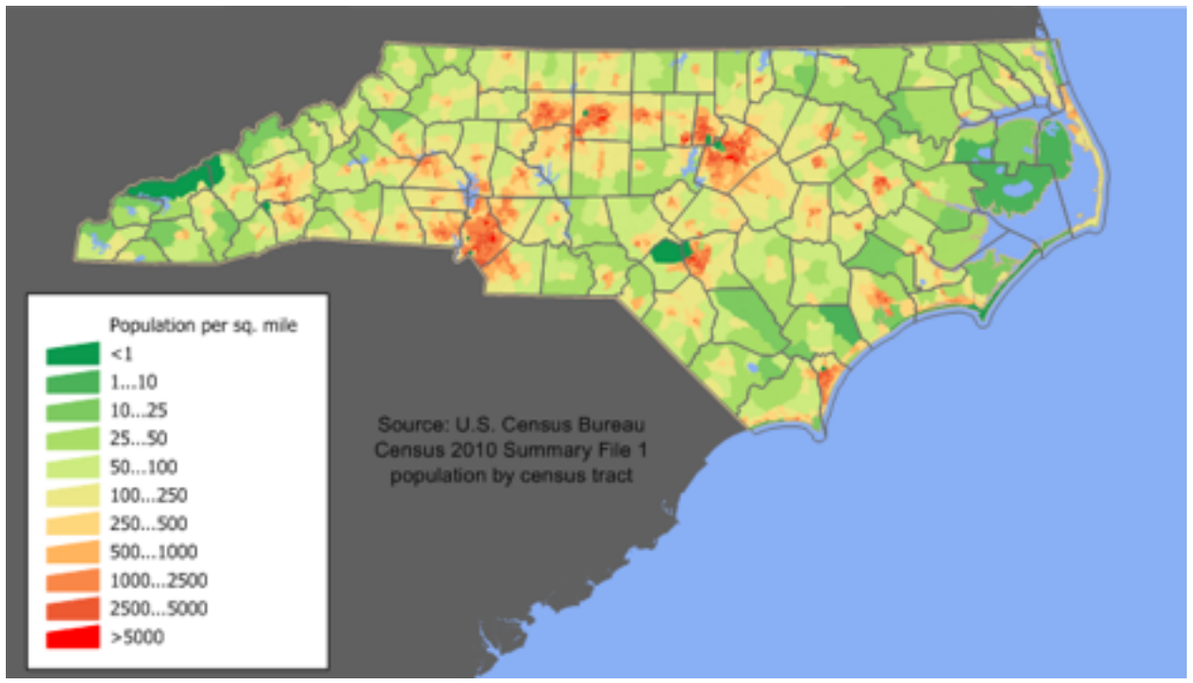


13 Congressional Representatives

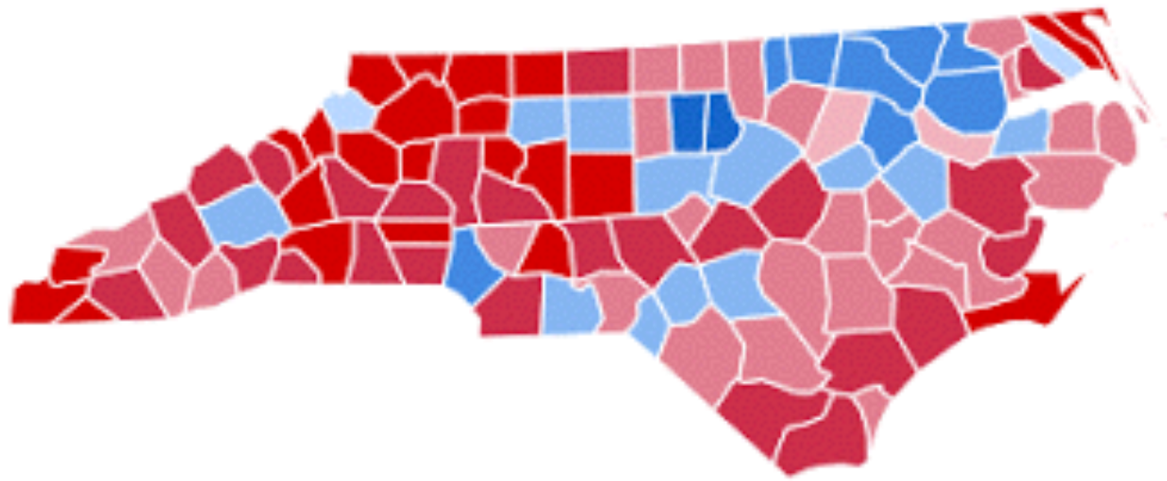
NC has around 10.2 million people

Every decade, required to redo the (13) congressional districts

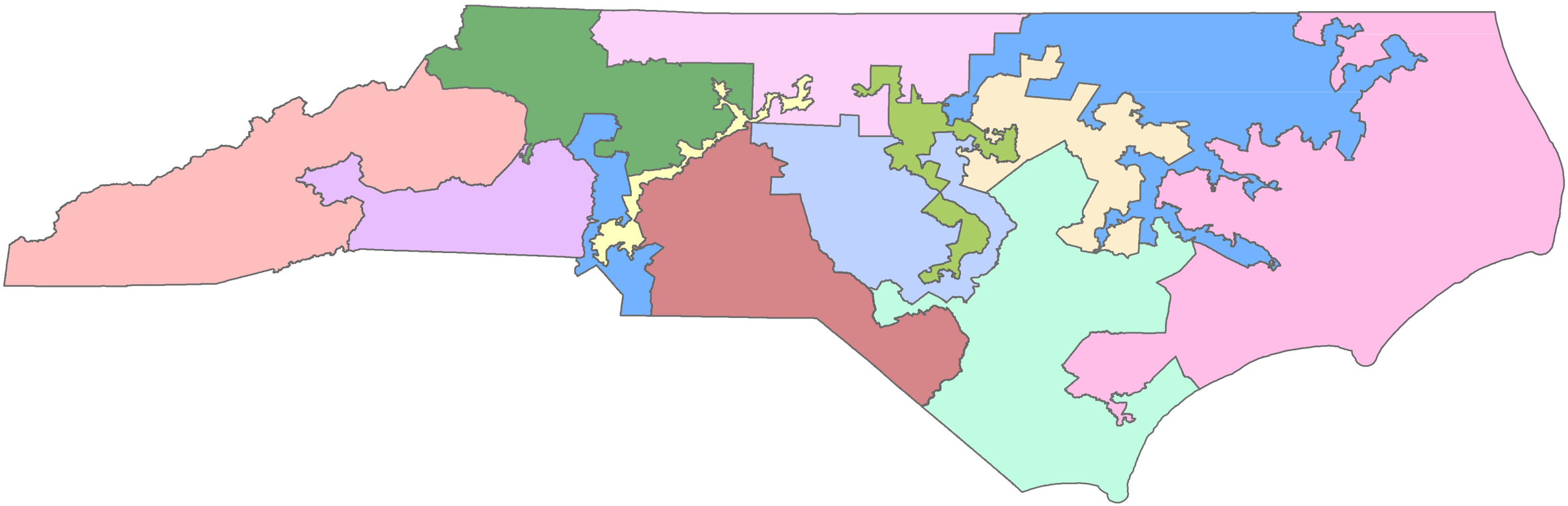
Population Density



Presidential Election 2016



Charlotte Area: *Charlotte-Gastonia-Salisbury*- population 2,402,623
The Triangle: Raleigh-Durham-Cary-Chapel Hill- population 1,749,525
The Piedmont Triad: *Greensboro—Winston-Salem—High Point*- population 1,589.200



U.S. House congressional districts for 2012 election

DEMOCRACY

DEMOCRACY

The will of the people is expressed

every person can

VOTE

every vote is

COUNTED

(once)

one person one vote principle

"By ensuring that **each representative** is subject to requests and suggestions from the **same number of constituents**, total population apportionment promotes equitable and effective representation."

–Justice Ruth Bader Ginsburg
Evenwel v. Abbott, April 2016

2012 North Carolina Elections for U.S. House

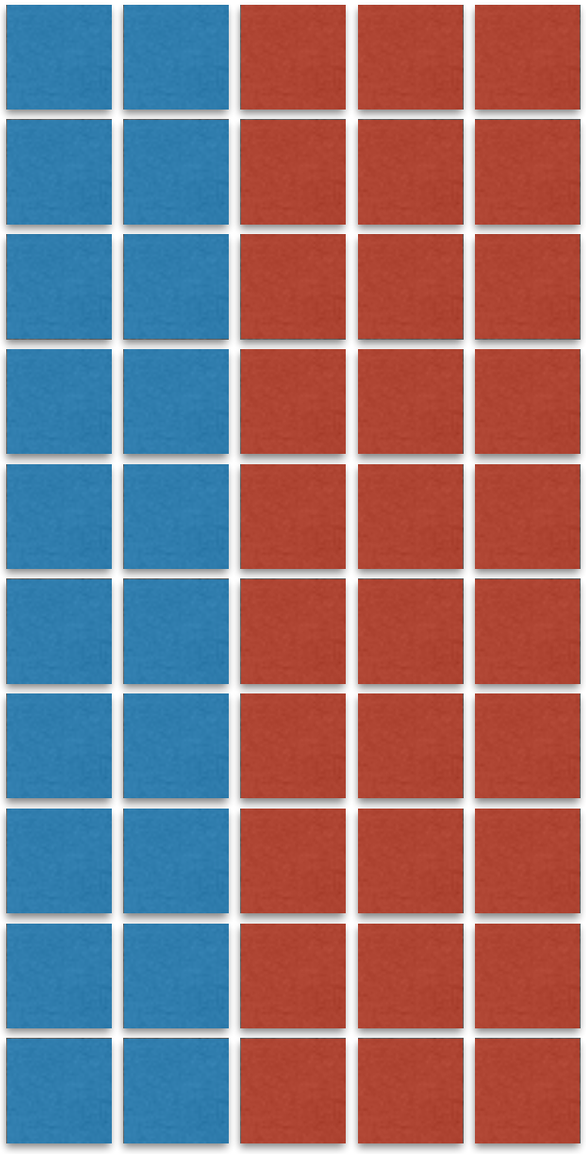
	VOTES	PERCENTAGE	SEATS
Democratic	2,218,357	50.65%	4
Republican	2,137,167	48.80%	9
Libertarian	24,142	0.55%	0

-The most Democratic district had 79.63% Democratic votes.
-The most Republican district had 63.11% Republican votes.

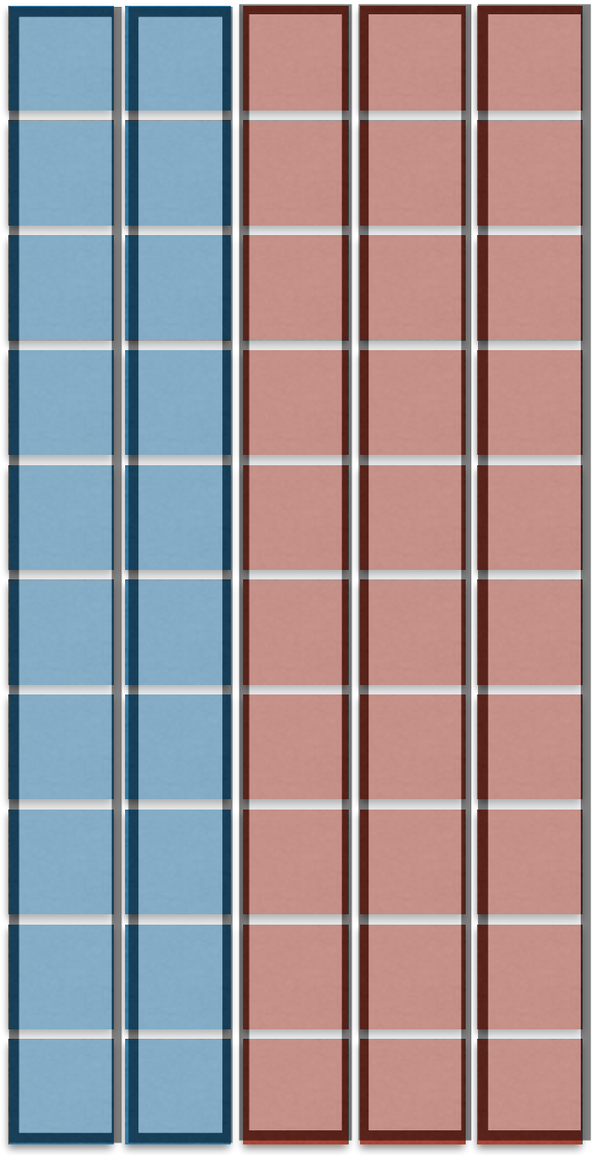
Are these results due to political gerrymandering ?

or

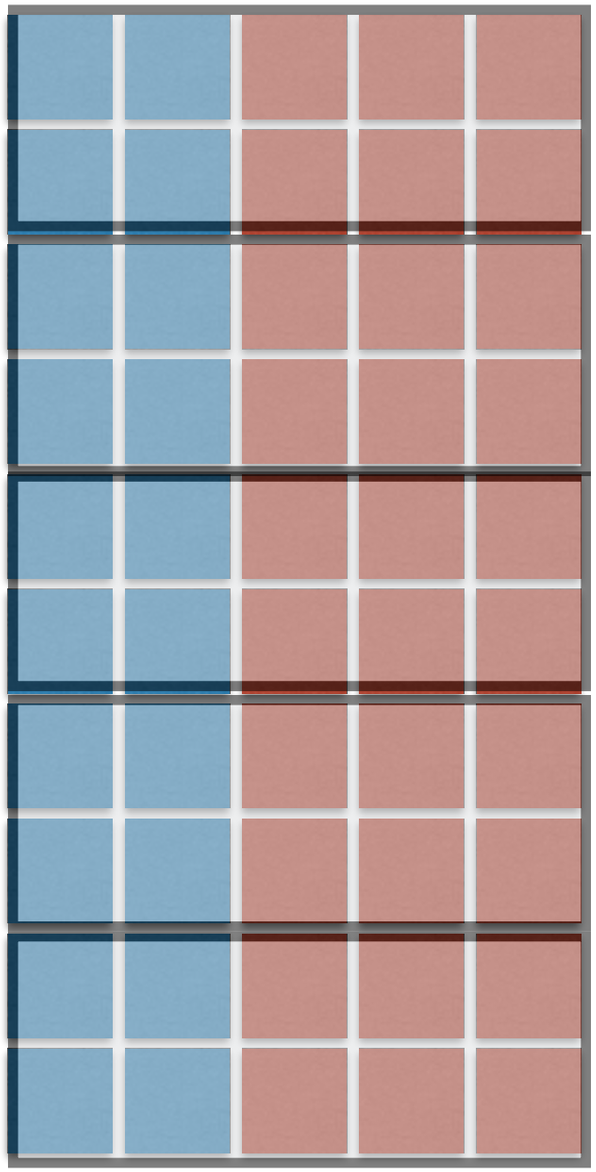
Are these results natural outcomes of NC's geopolitical structure of the spatial distribution of partisan votes ?



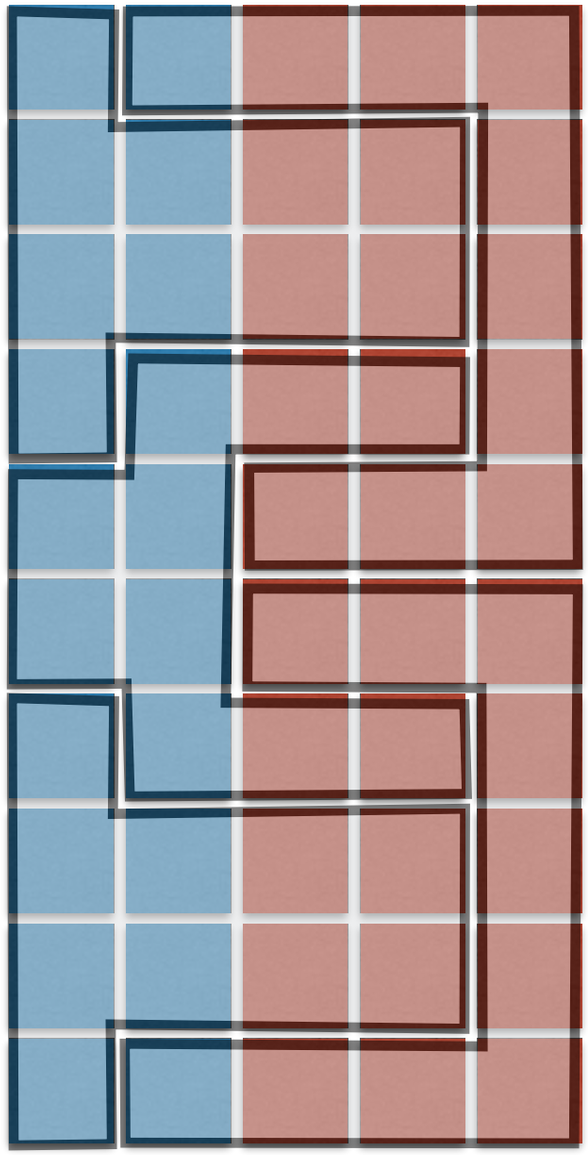
40% Blue
60% Red



Red wins 3
Blue wins 2

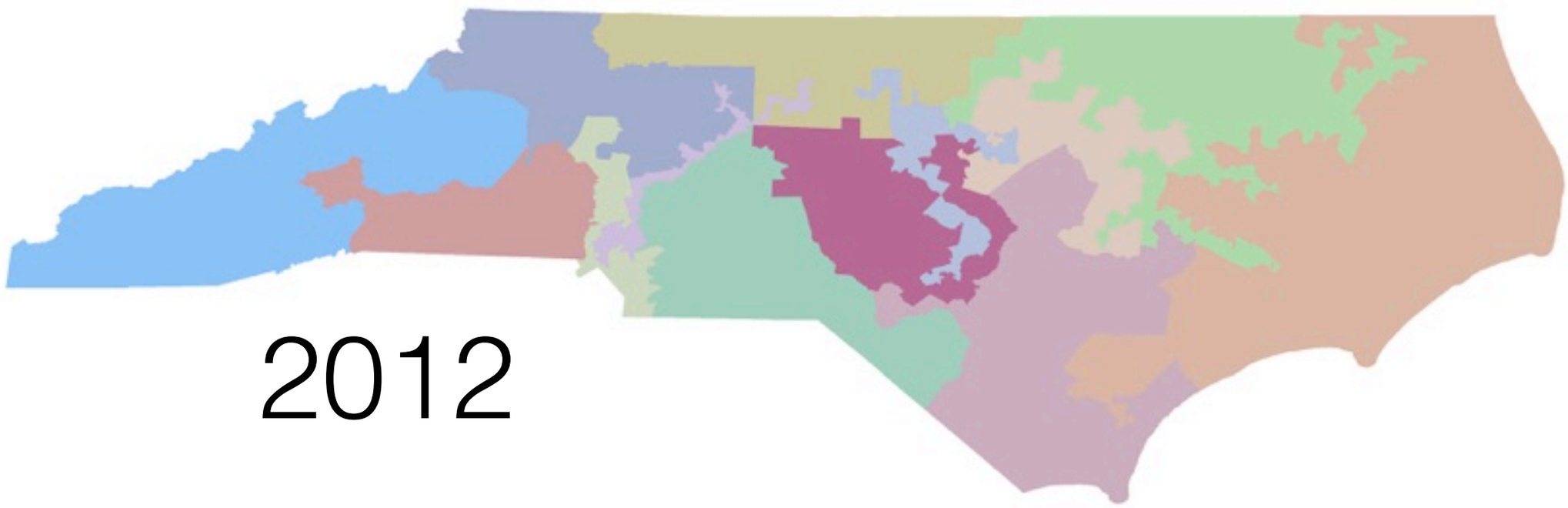


Red wins 5
Blue wins 0

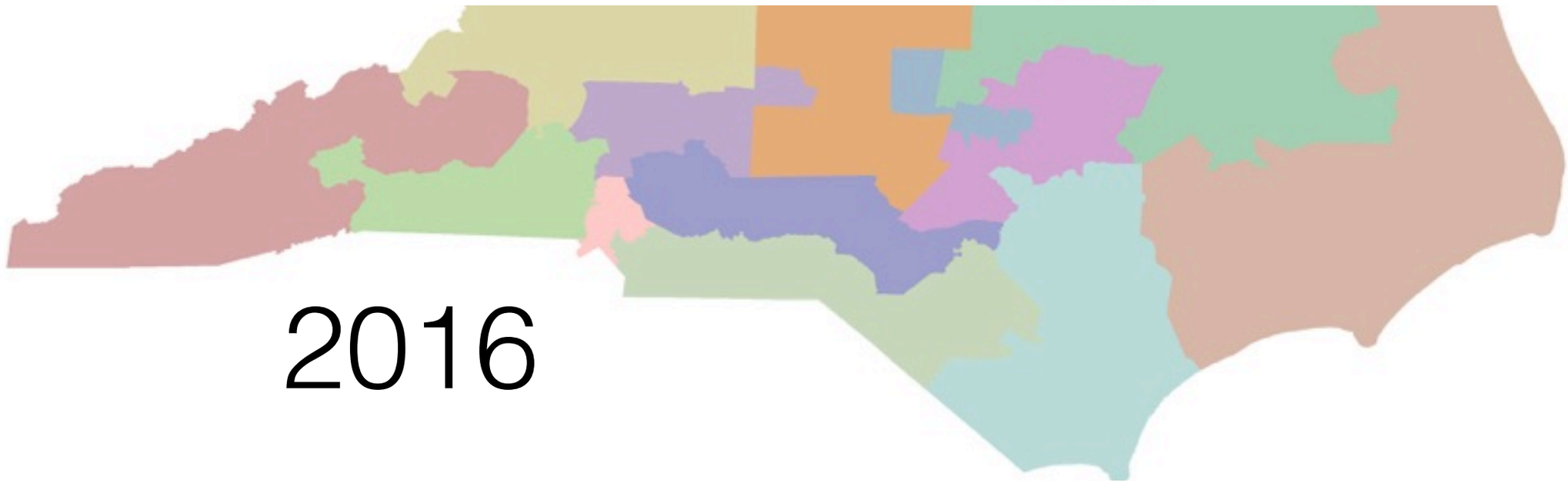


Red wins 2
Blue wins 3

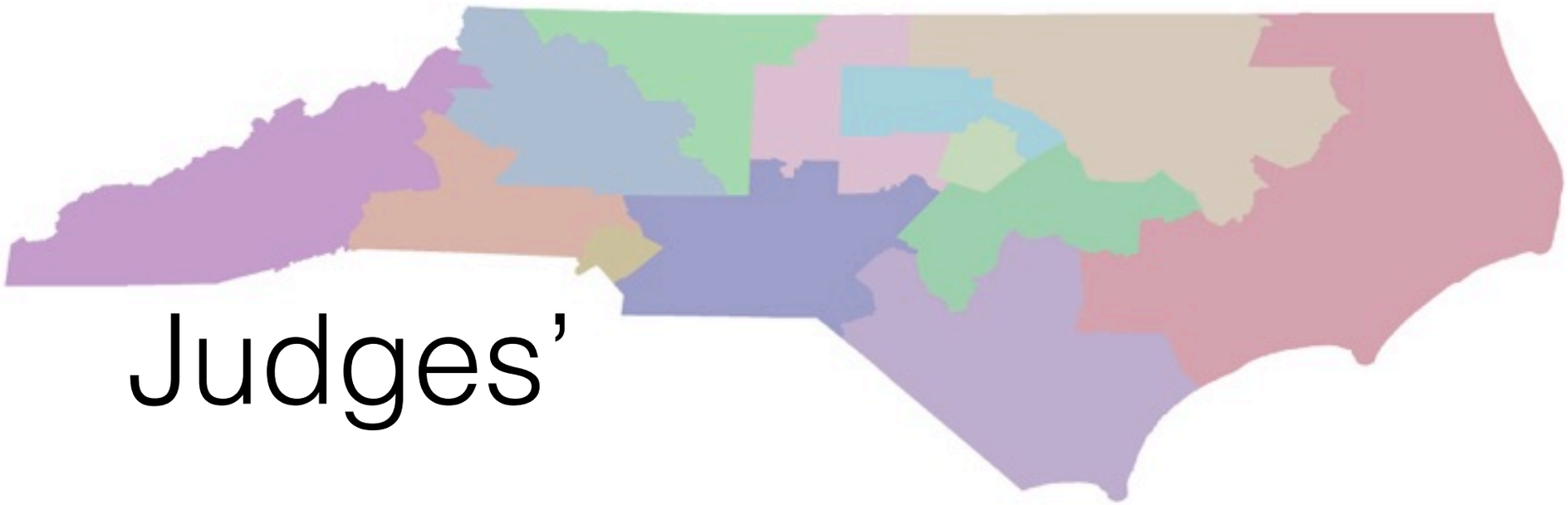
from Wikipedia after an image by Steven Nass



2012



2016



Judges’

How to quantify gerrymandering?

How to reveal a state's geopolitical structure?

How does one find the true message in an election ?

What if we drew the districts randomly ?

with no regard for party registration or most demographics

reveal the geopolitical structure encoded in the votes

Many Groups using algorithmic generated maps to benchmark

- Jowei Chen (Univ Michigan)
- Wendy Cho (UIUC)
- Samuel Wang (Princeton)
- Kosuke Imai, Benjamin Fifield (Princeton)
- Alan Frieze, Wesley Pegden, Maria Chikina (CMU)

Not all the same. Not all Random.

Some generating alternative maps.
Some Sampling a defined distribution.
some using actual surrogate districts.

Focus on our group at Duke

Impact of Duke Team work

Gill v. Whitford *(WI State Assembly) :*

- Oral argument held in Supreme Court (SCOTUS) October 2
- Provide report supporting Amicus Brief by Eric S. Lander

Common Cause v. Rucho *(N.C. Congressional)*:

- 3 judge conditional panel. Direct appeal to SCOTUS.
- Provide expert testimony and report in lawsuit.
  Closing arguments on October 16

North Carolina v. Covington *(N.C. State Assembly)*:

- 3 judge conditional panel rule racial gerrymander. Affirmed by SCOTUS in June.
- Provide expert testimony on new maps produces at courts order
- Preparing for partisan gerrymander

The Recipe

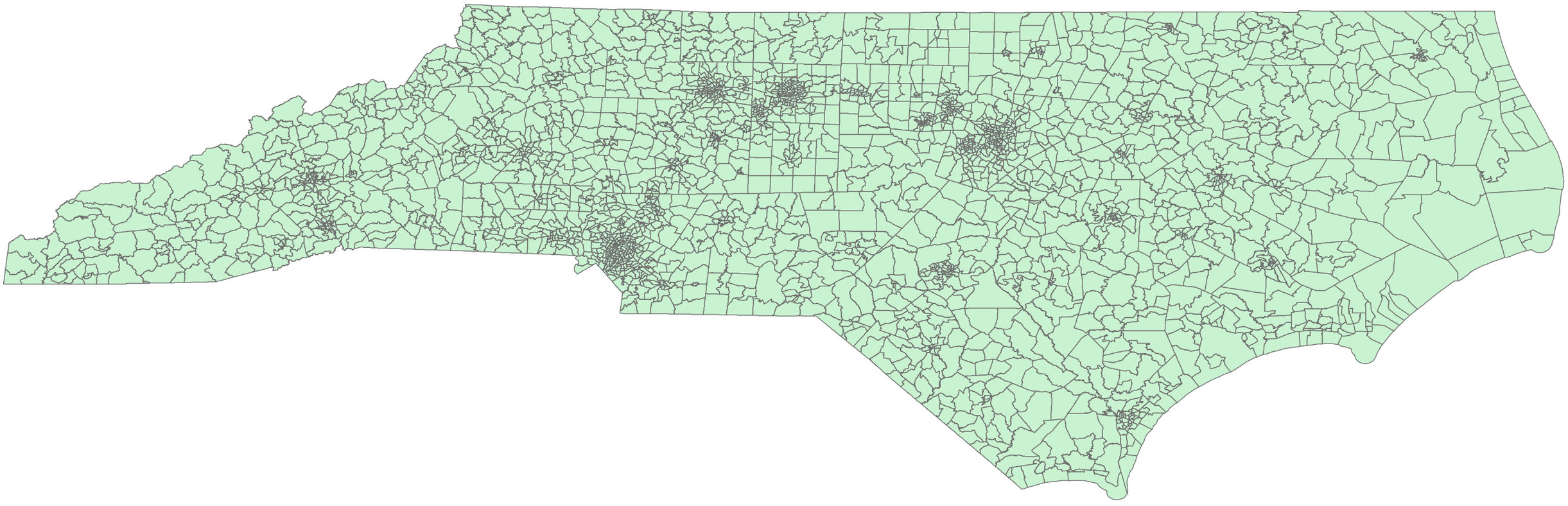
1. Make a good random redistricting of N.C. into 13 U.S. house districts.

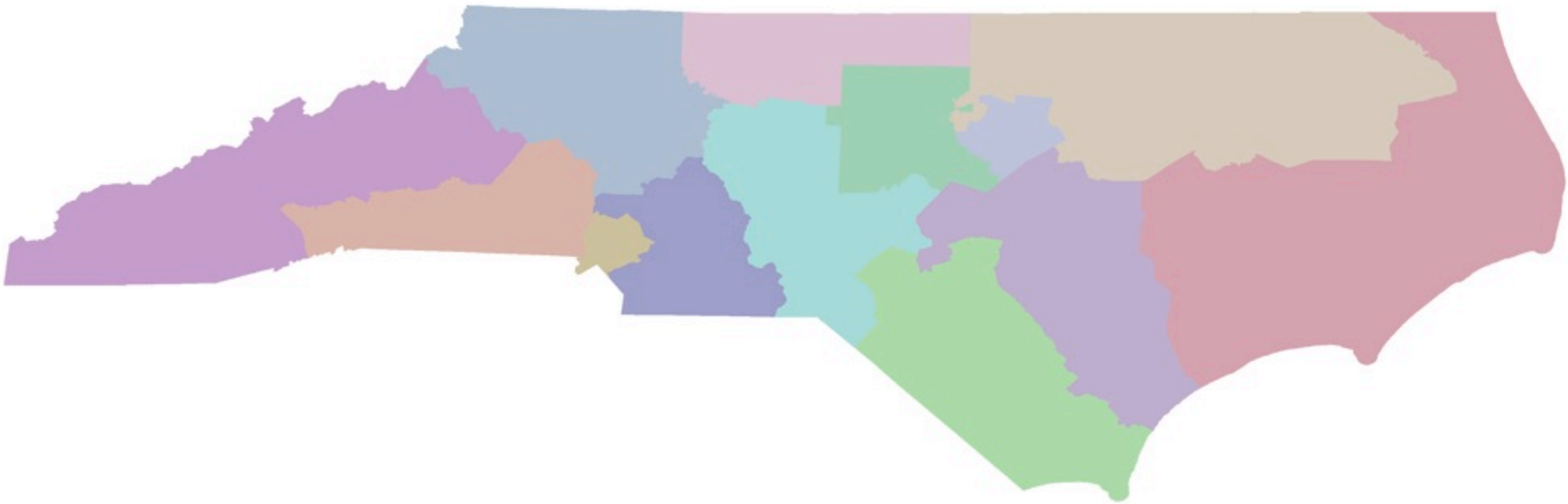
2. Count number of Democratic and Republican votes in each of the new districts using the actual 2012 votes.

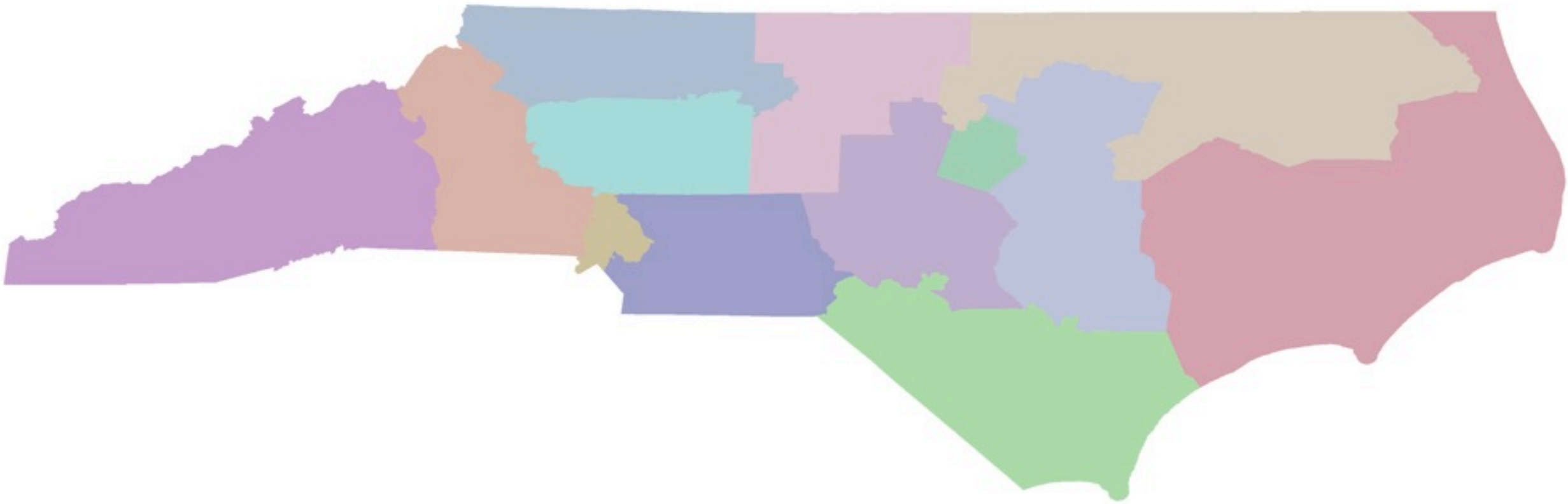
3. Determine winner in each district of the random redistricting.

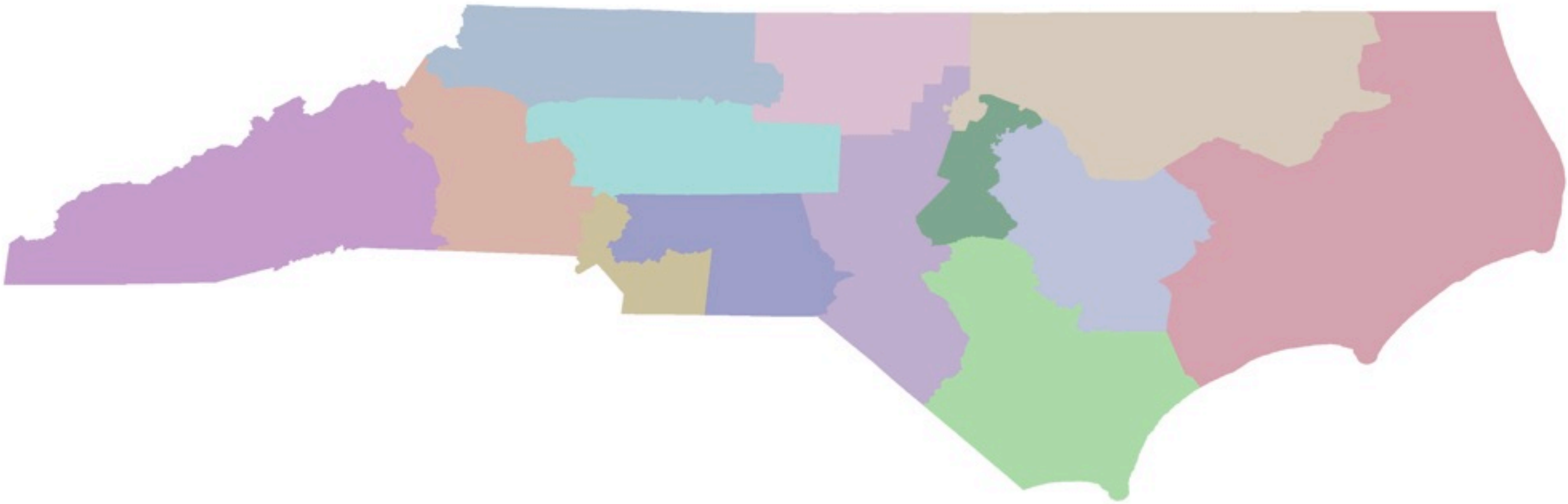
4. Return to step 1.

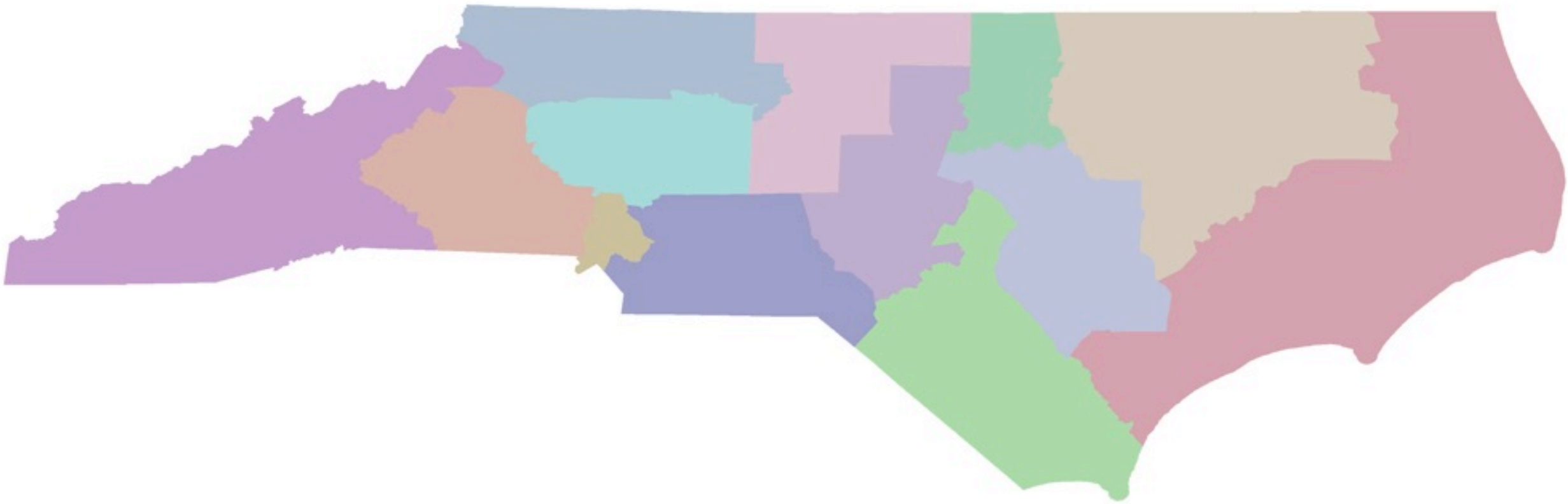
Use Markov Chain Monte Carlo to sample a distribution on redistrictings

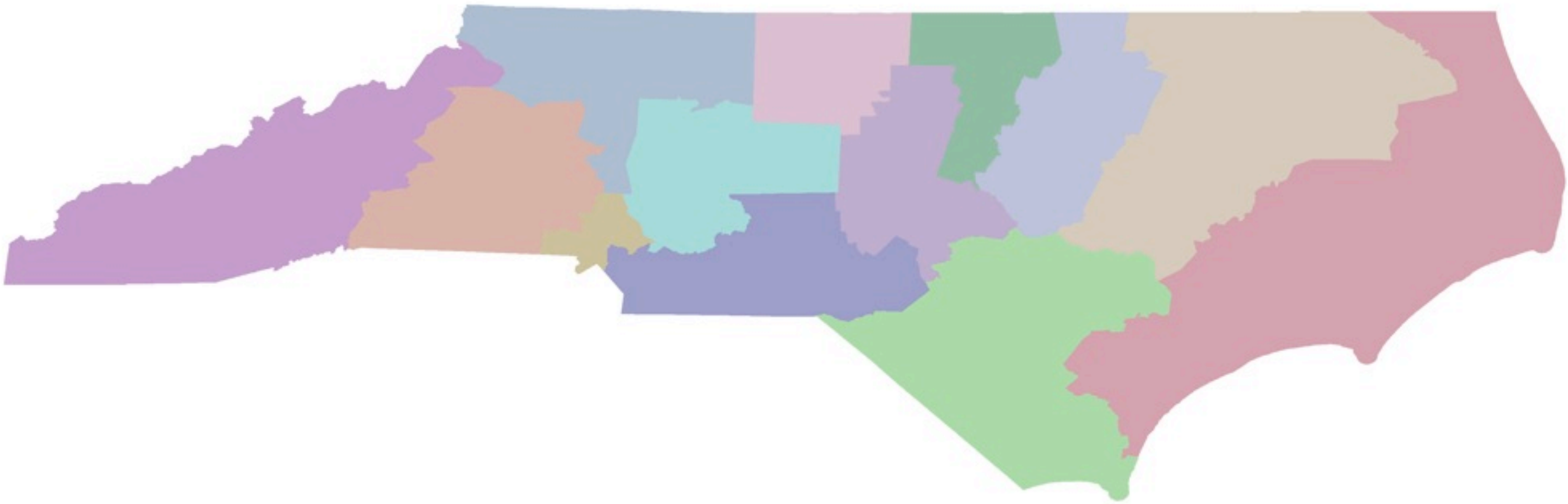


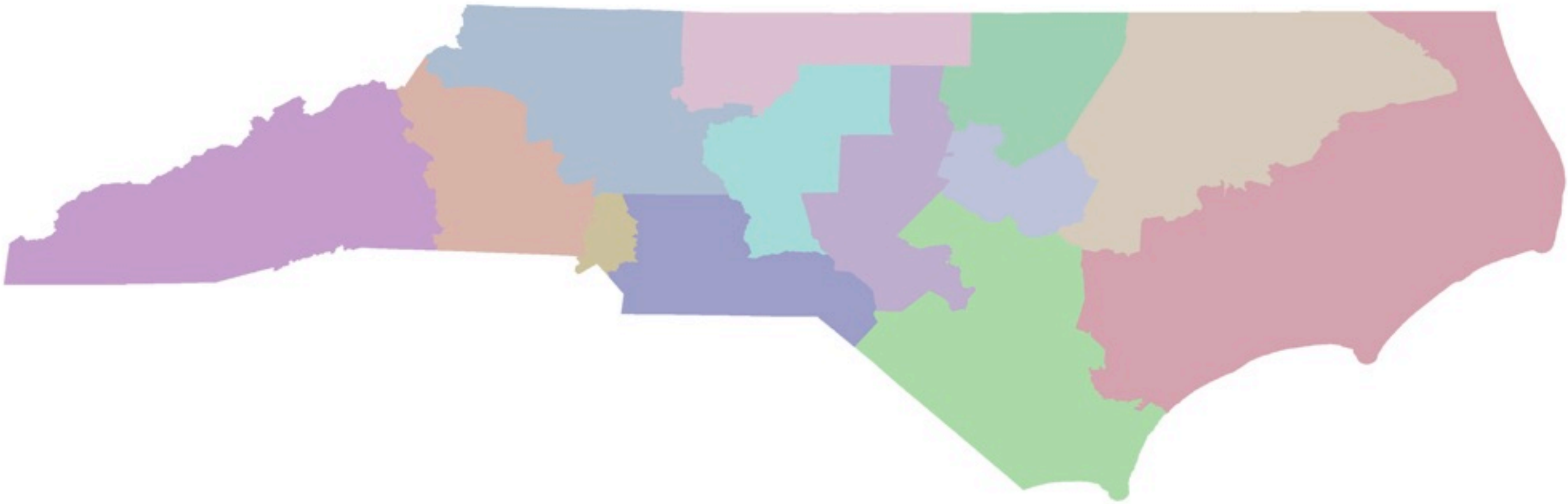


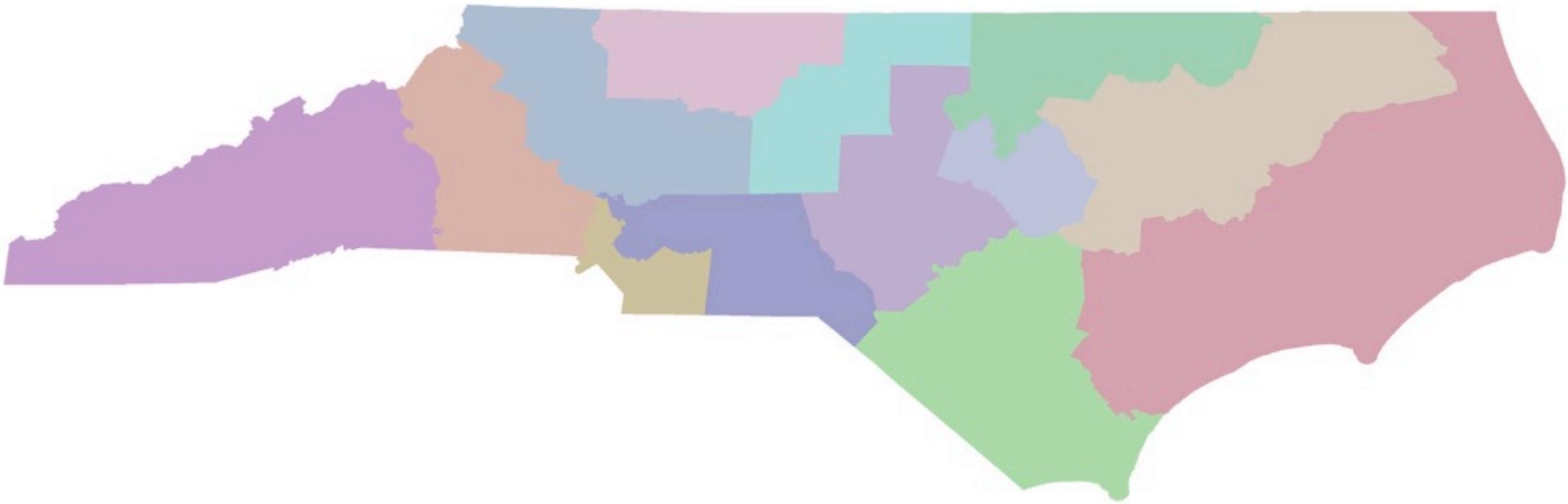












Criteria for Sampling

non-partisan design criteria (HB 92)

1. districts have equal population

2. the districts are connected and compact,

3. splitting counties is minimized, and

4. African American voters are sufficiently concentrated in 2 districts to affect the winner.

Use Markov Chain Monte Carlo to sample from redistricting with good scores.

$$\text{Sample: } (\text{density}) \propto e^{-\beta(\text{score of redistricting})}$$

Know what distribution we are sampling from.

Not just generating a large number of alternatives.

N.C. HOUSE BILL 92 REDISTRICTING STANDARDS

- Districts within 0.1% of equal population
- Districts shall be reasonably compact
- Contiguous territory, attempting not to split cities or counties
- Comply with the Voting Rights Act of 1965
- Ignore: Incumbency, party affiliation, demographics

N.C. HOUSE BILL 92
REDISTRICTING STANDARDS

- Districts within 0.1% of equal population (we get close)
- Districts shall be reasonably compact
- Contiguous territory, attempting not to split cities or counties
- Comply with the Voting Rights Act of 1965
- Ignore: Incumbency, party affiliation, demographics

Score function

$$P(\xi) = \frac{1}{Z} e^{-\beta J(\xi)}$$

$$\xi \colon \{\text{Precincts}\} \mapsto \{1, \ldots, 13\}$$

$$J(\xi) = w_p J_{pop}(\xi) + w_I J_{compact}(\xi) + w_c J_{county}(\xi) + w_m J_{mino}(\xi)$$

(a 13 color Potts Model with an unusual energy)

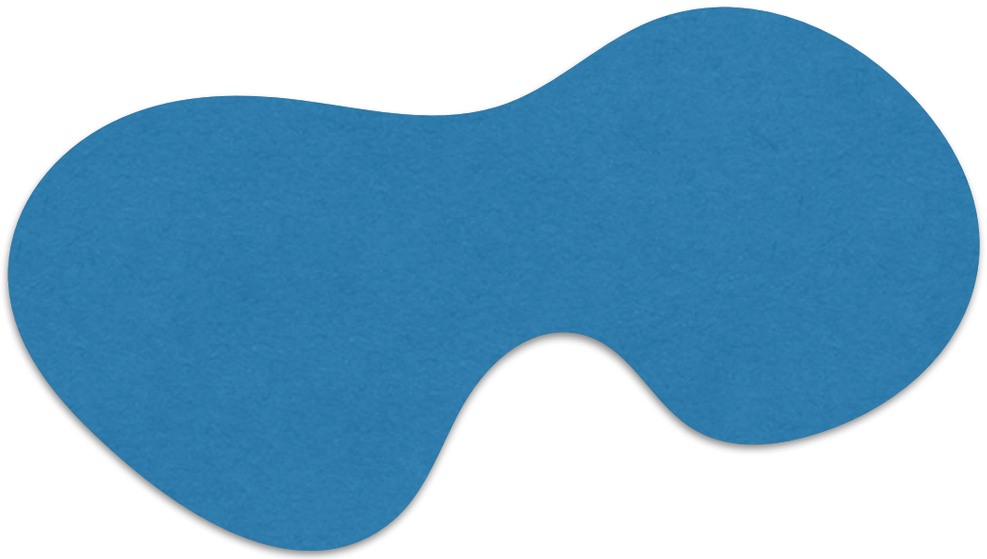
Population Score

Sum of square deviation from ideal district population

$$\sum_{n=1}^{13} \Big[\text{Ideal} - (\text{Pop in district } n)\Big]^2$$

$$\text{Ideal} = \frac{\text{Population of N.C.}}{13} \approx 733,499$$

Compactness score

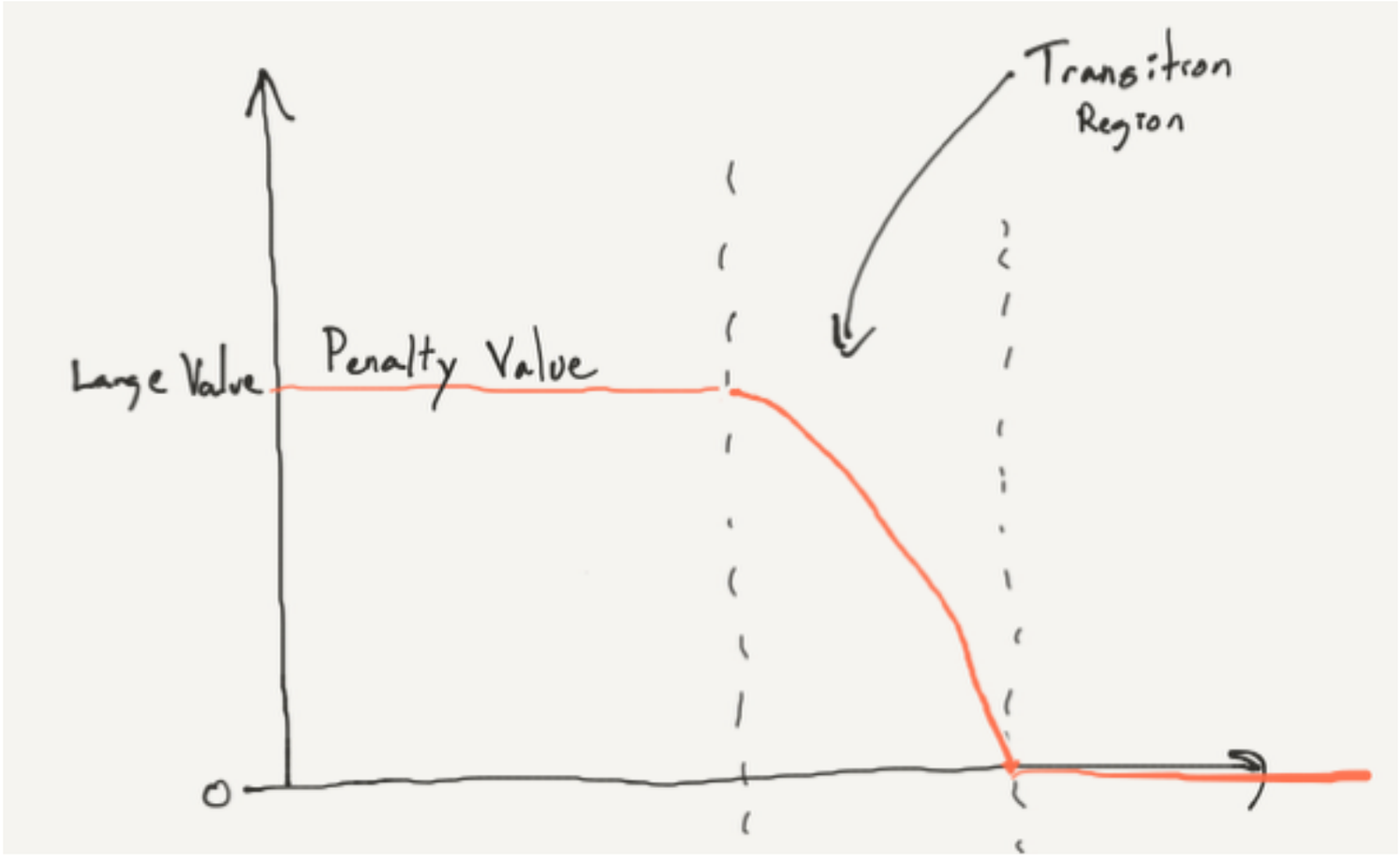


$$\frac{(\text{Perimeter})^2}{\text{Area}} \geq 4\pi \approx 12.5$$

Minimized for a circle

Also considered the ratio of district's area to the smallest circumscribing rectangle

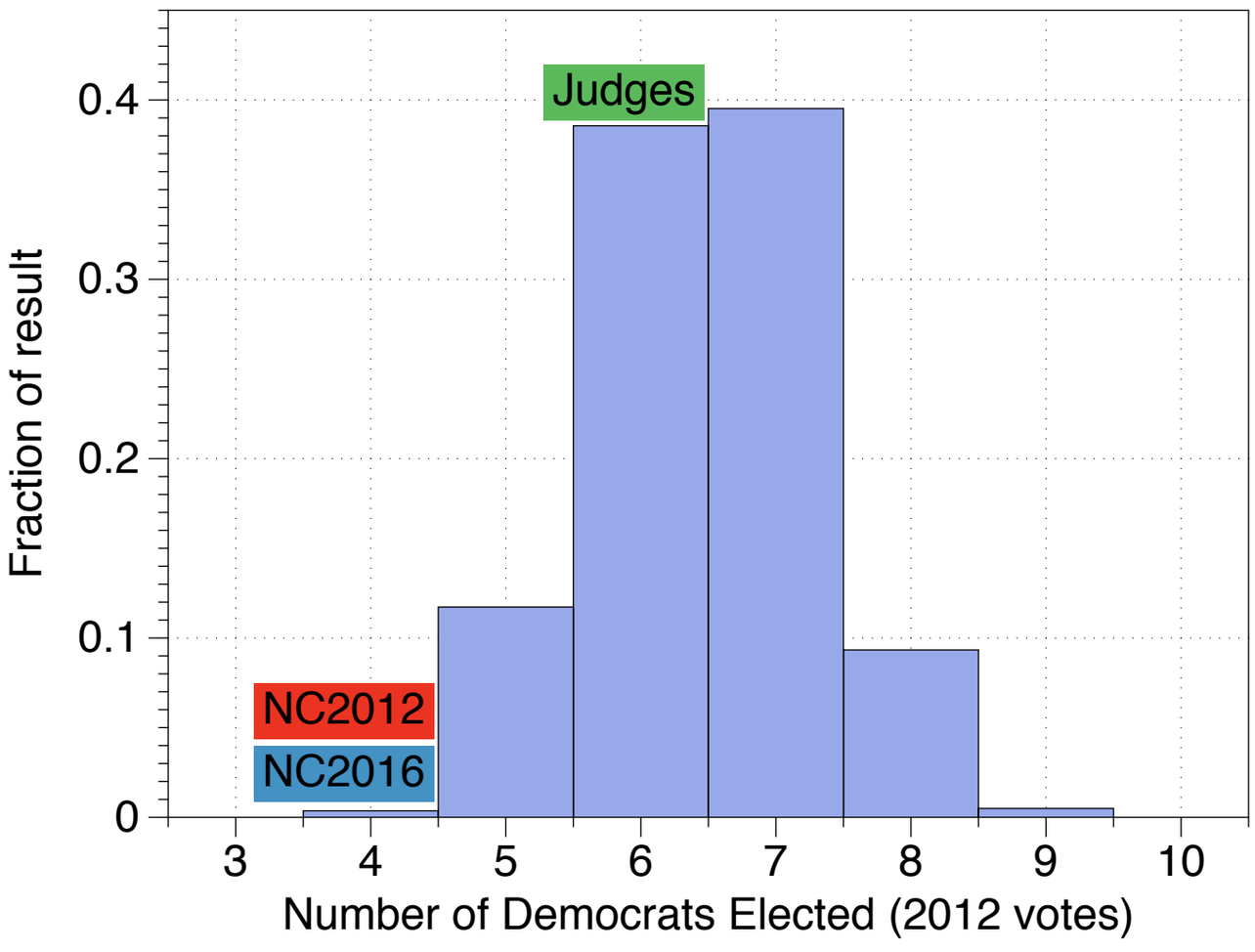
Also include score terms for Voting Rights Act and Preserving County Boundaries

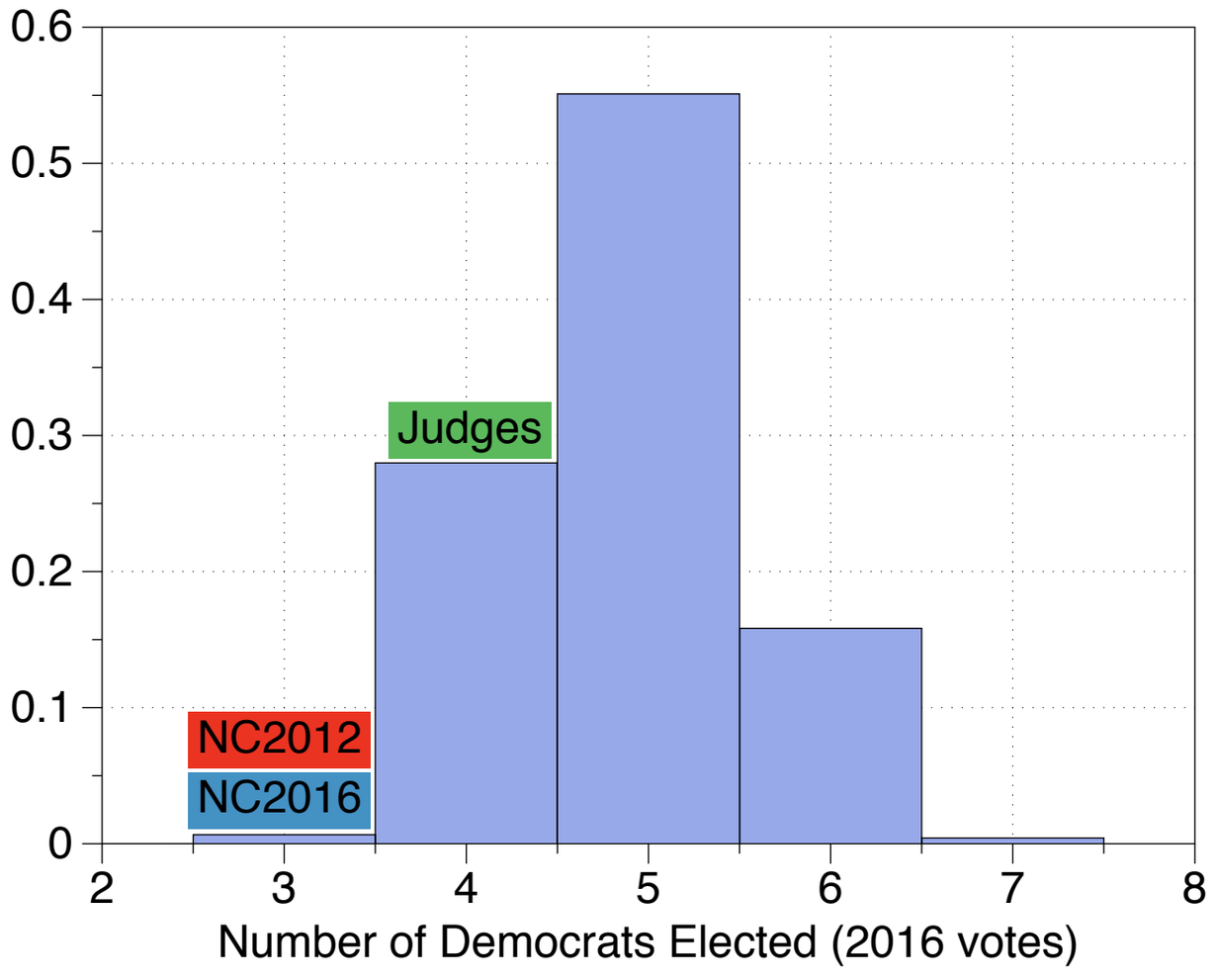


Soft penalization :

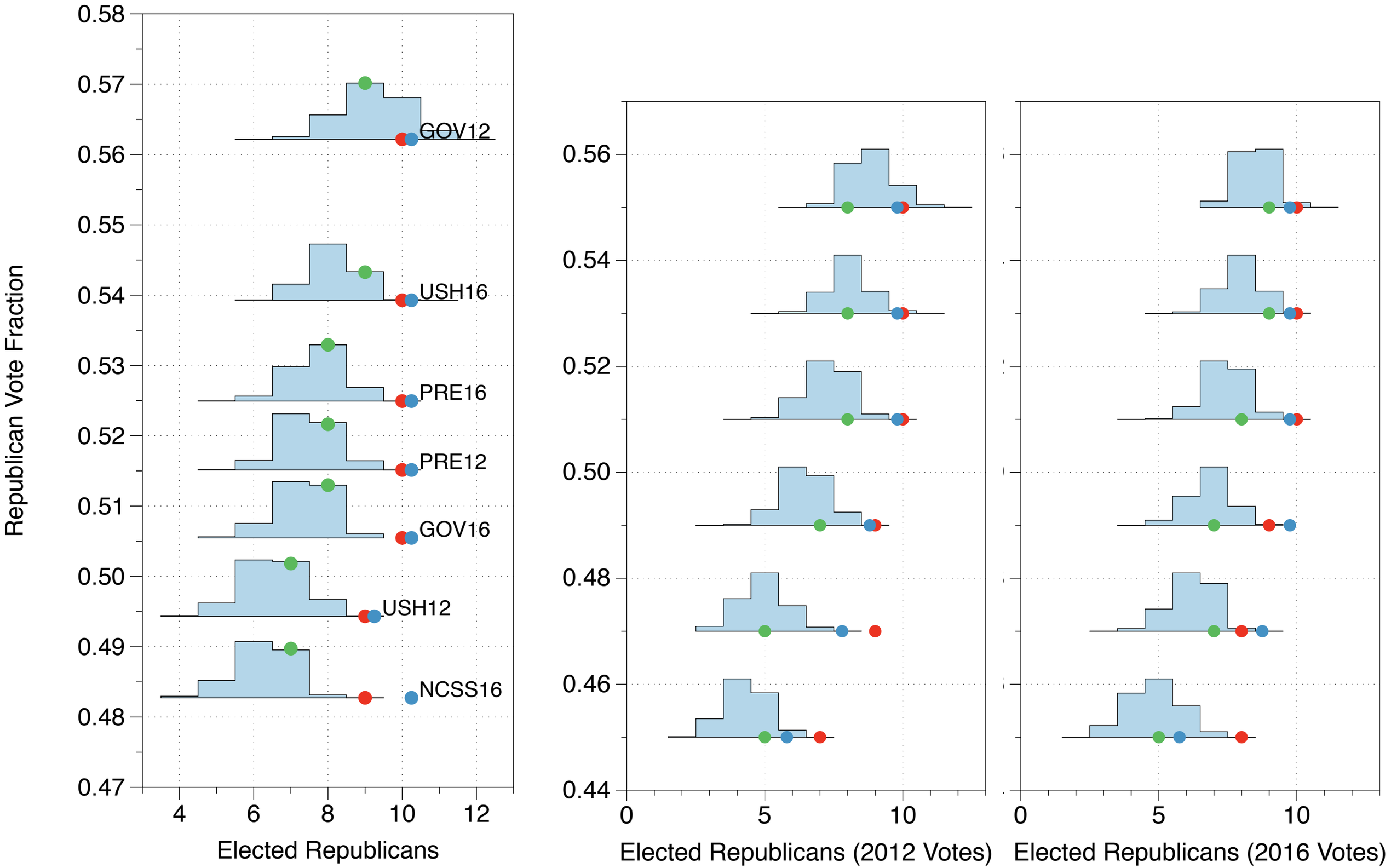
- for number of split counties of different sizes
- redistricting plans without two districts meeting minimal voting age black population.

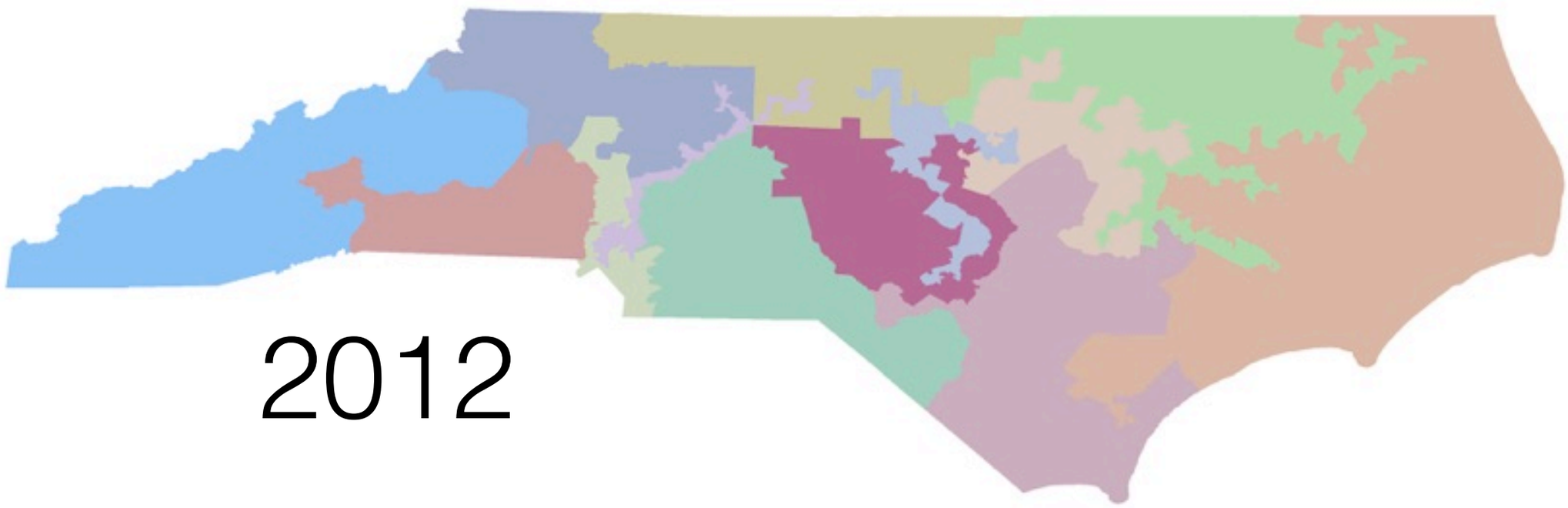
Election Results for Ensemble of Redistricting Plans



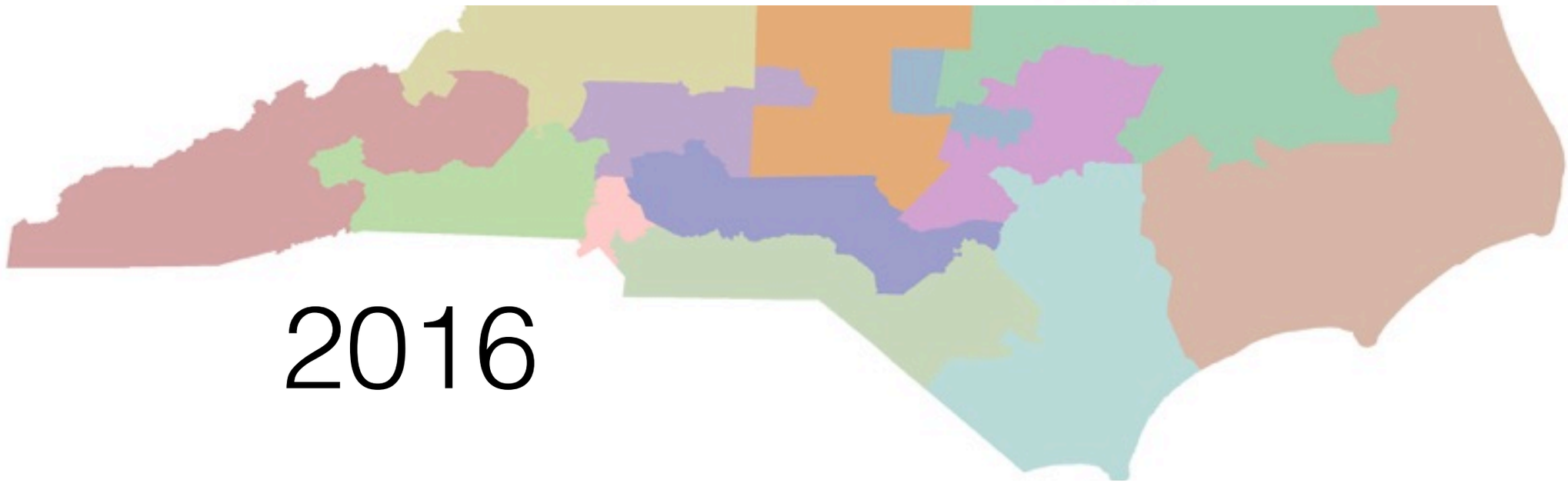


Historical and Shifted Elections

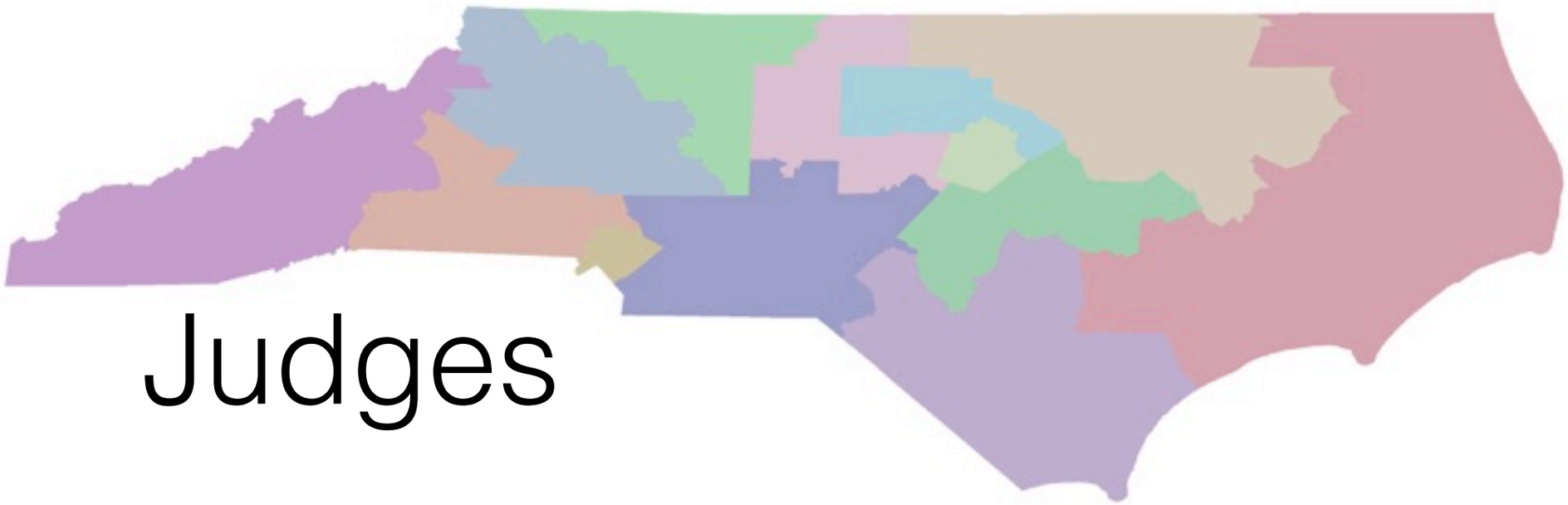




2012



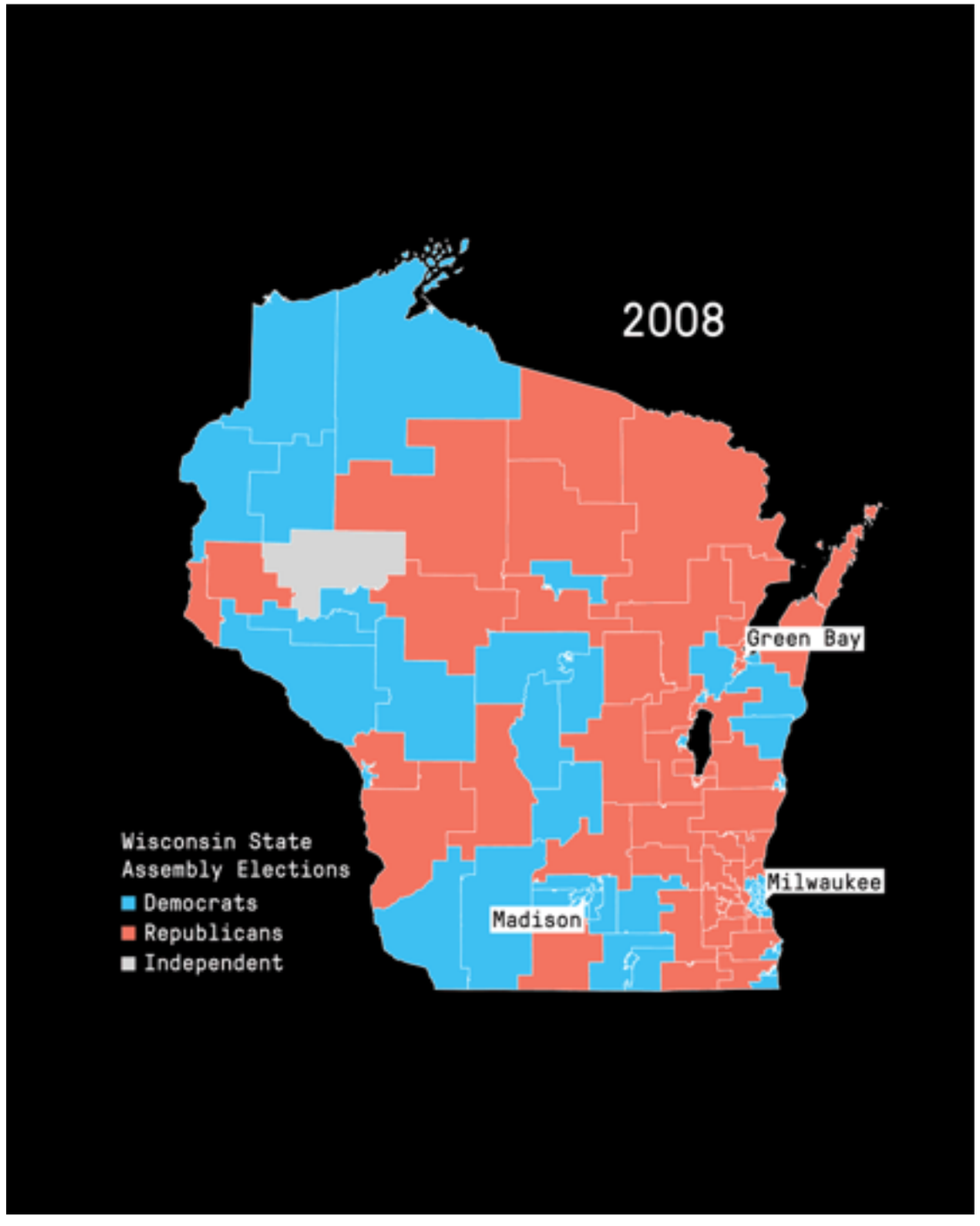
2016

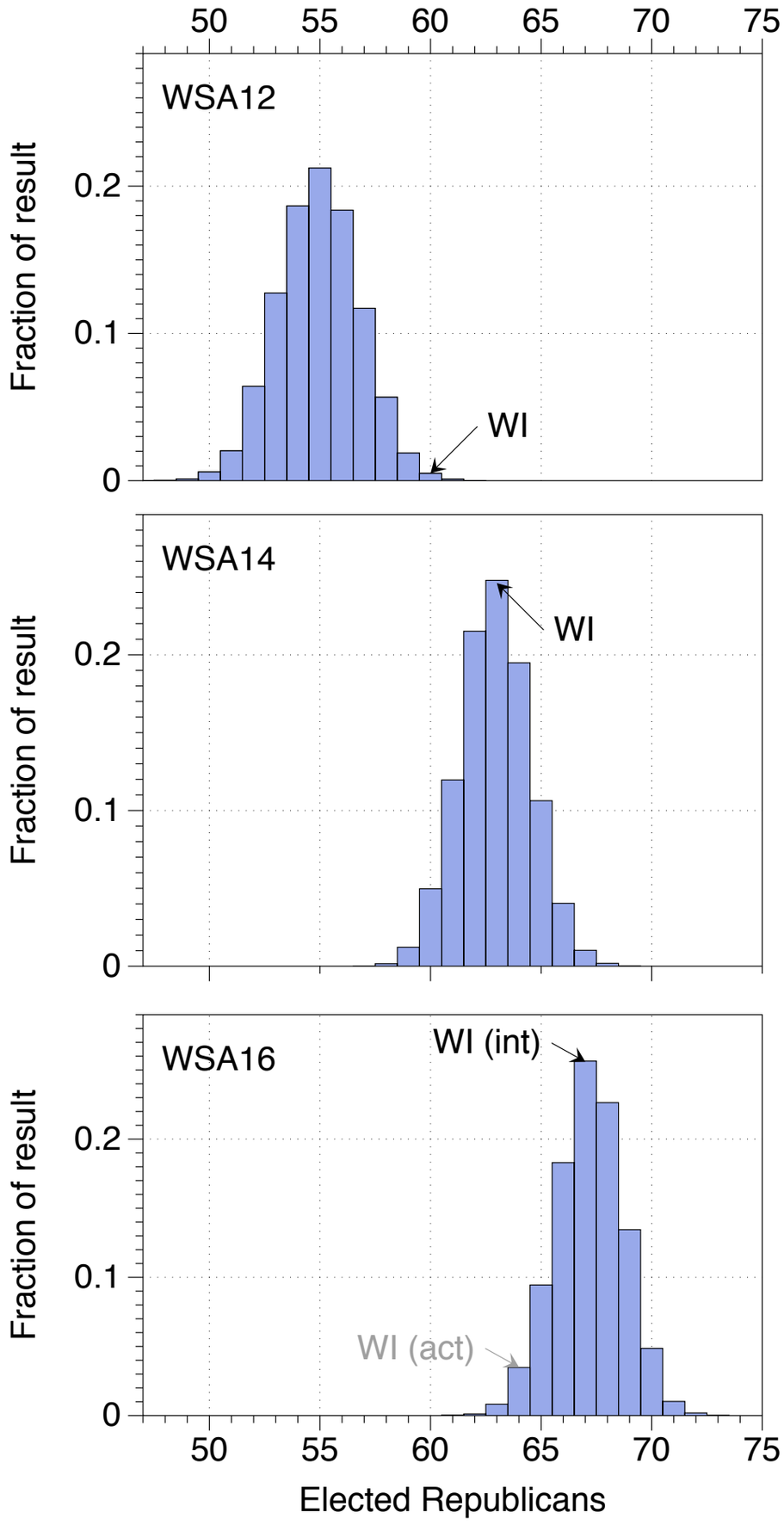


Judges

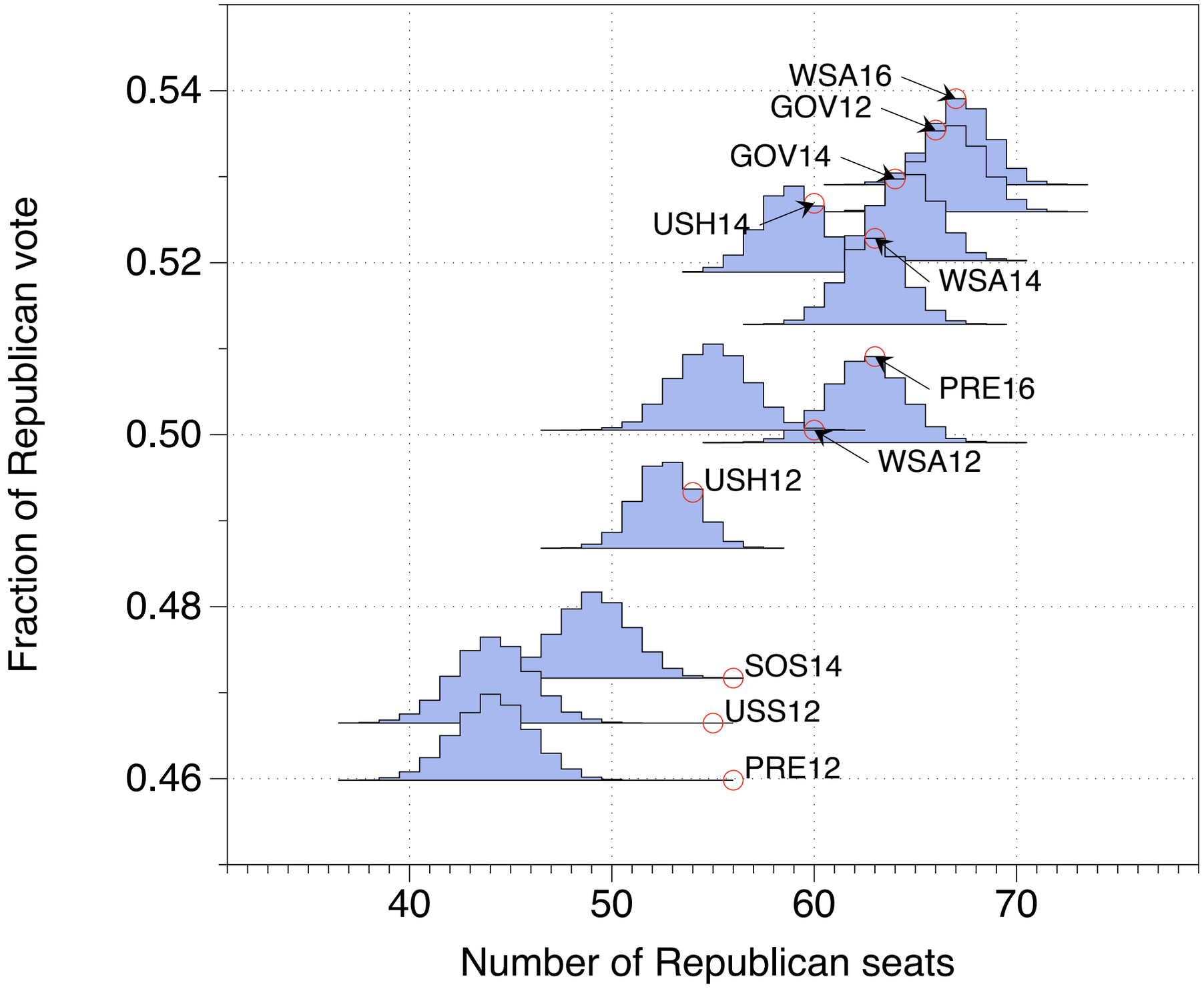
Other states ?

Wisconsin General Assembly

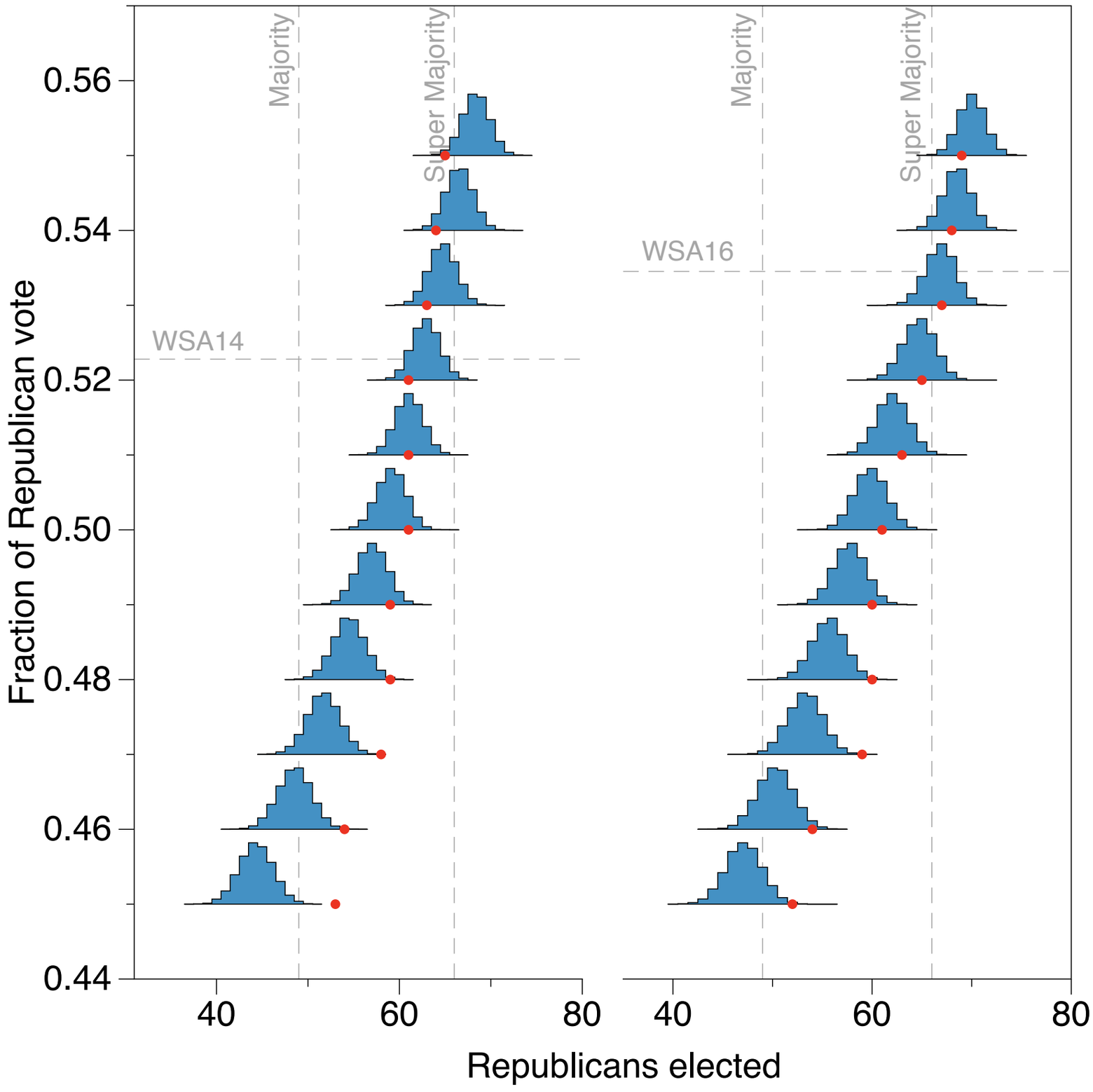




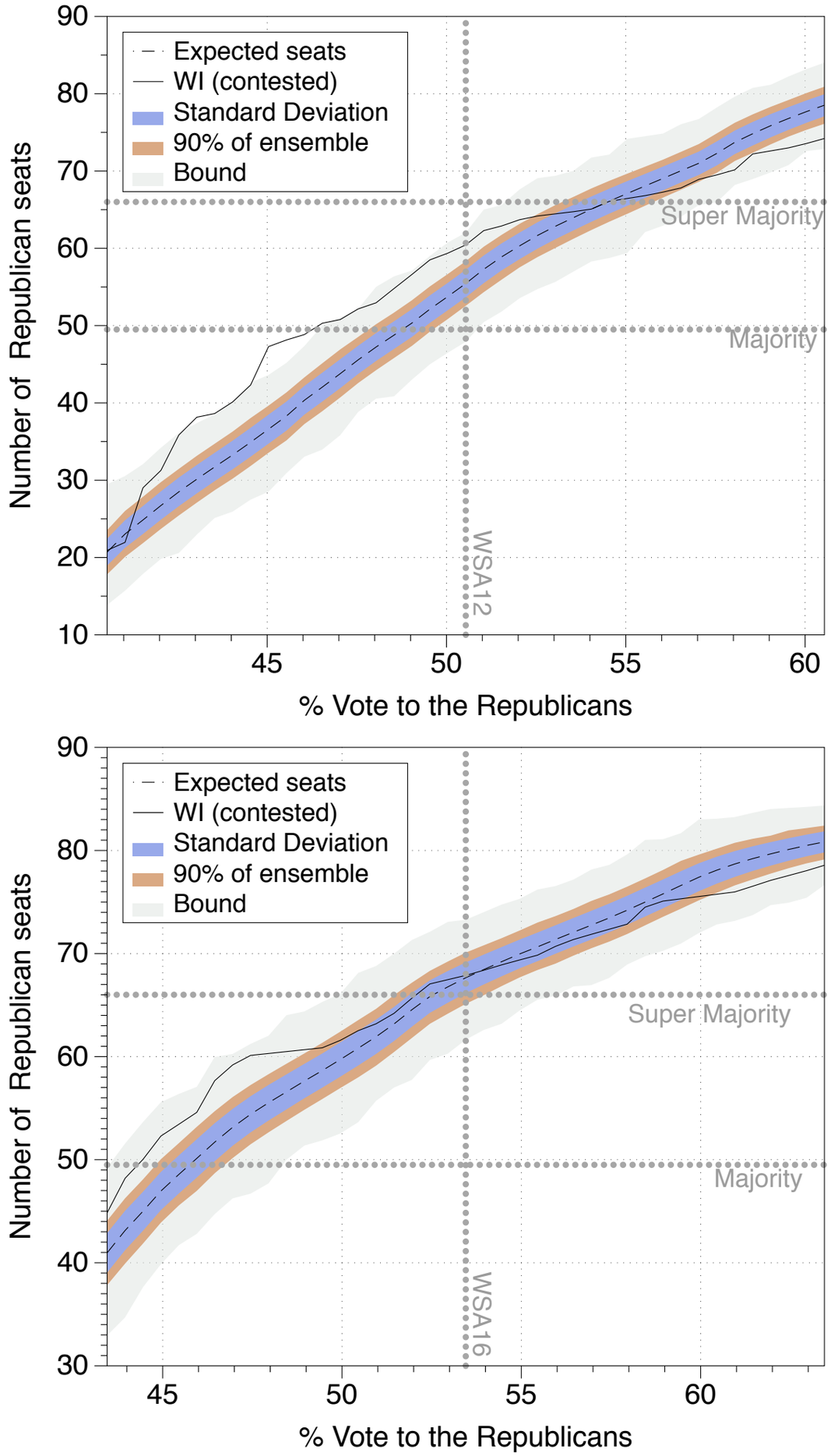
Wisconsin historical elections

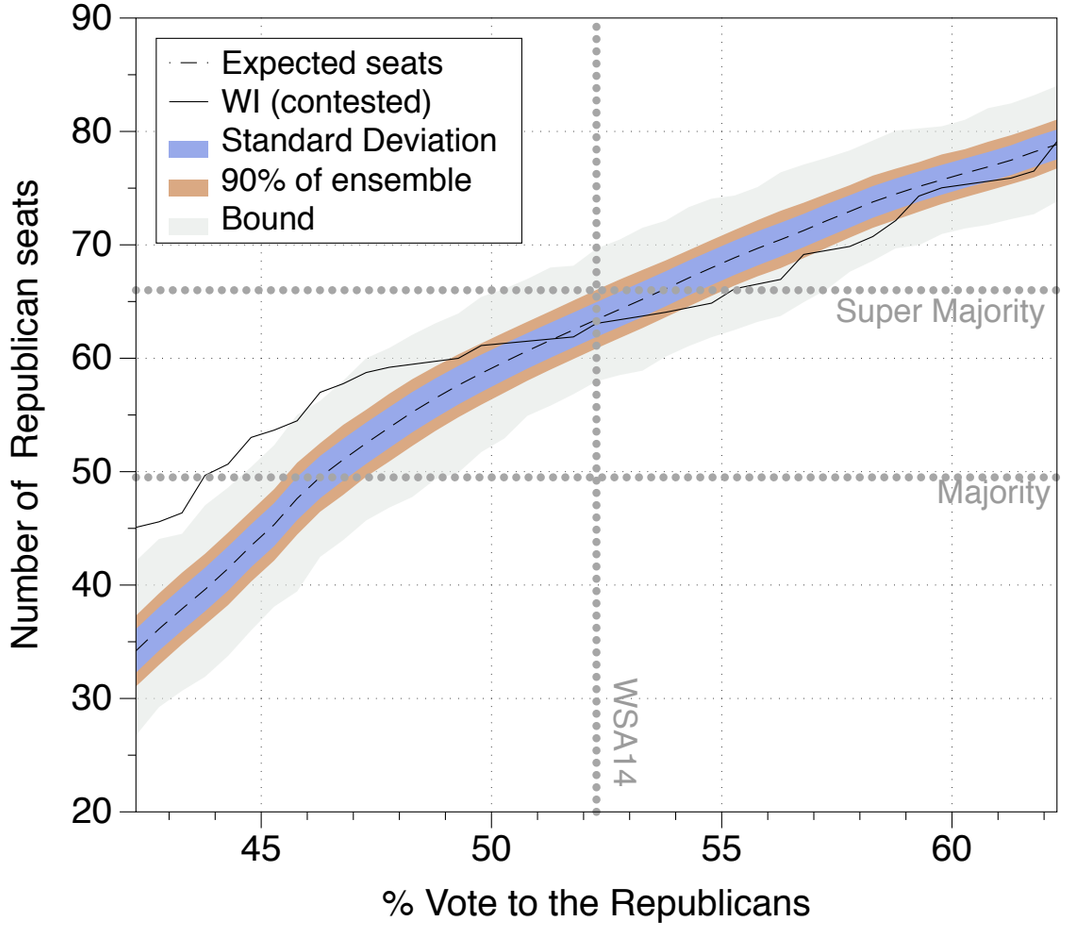


Shift the global percentages

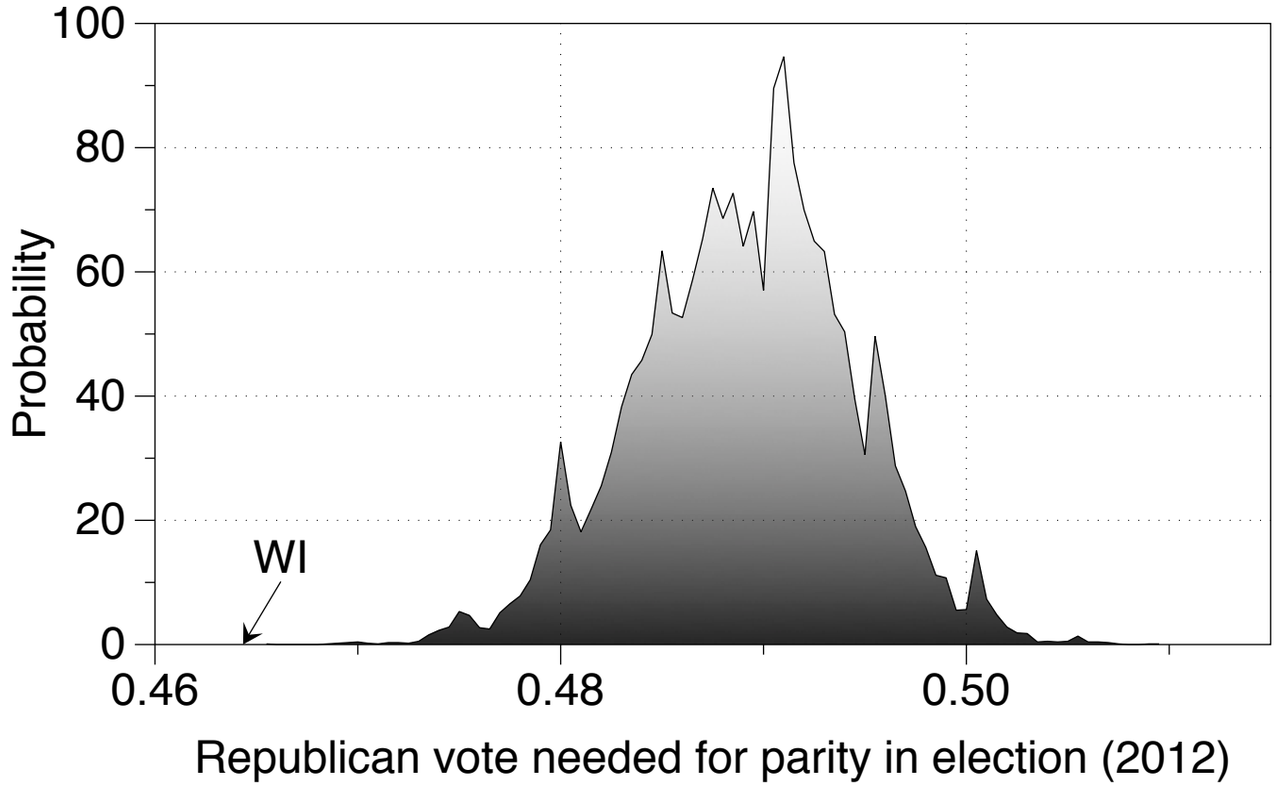


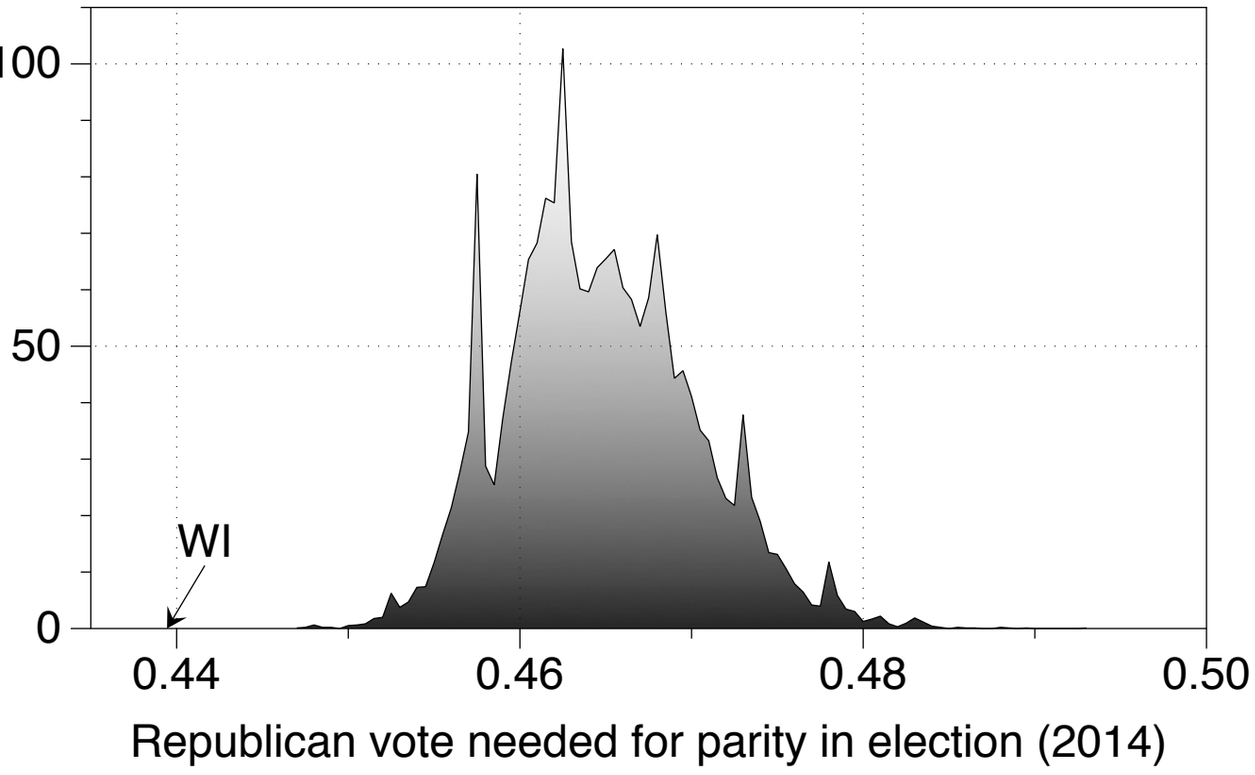
seats vs global vote

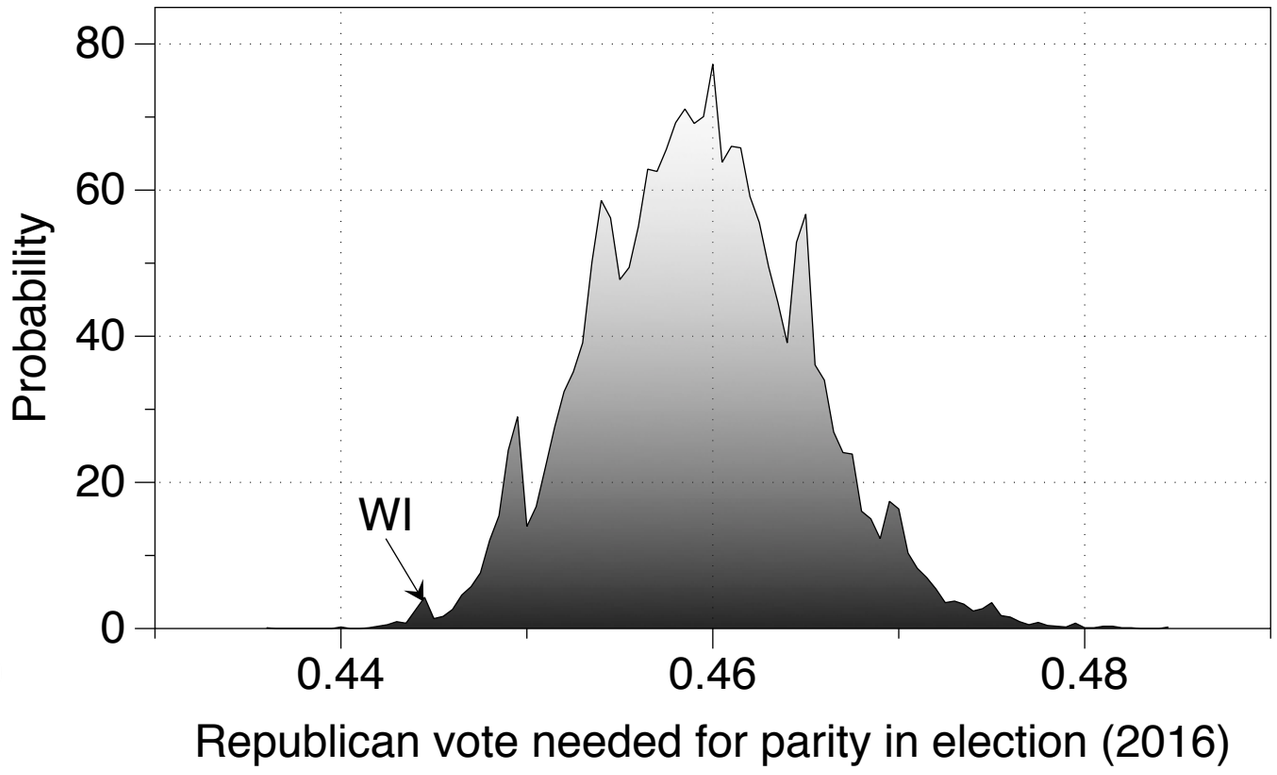




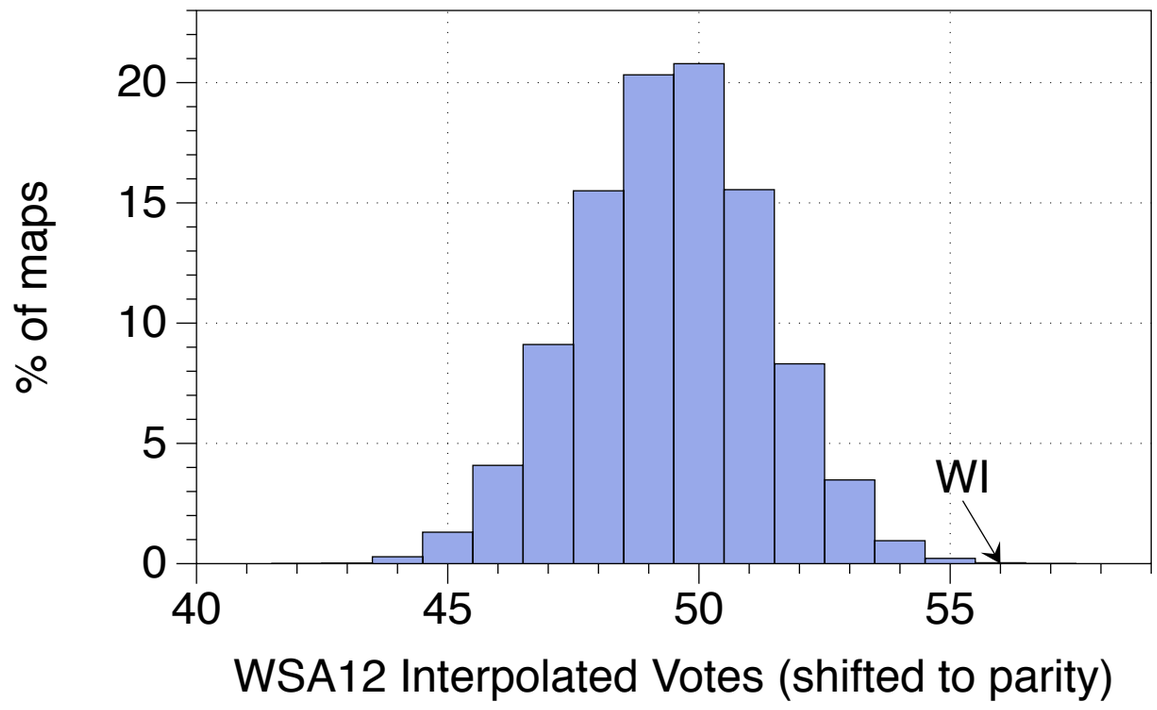
Structural advantage doesn't explain

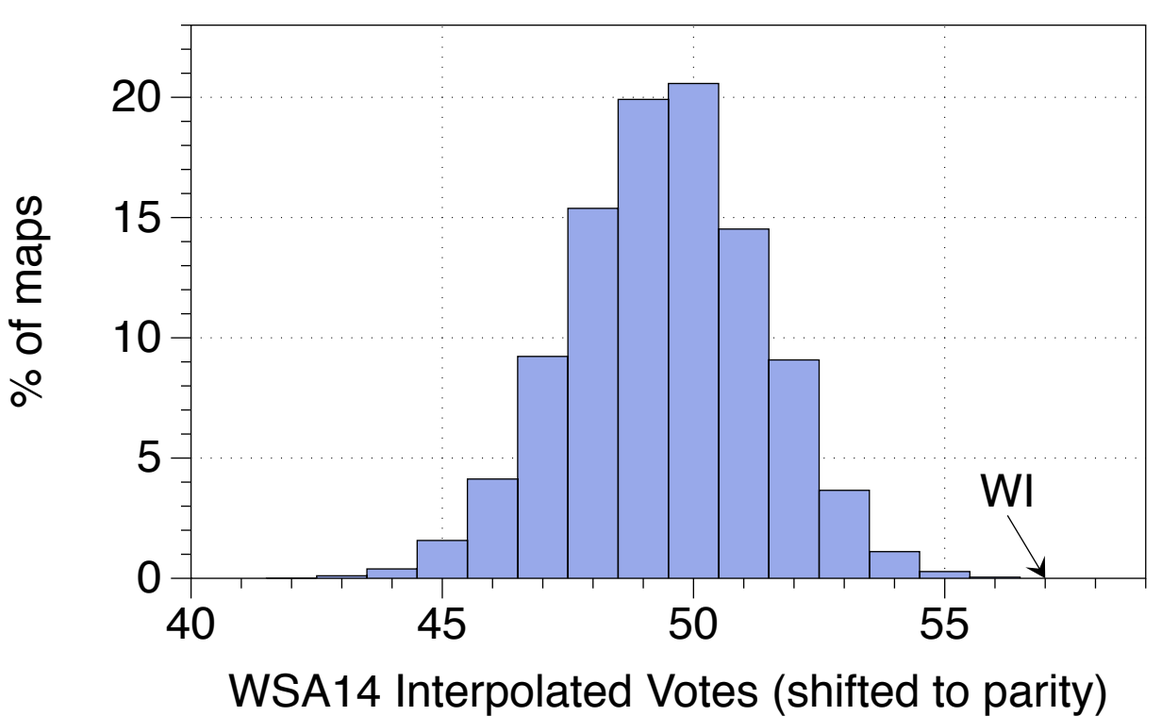


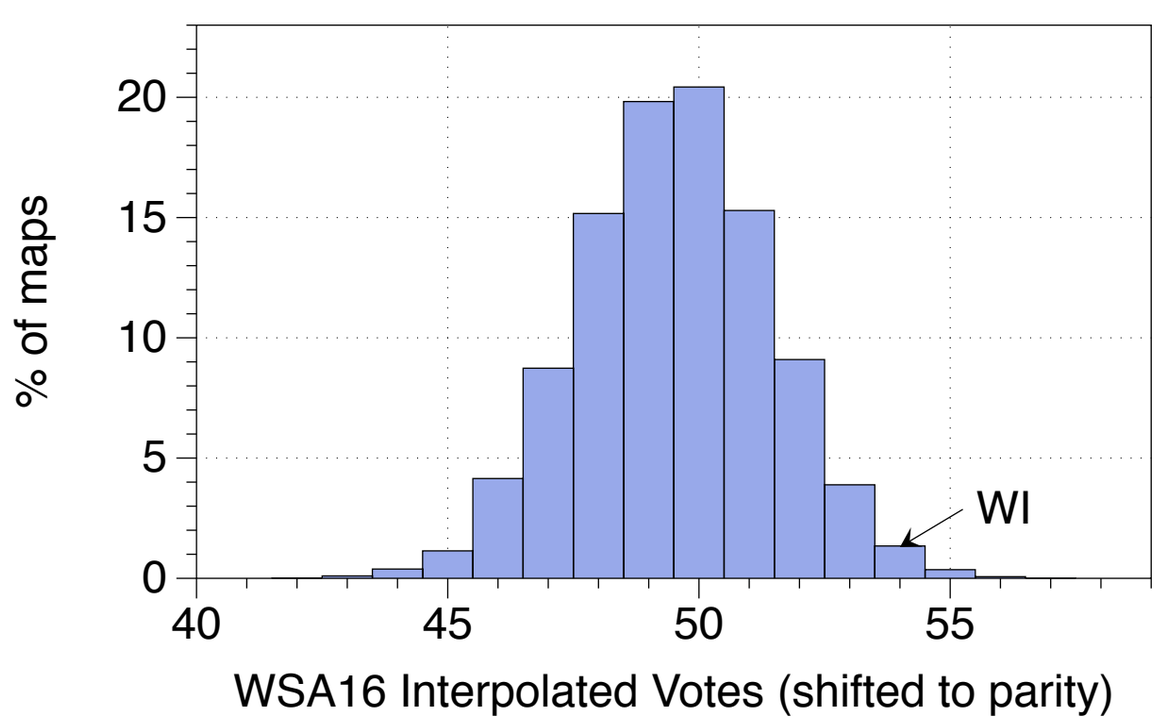




Shift to Median Parity

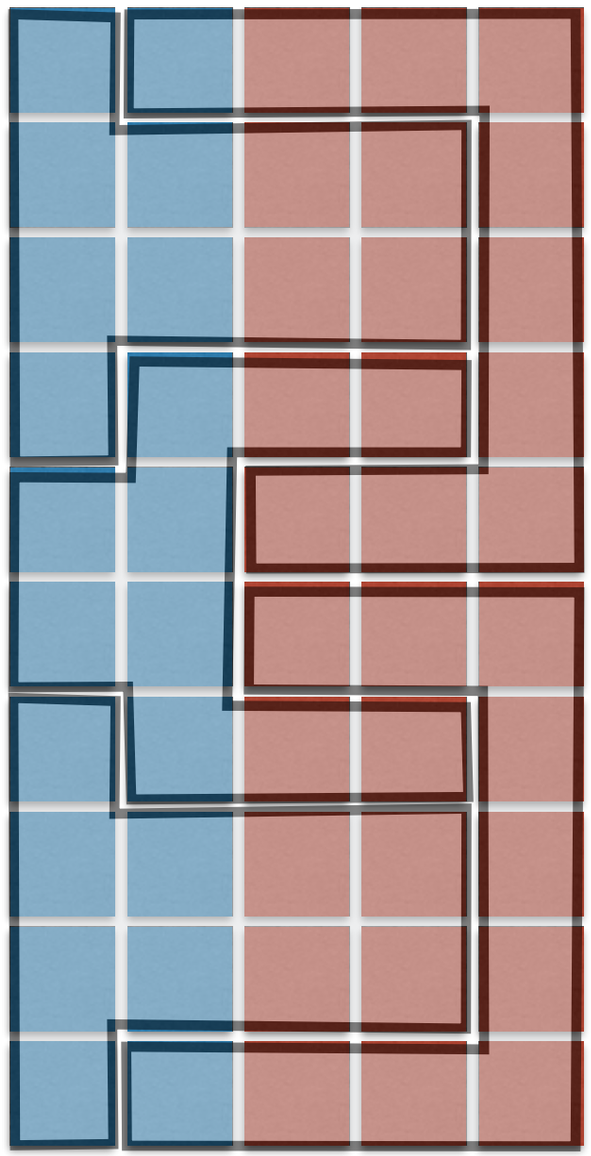






What produces these effects ?

What is the signature of gerrymandering ?



Red wins 2
Blue wins 3

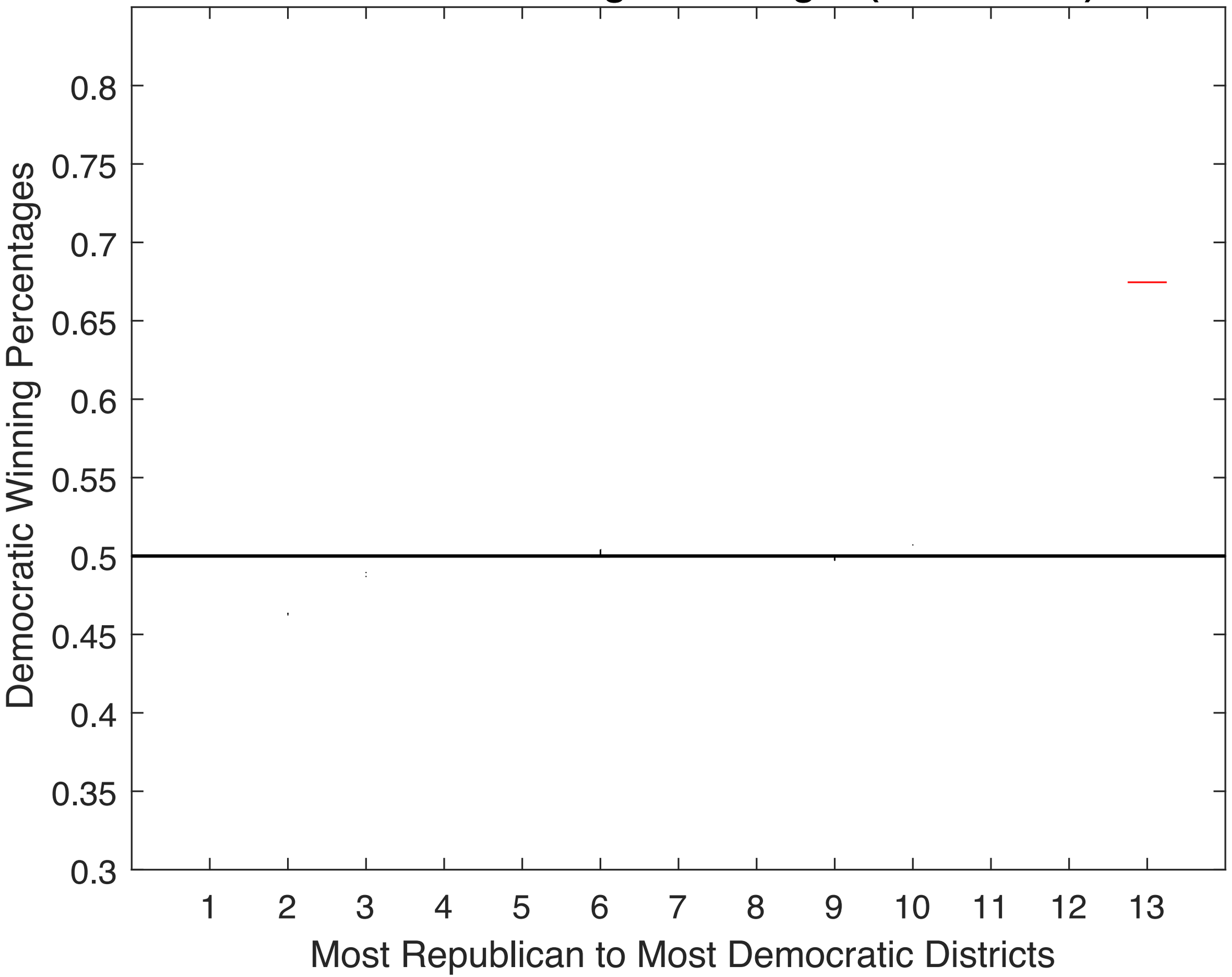
Red wins 2 districts by 8 votes each

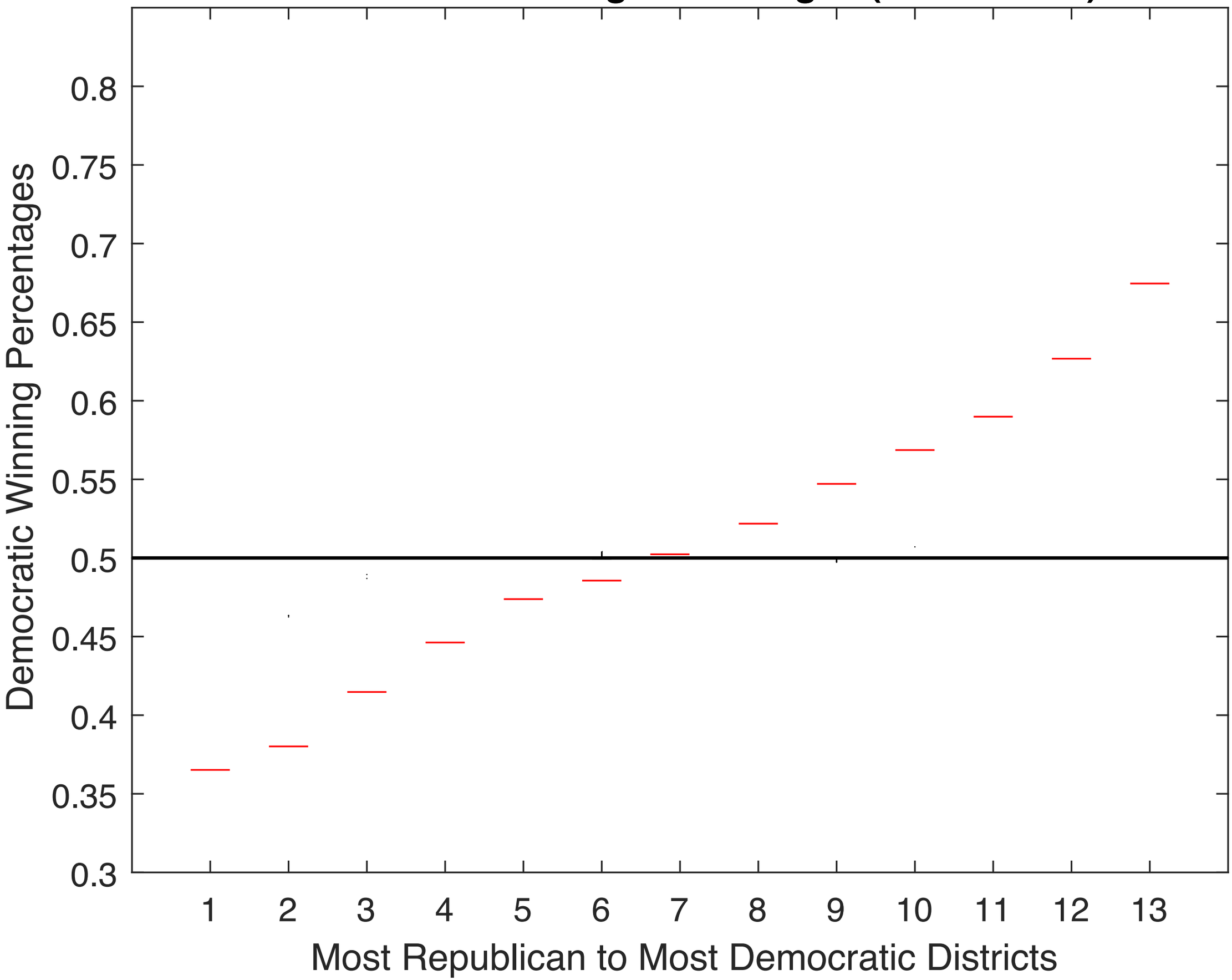
Blue wins 3 districts by 2 votes each

Percentage of Democrats from lowest to highest

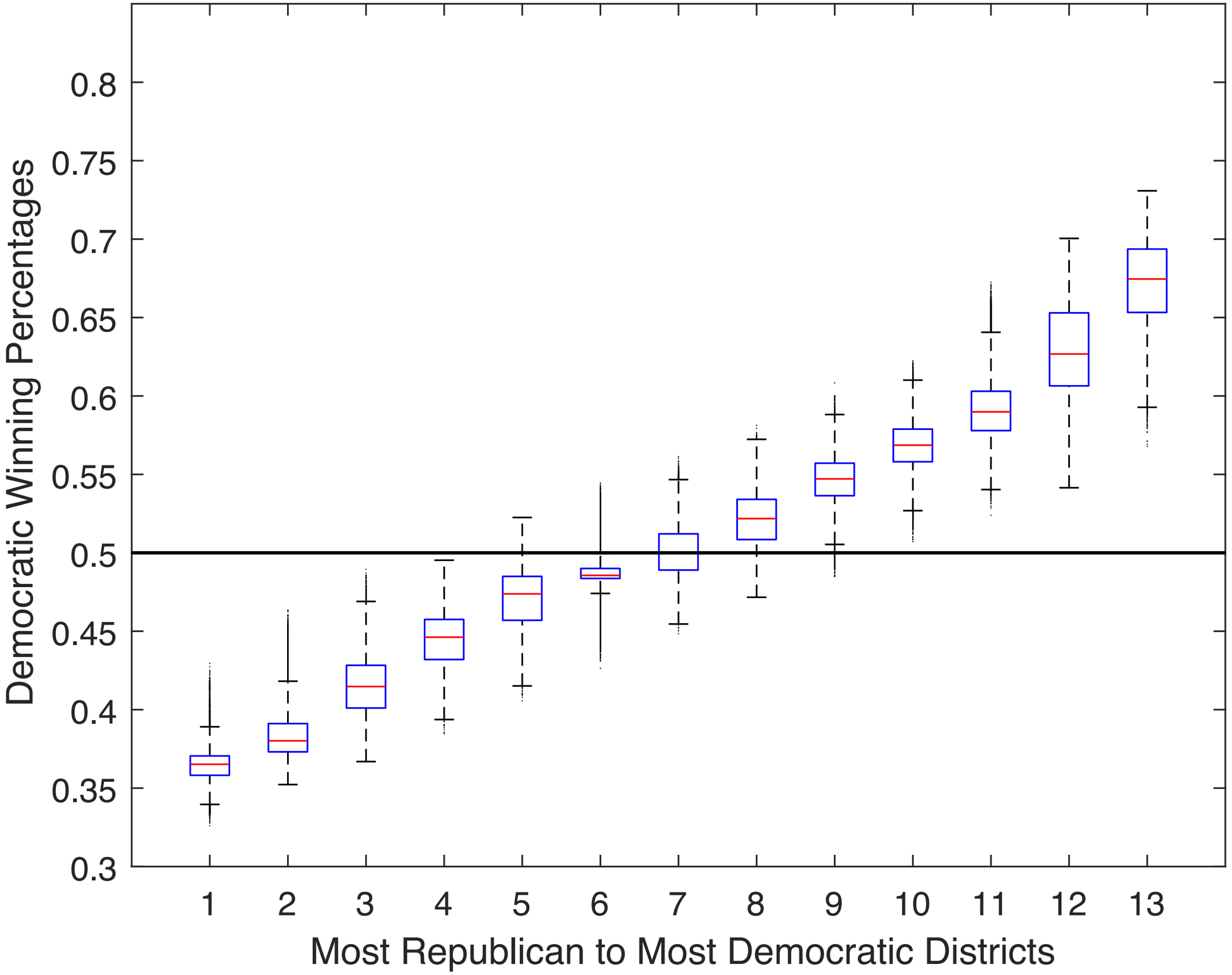
$$\begin{bmatrix} 10\% & 10\% & 60\% & 60\% & 60\% \end{bmatrix}$$

Democratic Winning Percentages (House 2012)

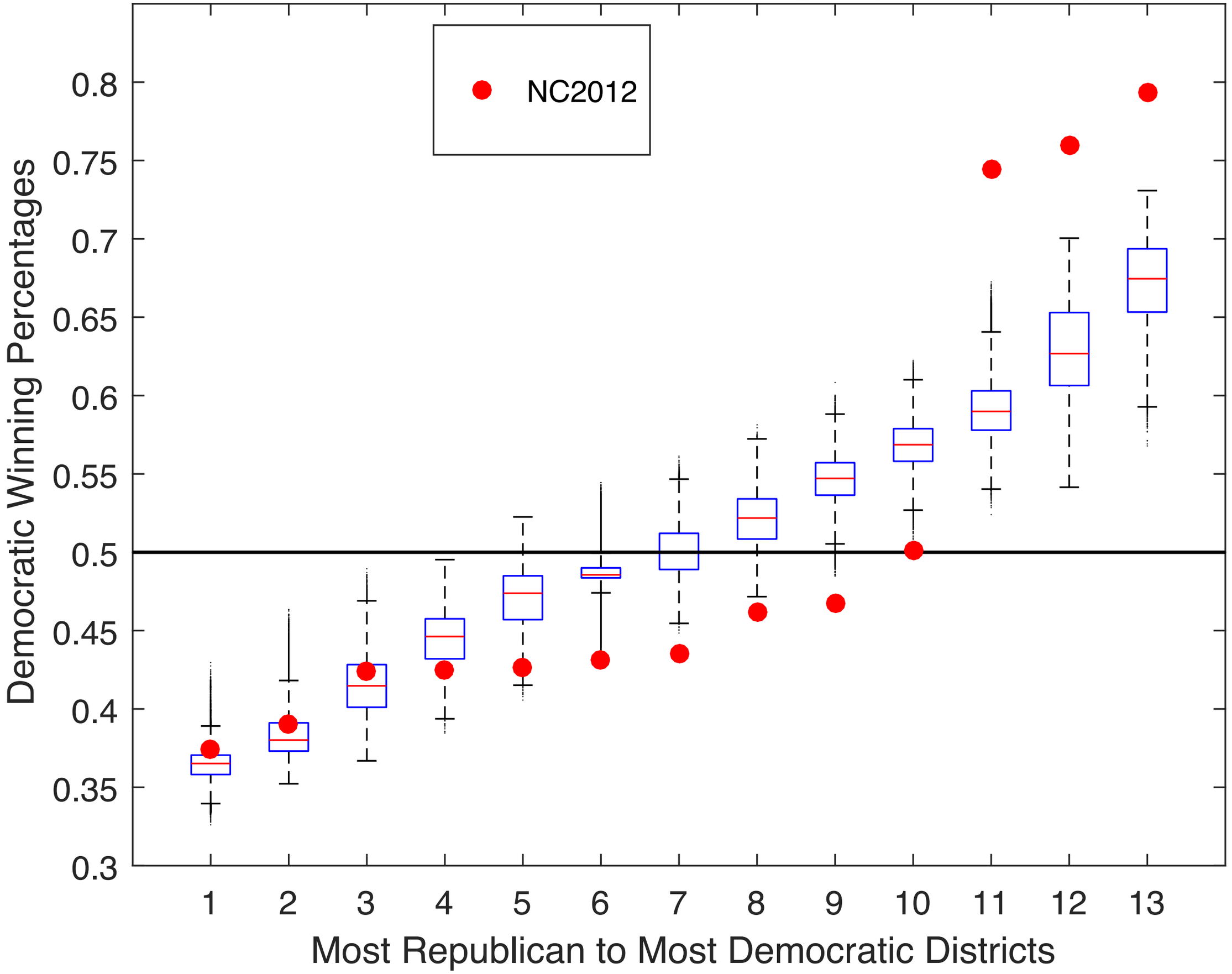



Democratic Winning Percentages (House 2012)
Democratic Winning Percentages
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
1
2
3
4
5
6
7
8
9
10
11
12
13
Most Republican to Most Democratic Districts

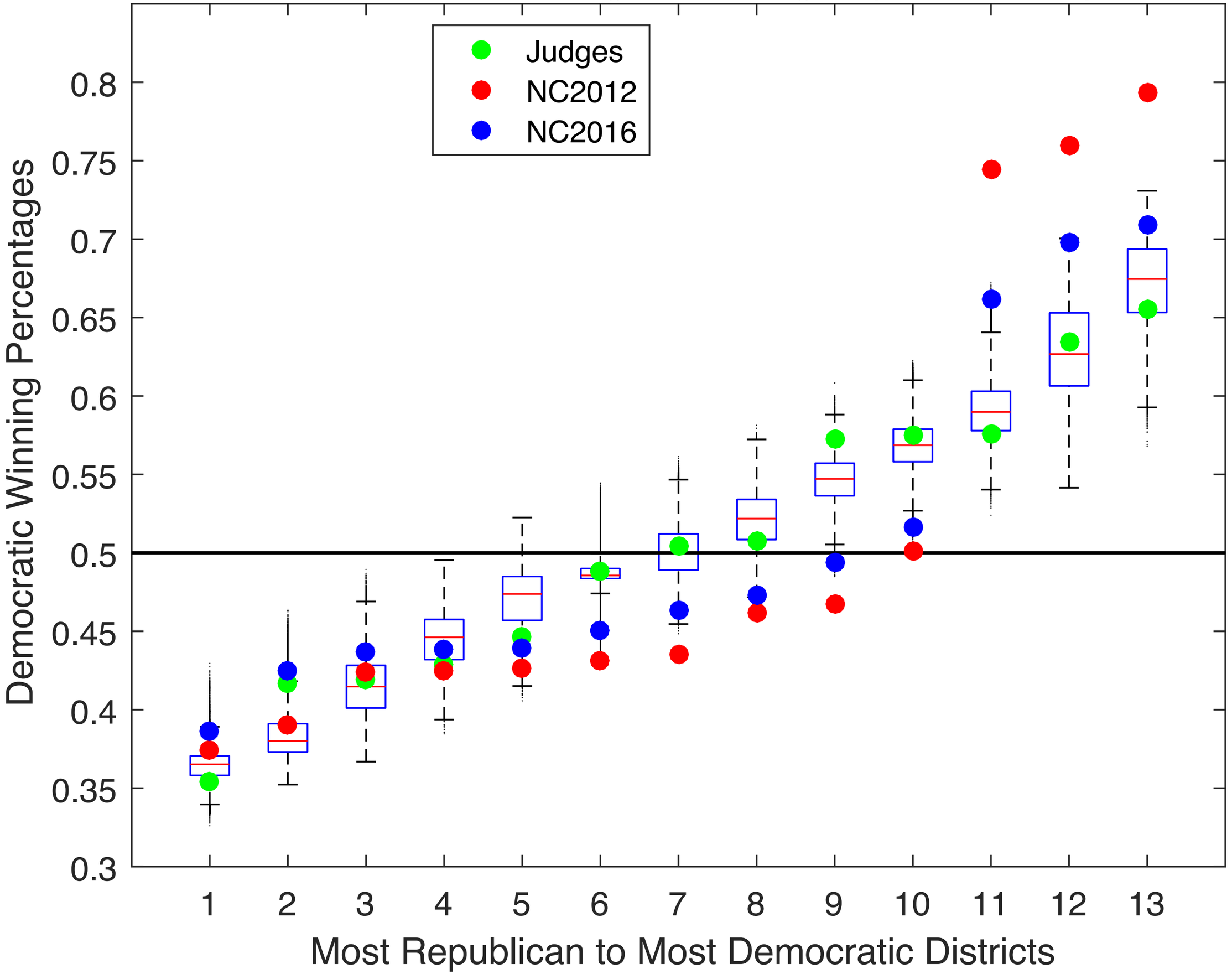
Democratic Winning Percentages (House 2012)

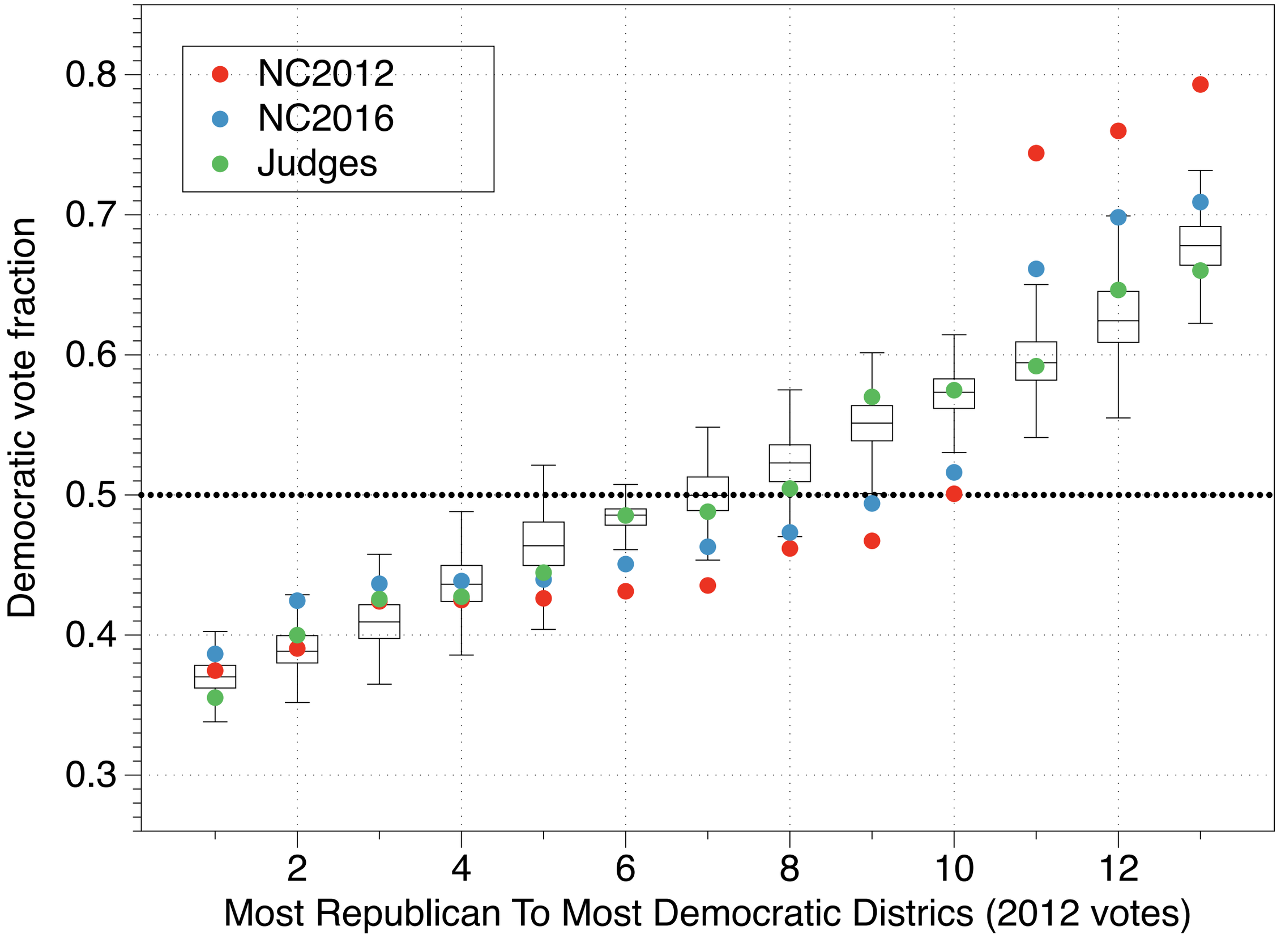


Democratic Winning Percentages (House 2012)

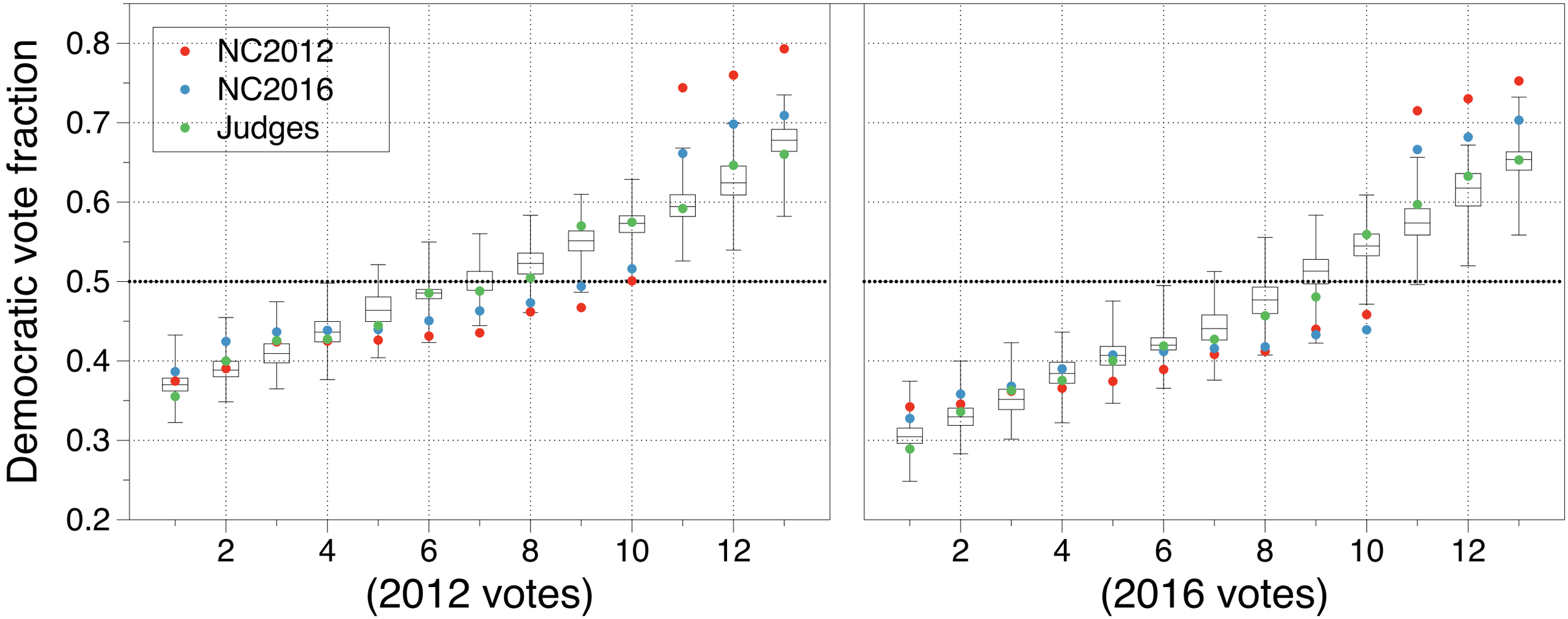


Democratic Winning Percentages (House 2012)



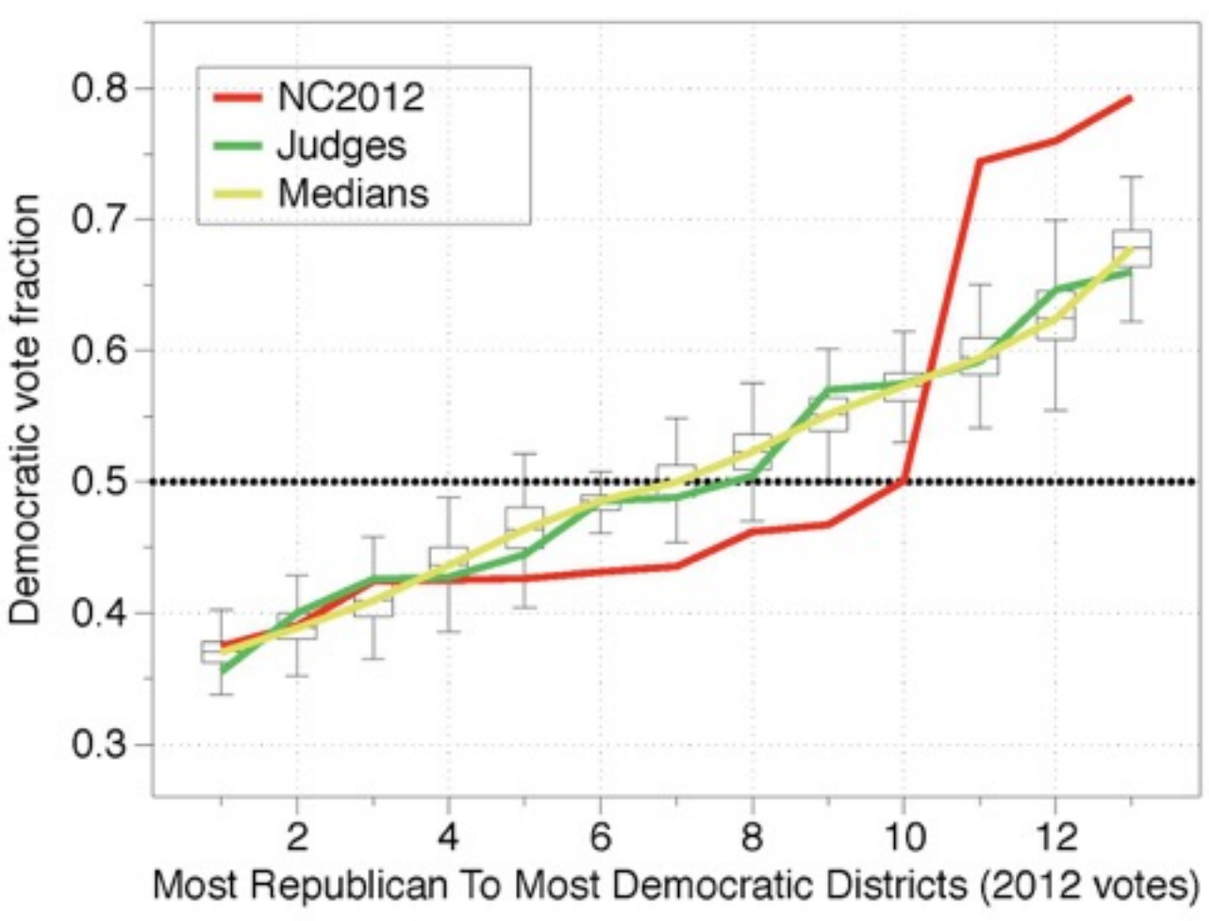

NC2012
NC2016
Judges
0.8
0.7
0.6
0.5
0.4
0.3
Democratic vote fraction
2
4
6
8
10
12
Most Republican To Most Democratic Districs (2012 votes)

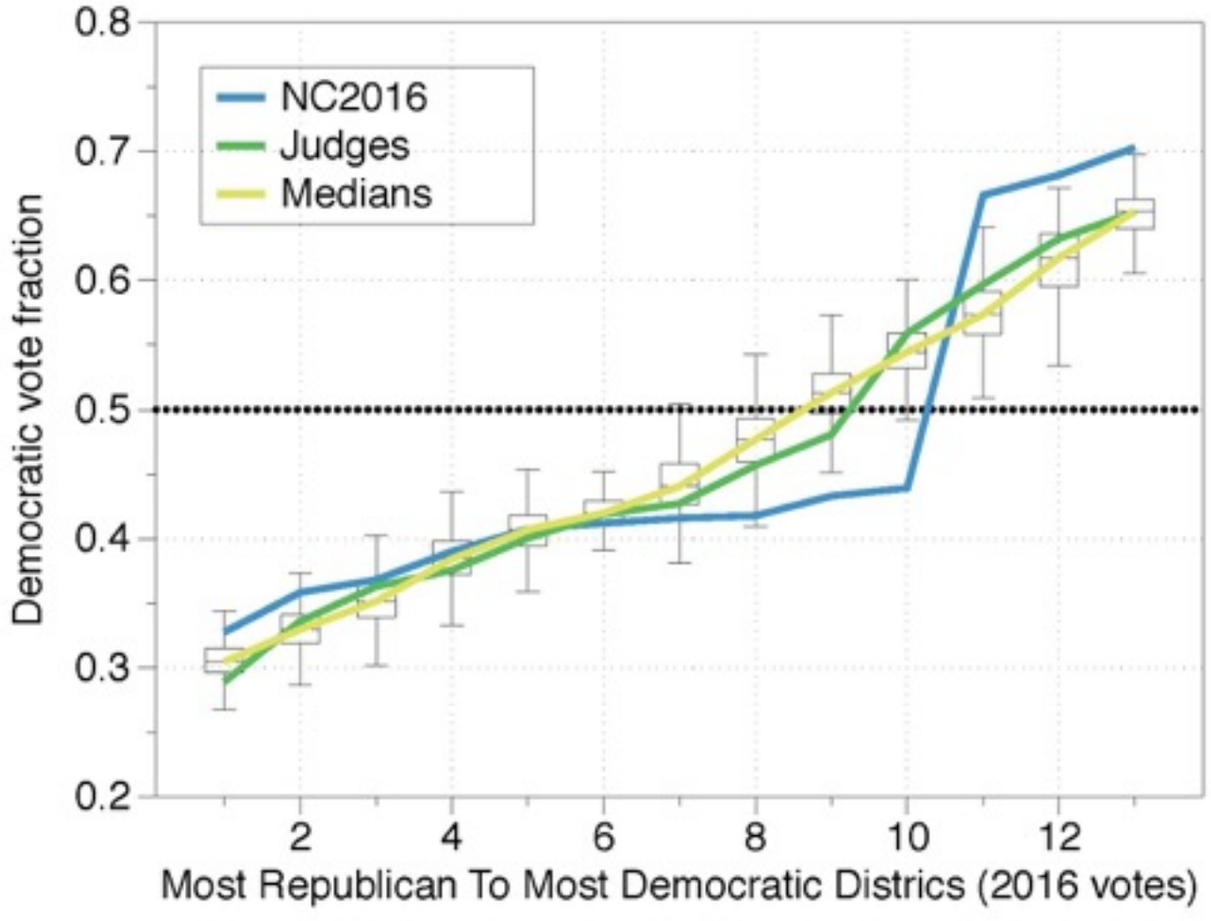
NC Congressional Delegation



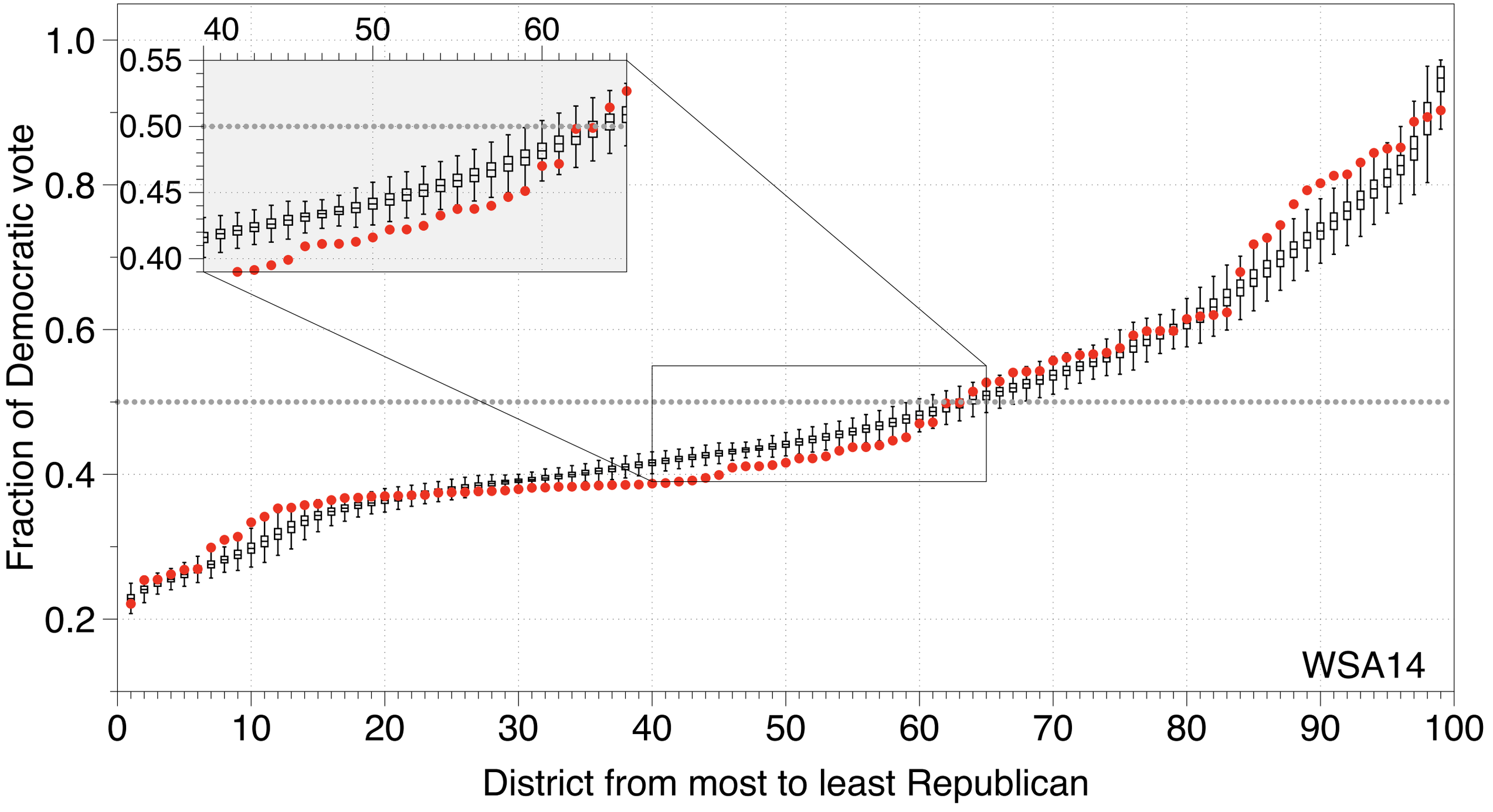
Identify Cracked and Packed districts

NC Congressional Delegation





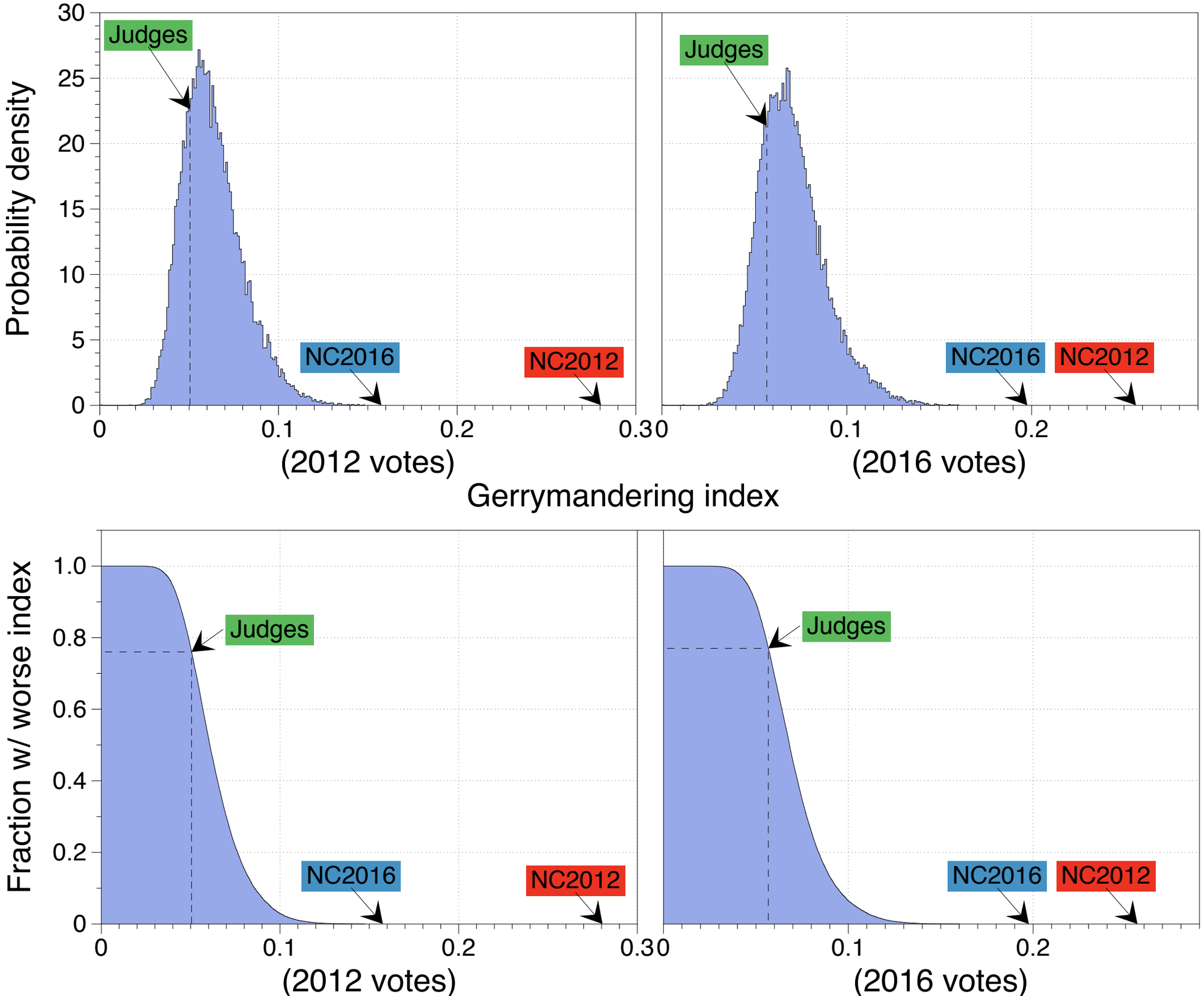
Wisconsin General Assembly



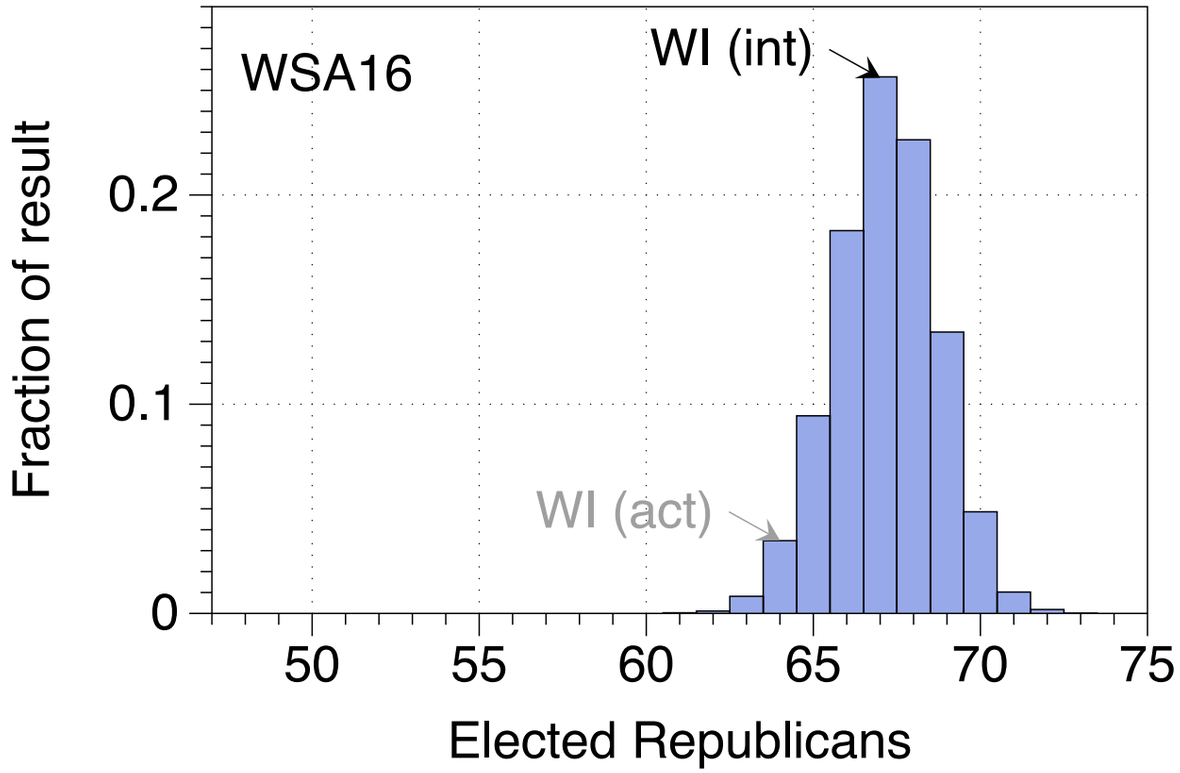
Metrics

Gerrymander Index

Measure deviation for expected district structure

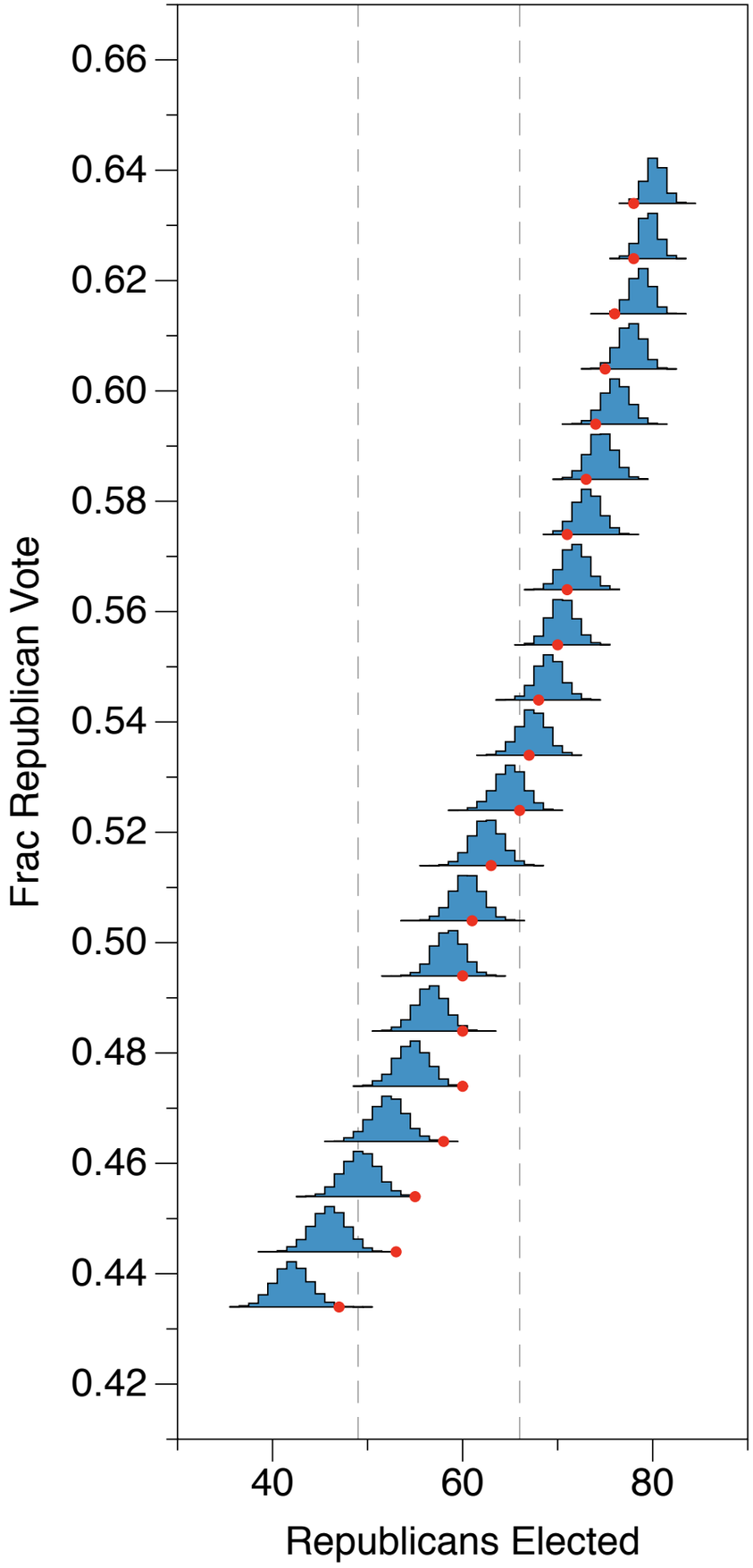


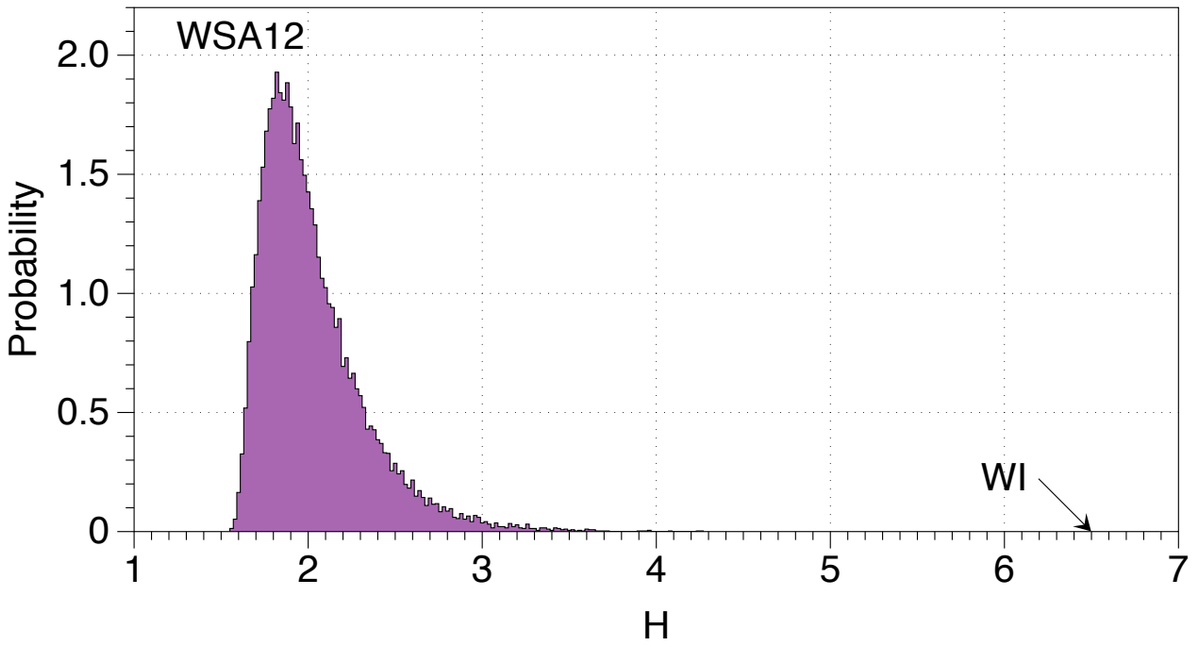
Measuring Representativeness

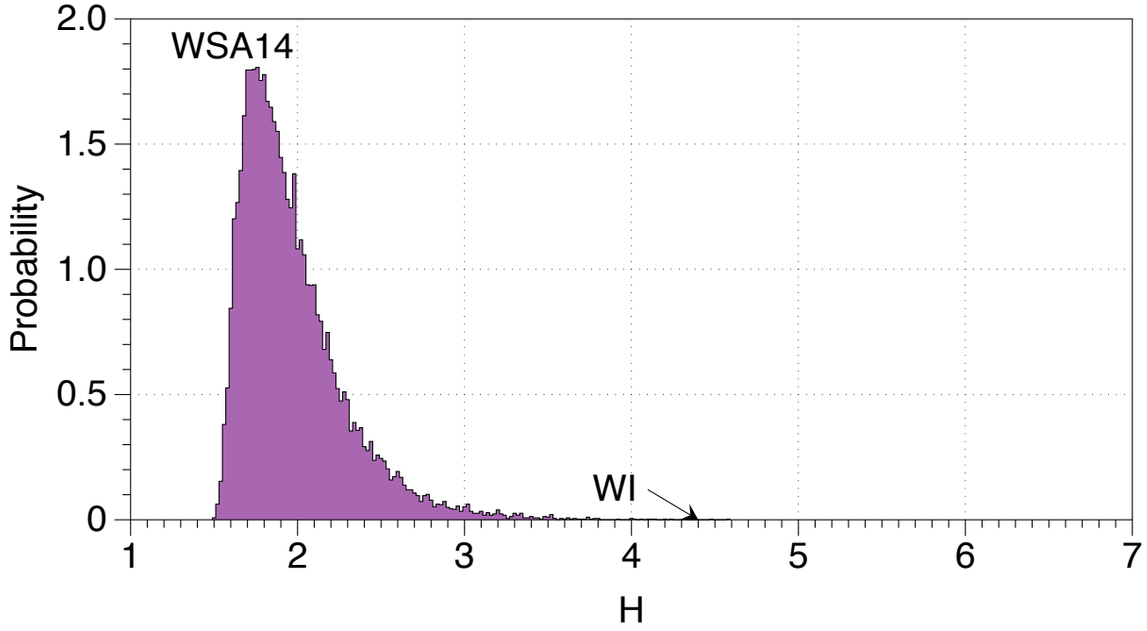


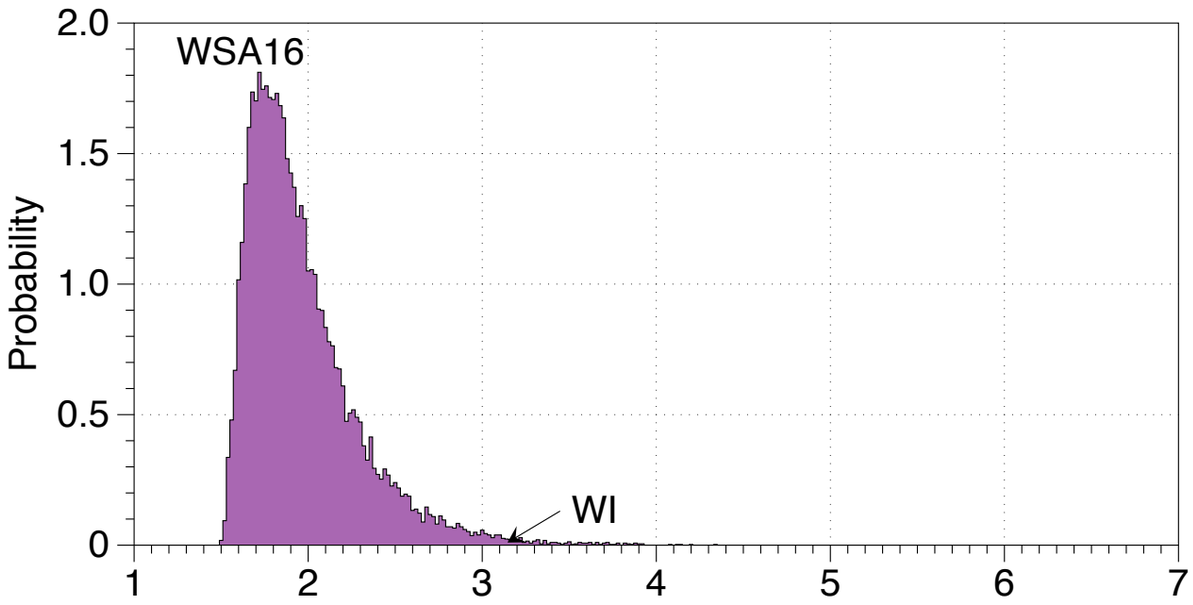
$$\ell(\text{map}) = -\log \text{Prob}(\text{outcome map produces})$$

Average $\ell(\text{map})$ over shift







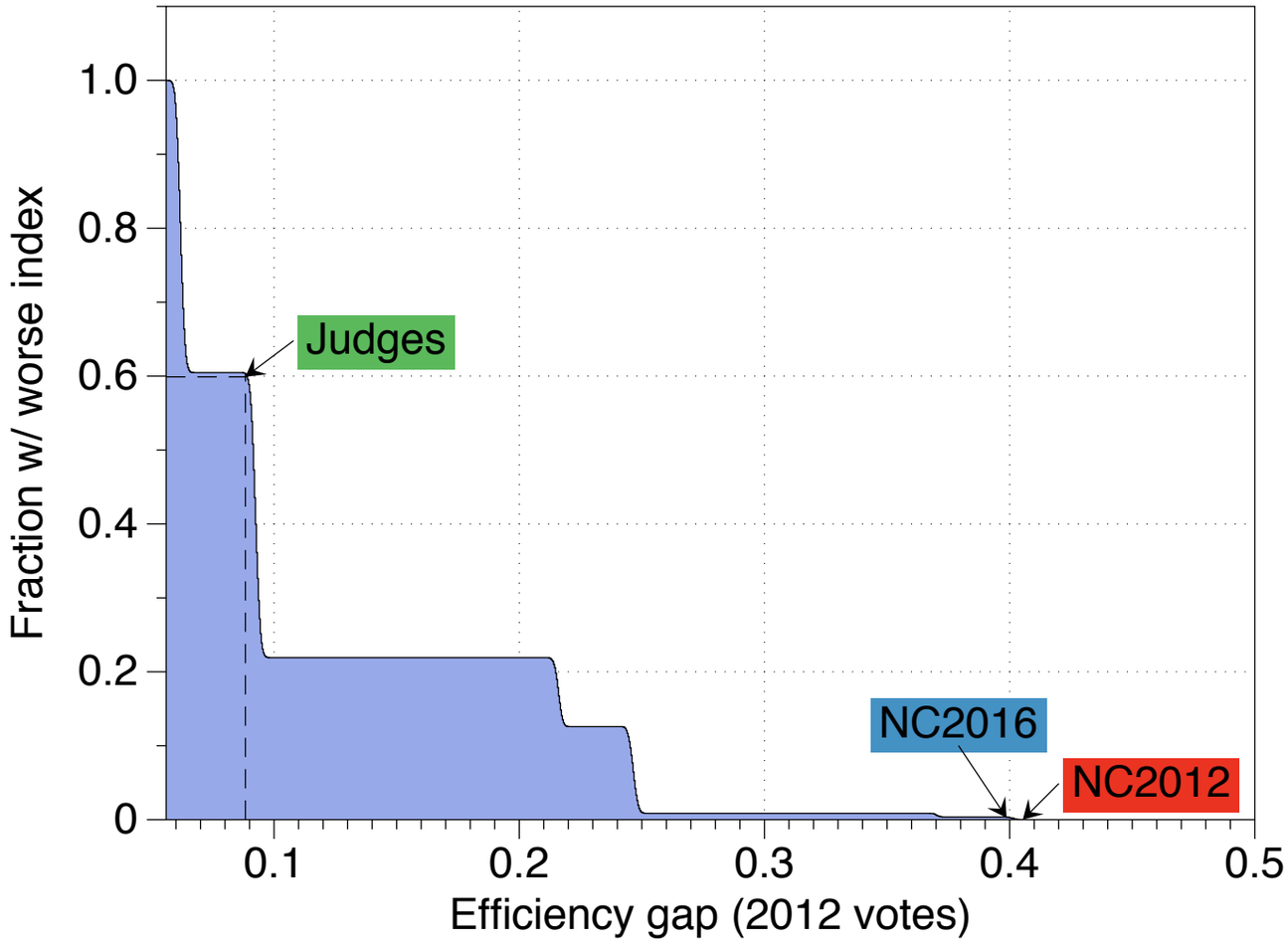


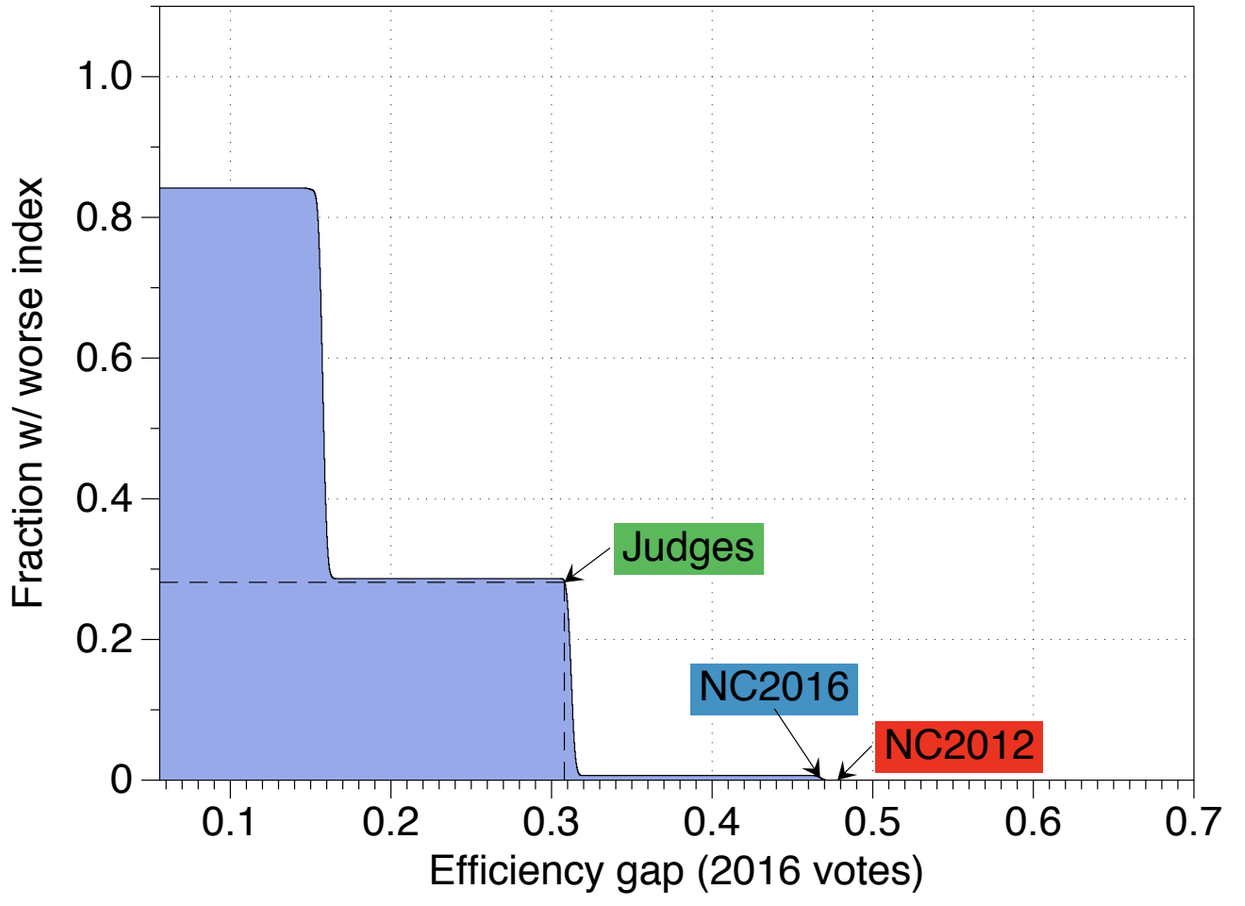
Measuring Representativeness

The Wisconsin plans are clearly an outlier for the average log likelihood over shifts 45%-55%

Efficiency Gap

Situate EG inside the ensemble

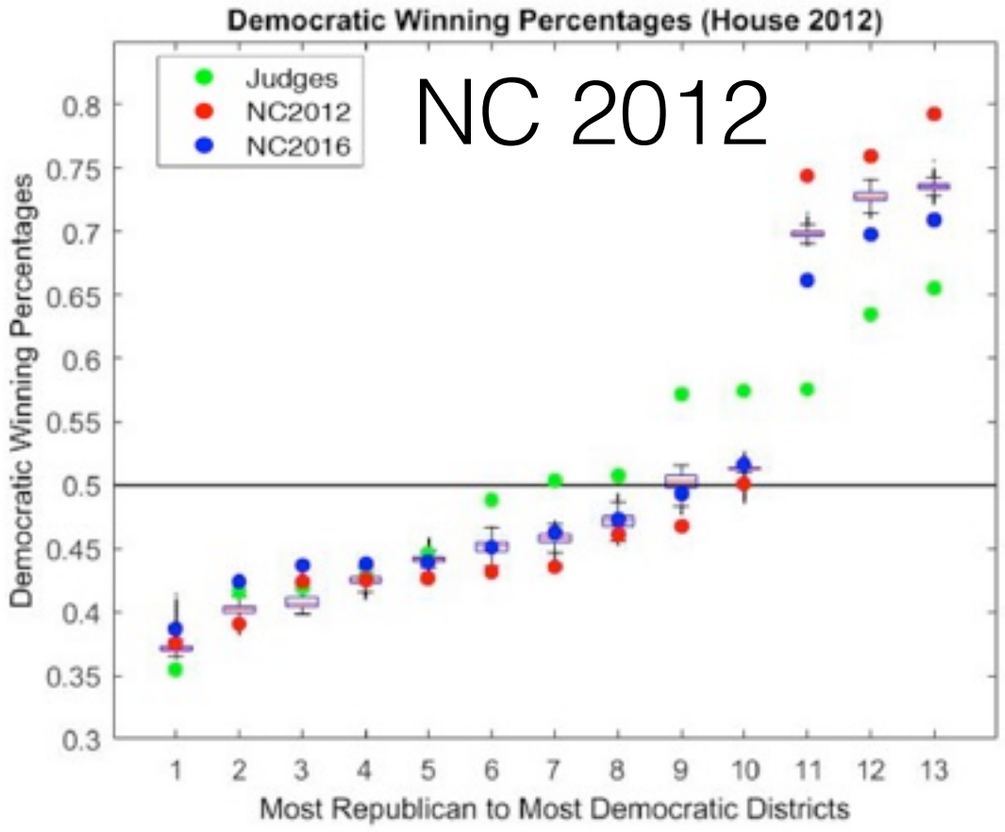


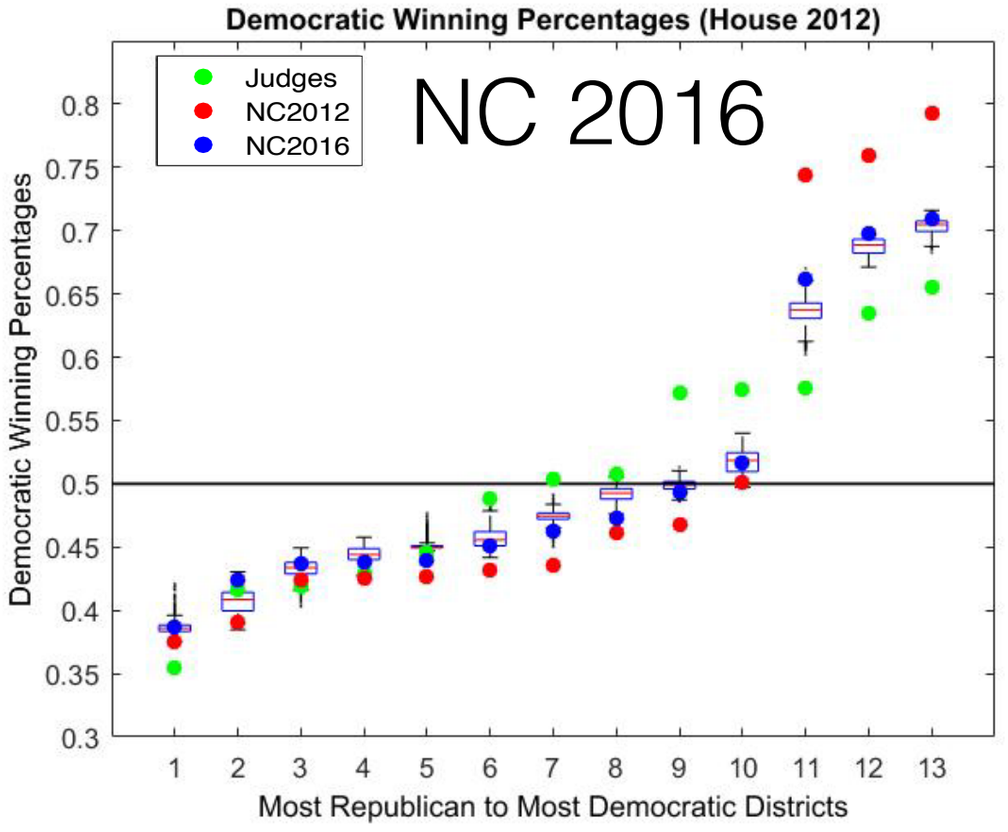


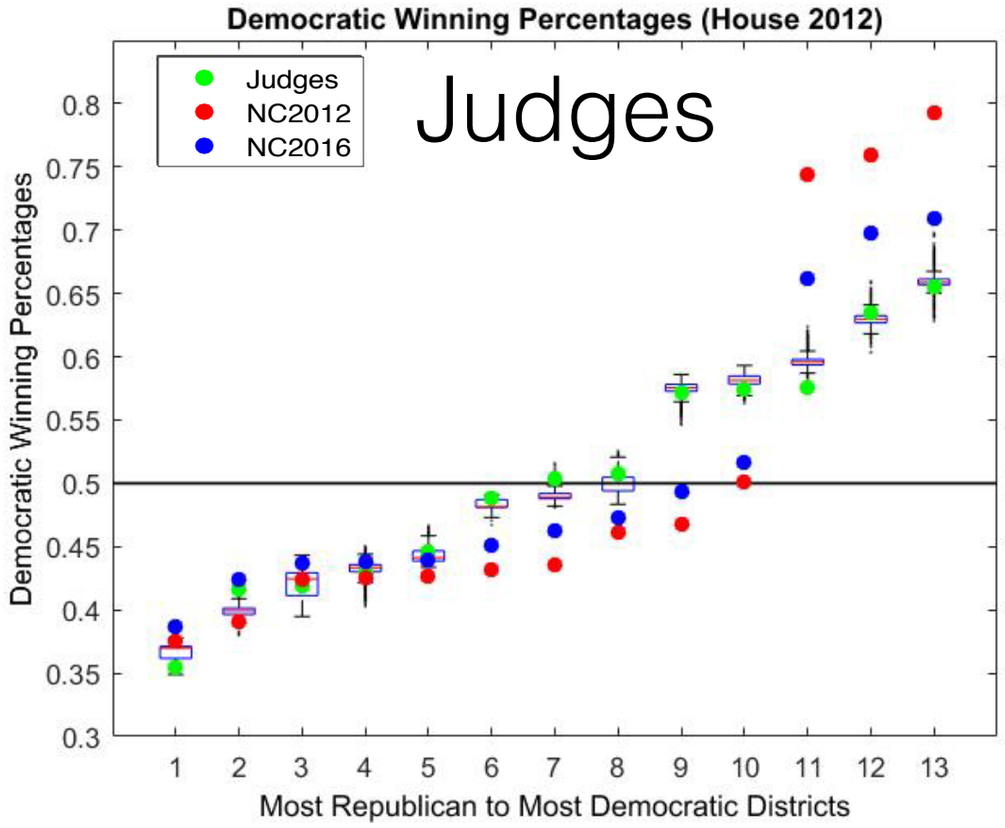
Engineered ?

results should be stable under small changes to districts

sample near by districts and observe changes





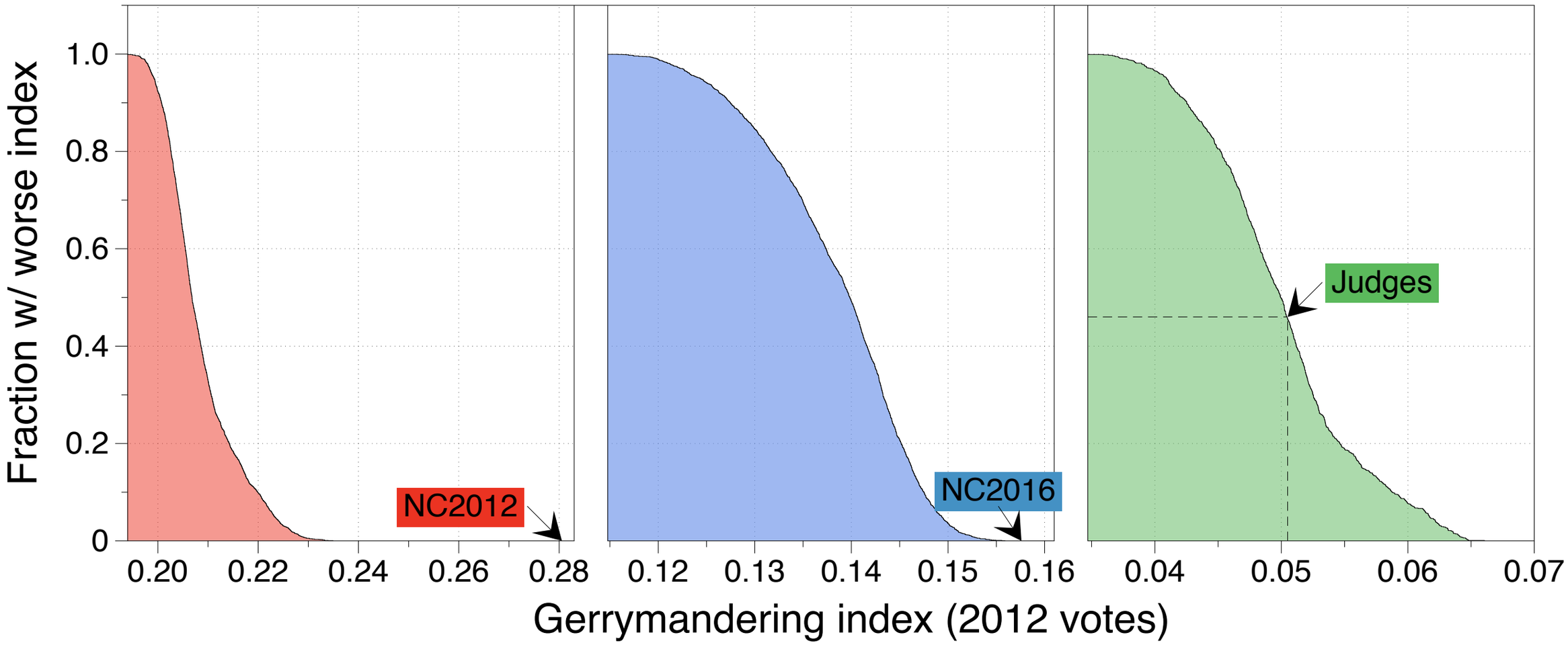


Judge's districts resemble near by districts

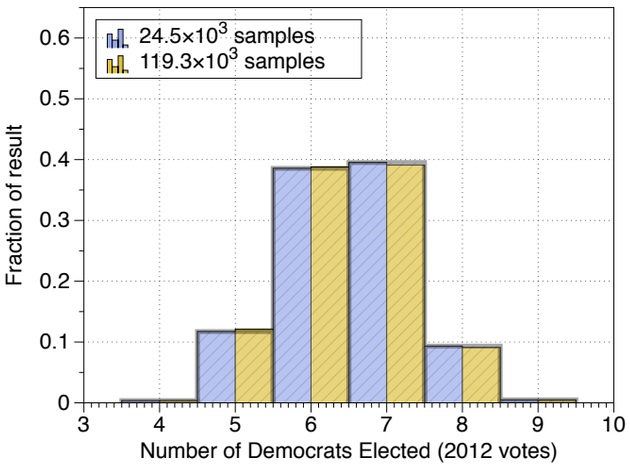
NC 2012 and NC 2016 do not

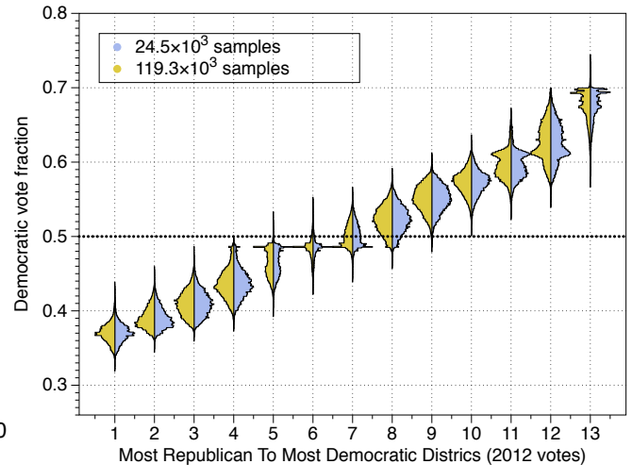
Gerrymander Index

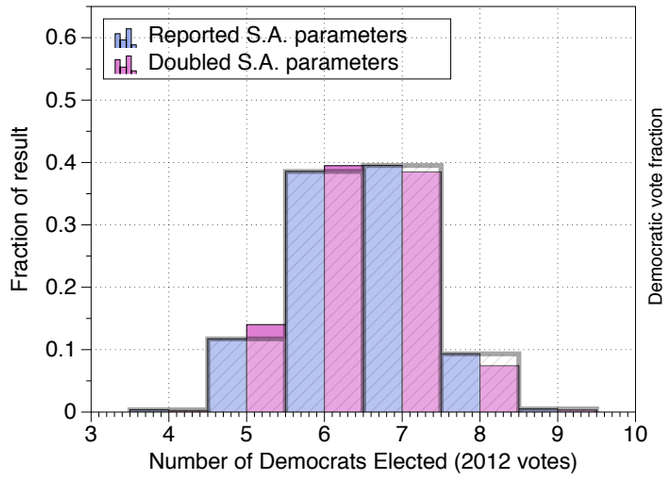
Local Perturbations

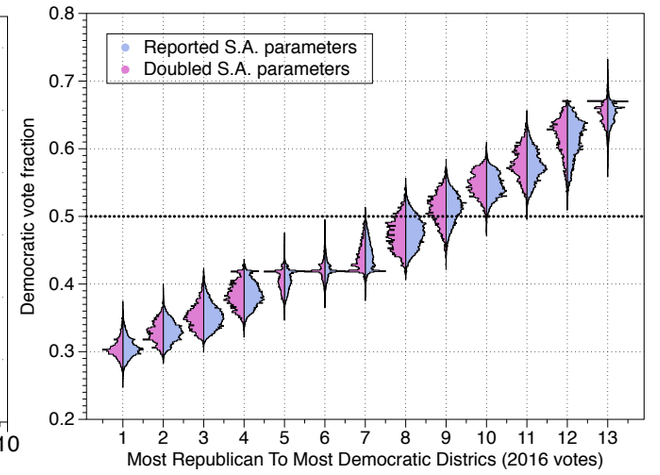


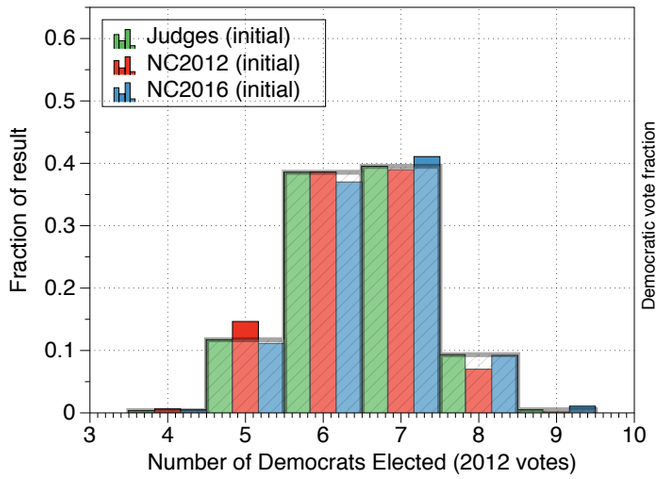
Stability of Conclusions

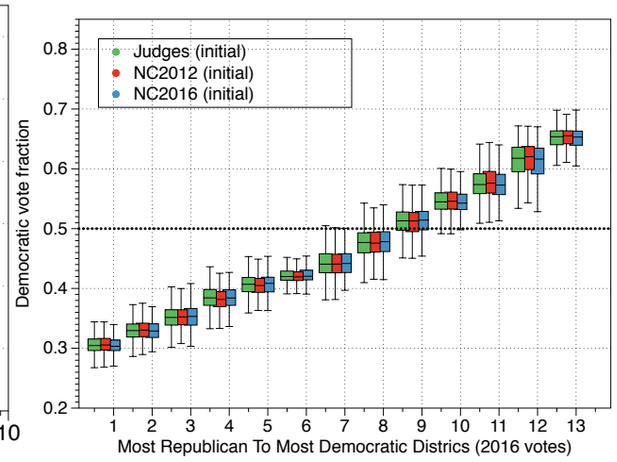

24.5×10^3 samples
119.3×10^3 samples
Fraction of result
Number of Democrats Elected (2012 votes)

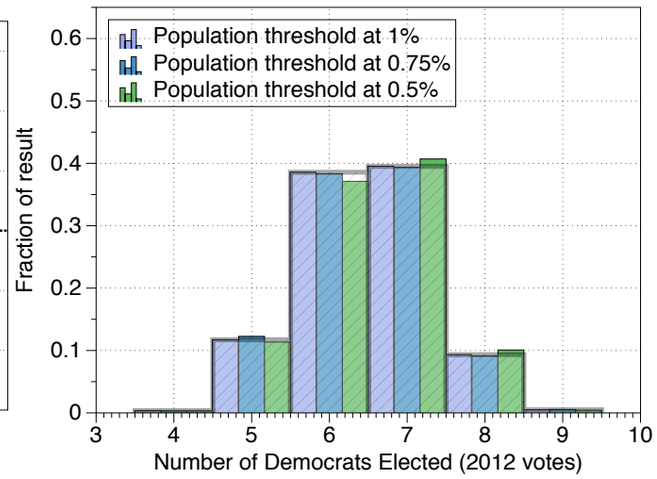

24.5×10^3 samples
119.3×10^3 samples
Democratic vote fraction
Most Republican To Most Democratic Districts (2012 votes)

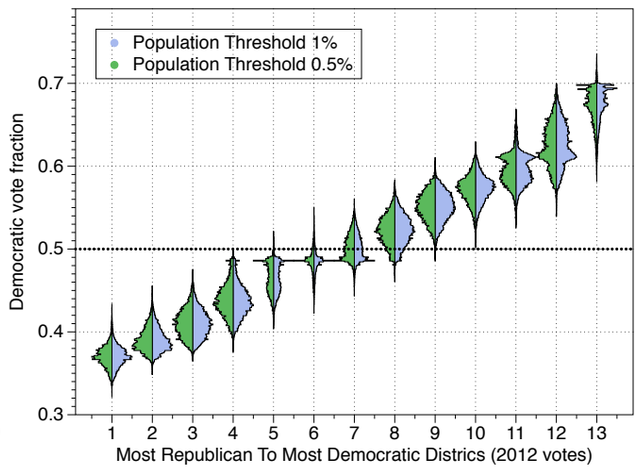

Reported S.A. parameters
Doubled S.A. parameters
Fraction of result
Number of Democrats Elected (2012 votes)

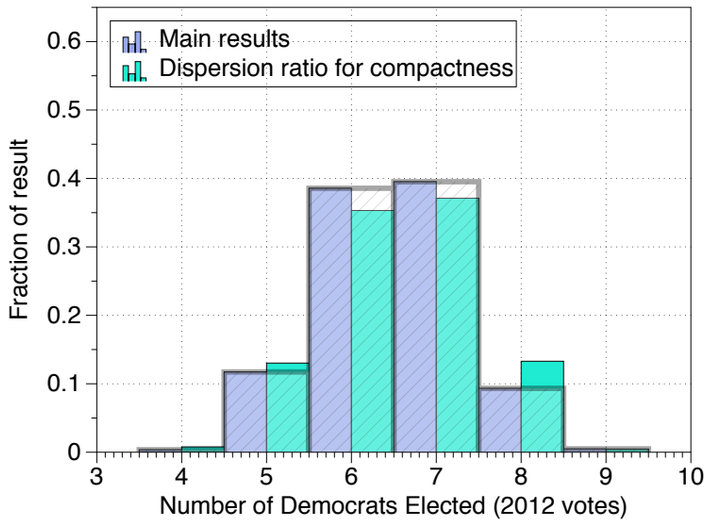

Reported S.A. parameters
Doubled S.A. parameters
Democratic vote fraction
Most Republican To Most Democratic Districts (2016 votes)

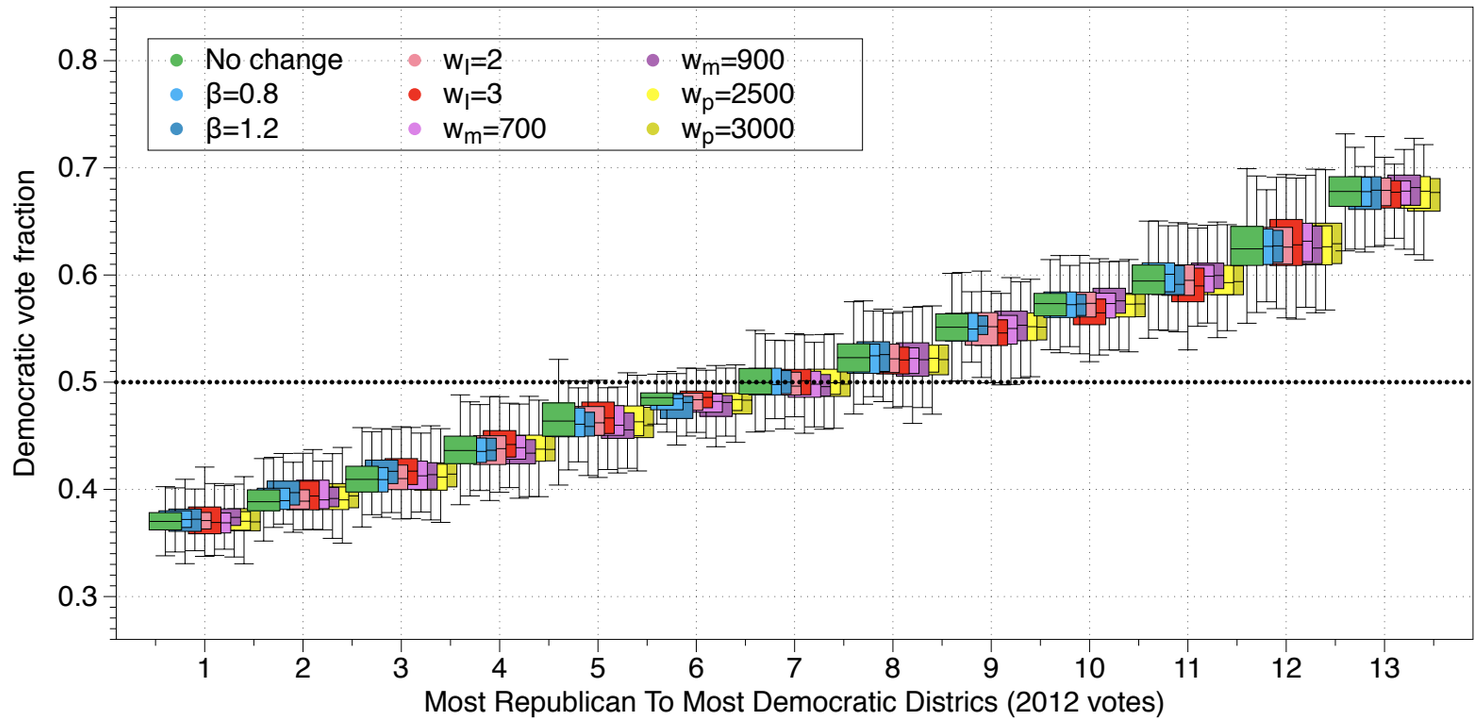

Judges (initial)
NC2012 (initial)
NC2016 (initial)
Fraction of result
Number of Democrats Elected (2012 votes)


Judges (initial)
NC2012 (initial)
NC2016 (initial)
Democratic vote fraction
Most Republican To Most Democratic Districts (2016 votes)

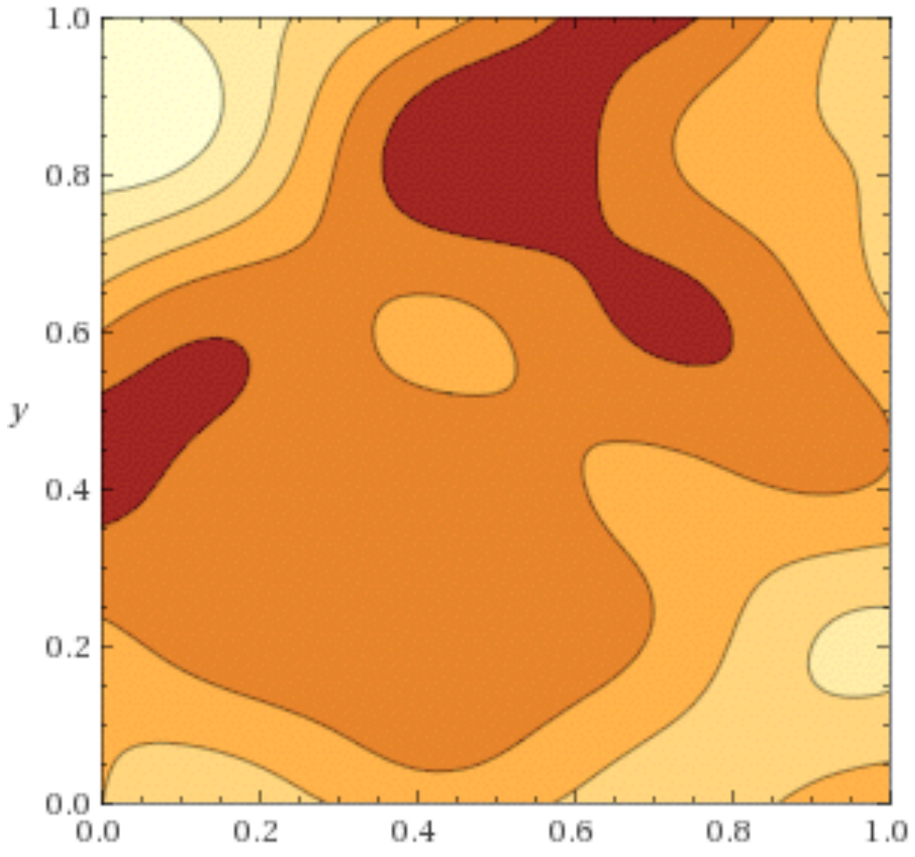

Population threshold at 1%
Population threshold at 0.75%
Population threshold at 0.5%
Fraction of result
Number of Democrats Elected (2012 votes)


Population Threshold 1%
Population Threshold 0.5%
Democratic vote fraction
Most Republican To Most Democratic Districts (2012 votes)


Main results
Dispersion ratio for compactness
Fraction of result
Number of Democrats Elected (2012 votes)


No change
β=0.8
β=1.2
w_l=2
w_l=3
w_m=700
w_m=900
w_p=2500
w_p=3000
Democratic vote fraction
Most Republican To Most Democratic Districts (2012 votes)

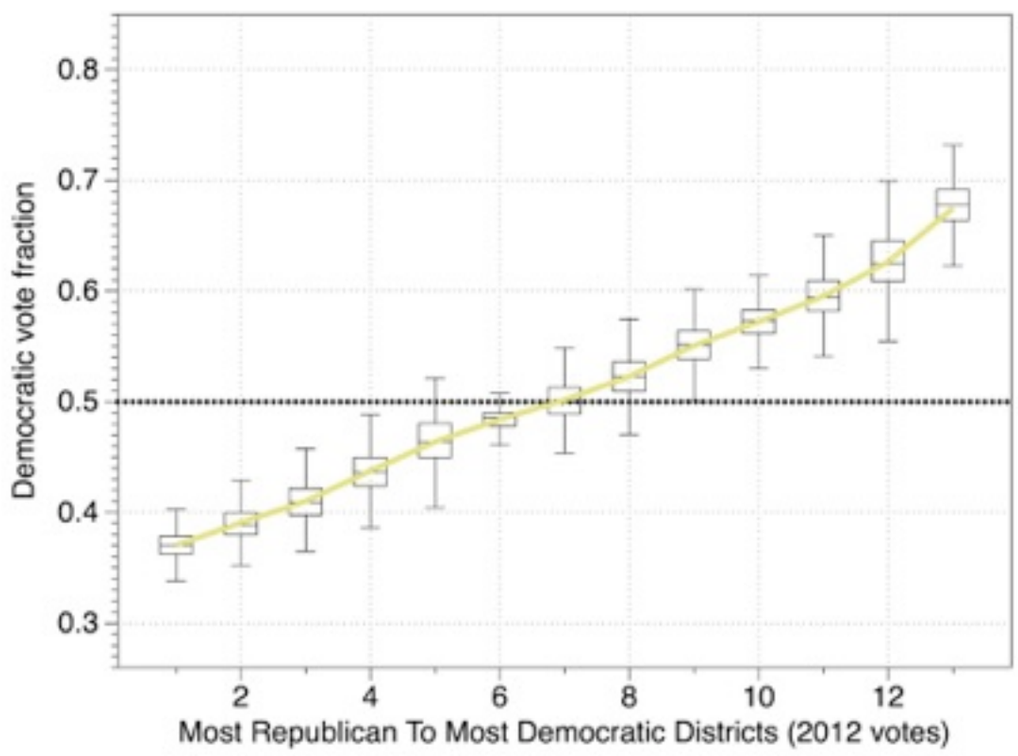
Math Questions ?

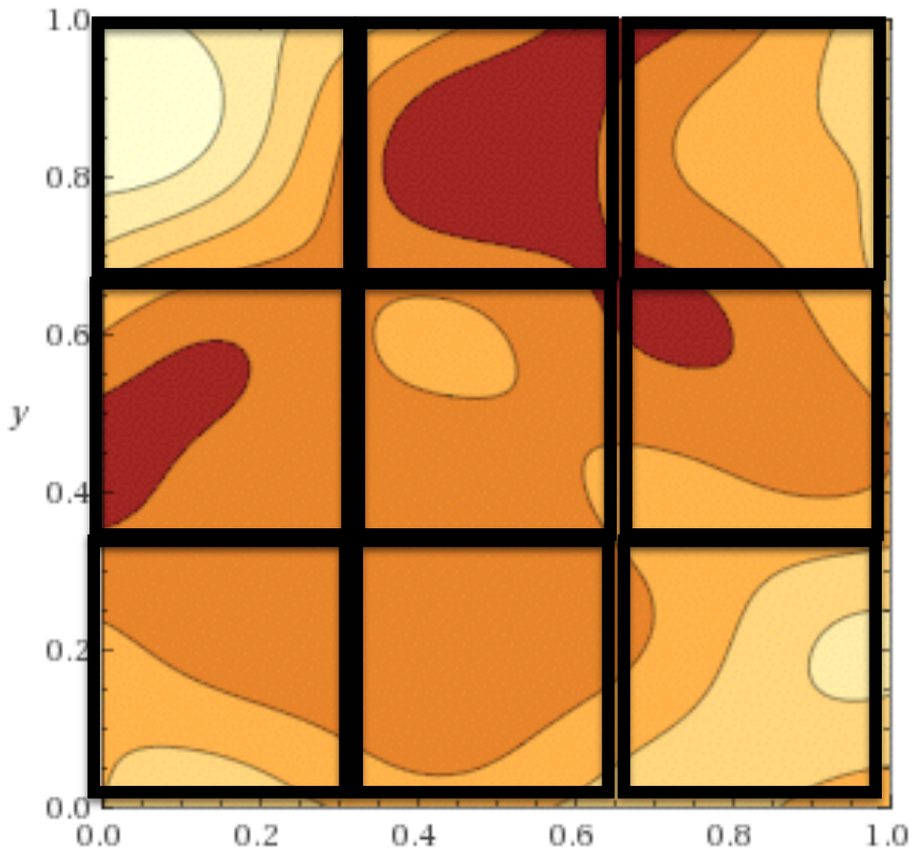


Assume the population is uniform

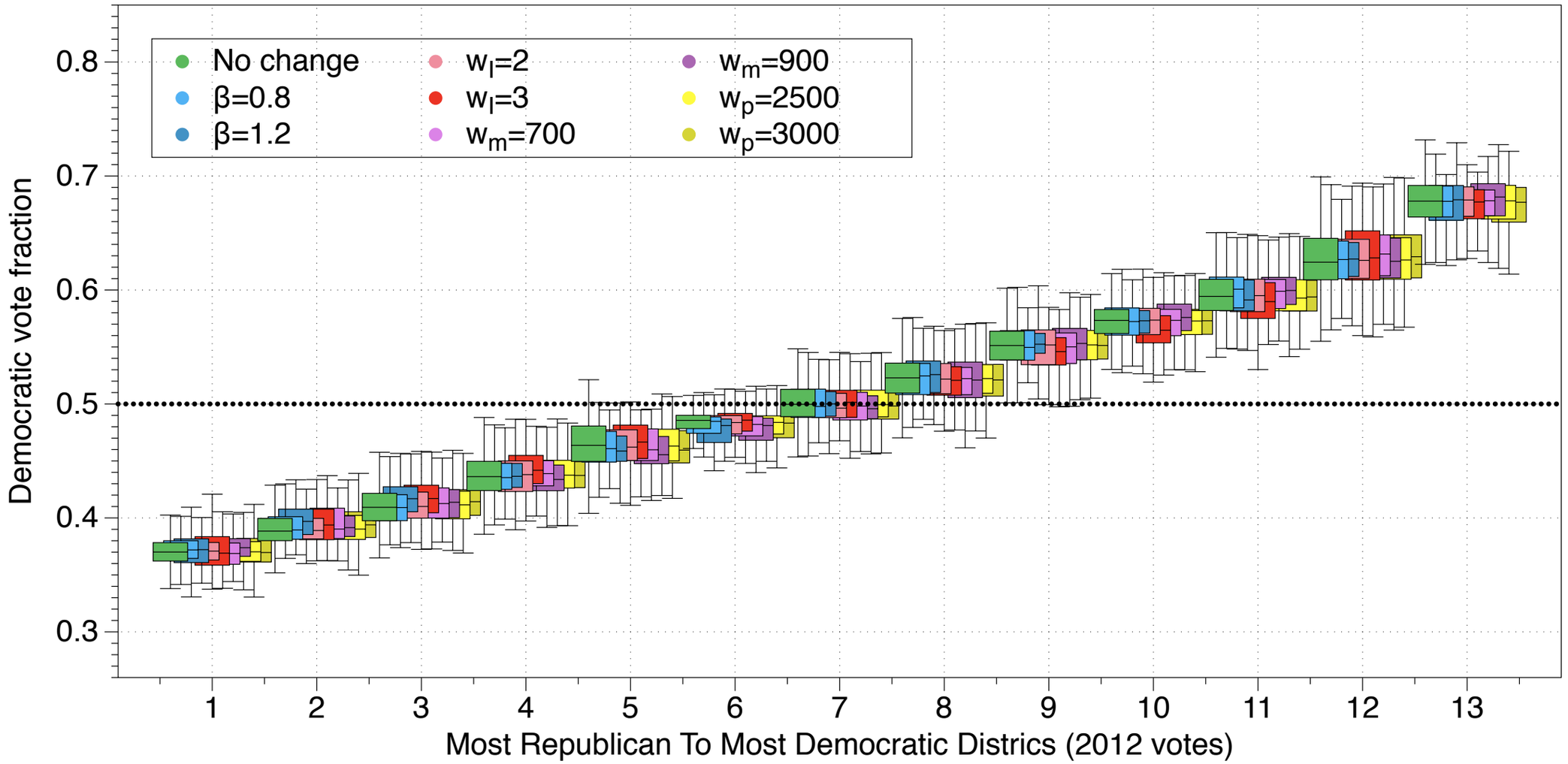
model a random distribution of political parties

Q: Find null distribution of order statistics for district make up





Q: Give some form of stability of plots over a class of energy functions which have certain marginal statistics.



Q: Characterize the structure of the energy landscape

Even with just population and compactness

Accelerate the sampling

- parallel tempering
- accelerated sampling
- hierarchical sampling
- parallel algorithms

The Team

Jonathan Mattingly

Christy Graves

Sachet Bangia

Sophie Guo

Bridget Dou

Justin Luo

Hansung Kang

Robert Ravier

Greg Herschlag

Michael Bell

Duke UNIVERSITY MATH · Data+ · iiD · Duke UNIVERSITY | POLIS

arXiv:1410.8796 arXiv:1709.01596 arXiv:1704.03360

Conclusions

Identify the background NULL hypothesis

Identify the background geopolitical structure of state.

Identify the outliers, the unreasonable maps.

Benchmark different proposed metrics

Interaction of geopolitical structure and metrics

Considered other states and effect of VRA

Lots of interesting math questions