Math For Unbiased Maps TX (MUM_TX)

Using Math and Computation to ensure a fair and transparent redistricting process

Andrea K. Barreiro, Associate Professor of Mathematics,
Southern Methodist University (SMU)

Other faculty from Texas universities
and interested community volunteers

MUM_TX is a project of the
Research Cluster on Political Decision-Making,
supported by the SMU Dedman College Interdisciplinary Insitutute

https://www.smu.edu/Dedman/Research/Institutes-and-Centers/DCII/Scholarship/Research-Cluster-on-Political-Decision-Making

Contact us:
MathForUnbiasedMapsTX@list.smu.edu

Members

Southern Methodist University (SMU)

Andrea K. Barreiro, Associate Professor of Mathematics
Matthew Lockard, Associate Professor of Philosophy
Scott Norris, Associate Professor of Mathematics
Brandilyn Stigler, Associate Professor of Mathematics

Collin College

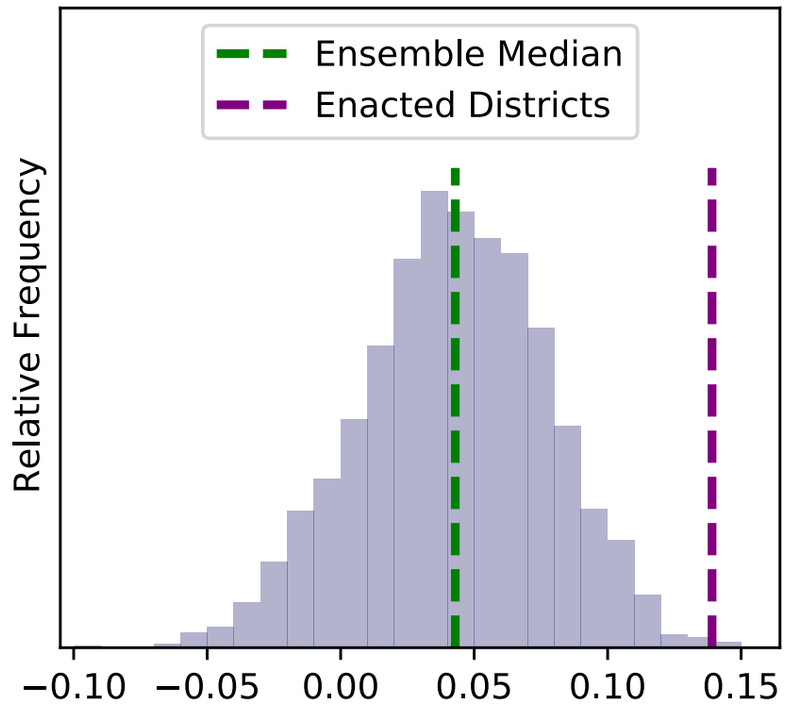
Dustin Potter, Professor of Mathematics

Community Members/Volunteers

Robert Meyers

What is ensemble sampling?

- We use *Markov Chain Monte Carlo* methods to generate a large number of legal plans.
- This provides an ***unbiased baseline*** for what a plan should look like.
- We can assess each plan on the measure(s) of interest such as:
  - partisan asymmetry metrics
  - VRA-relevant metrics
- Create a histogram to illustrate the distribution of outcomes.



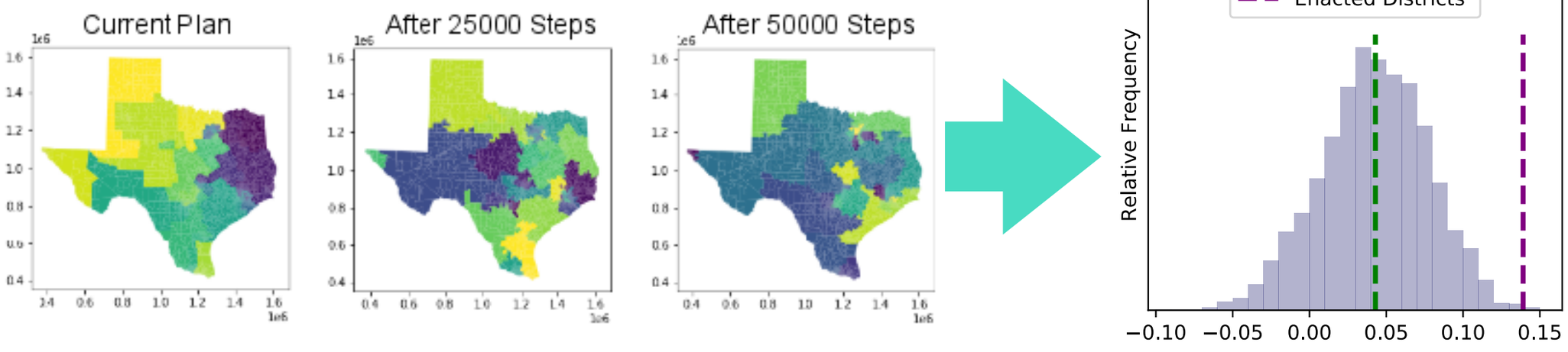
How can ensemble sampling help you assess/develop a plan?

- Compare the ensemble outcomes with the actual plan (or a proposed plan)
- ***If the outcome for the proposed plan is an outlier, this is evidence of gerrymandering***

What we will show you today

- Voting behavior in current maps using 2010 Census data
- New ensembles with ***2020 Census data***
  - Voting behavior in using the 2020 ensemble
- Using demographic data from your ensemble to help ensure compliance with the Voting Rights Act
- Benefits of ensemble sampling: it's fast, reliable, and validated.

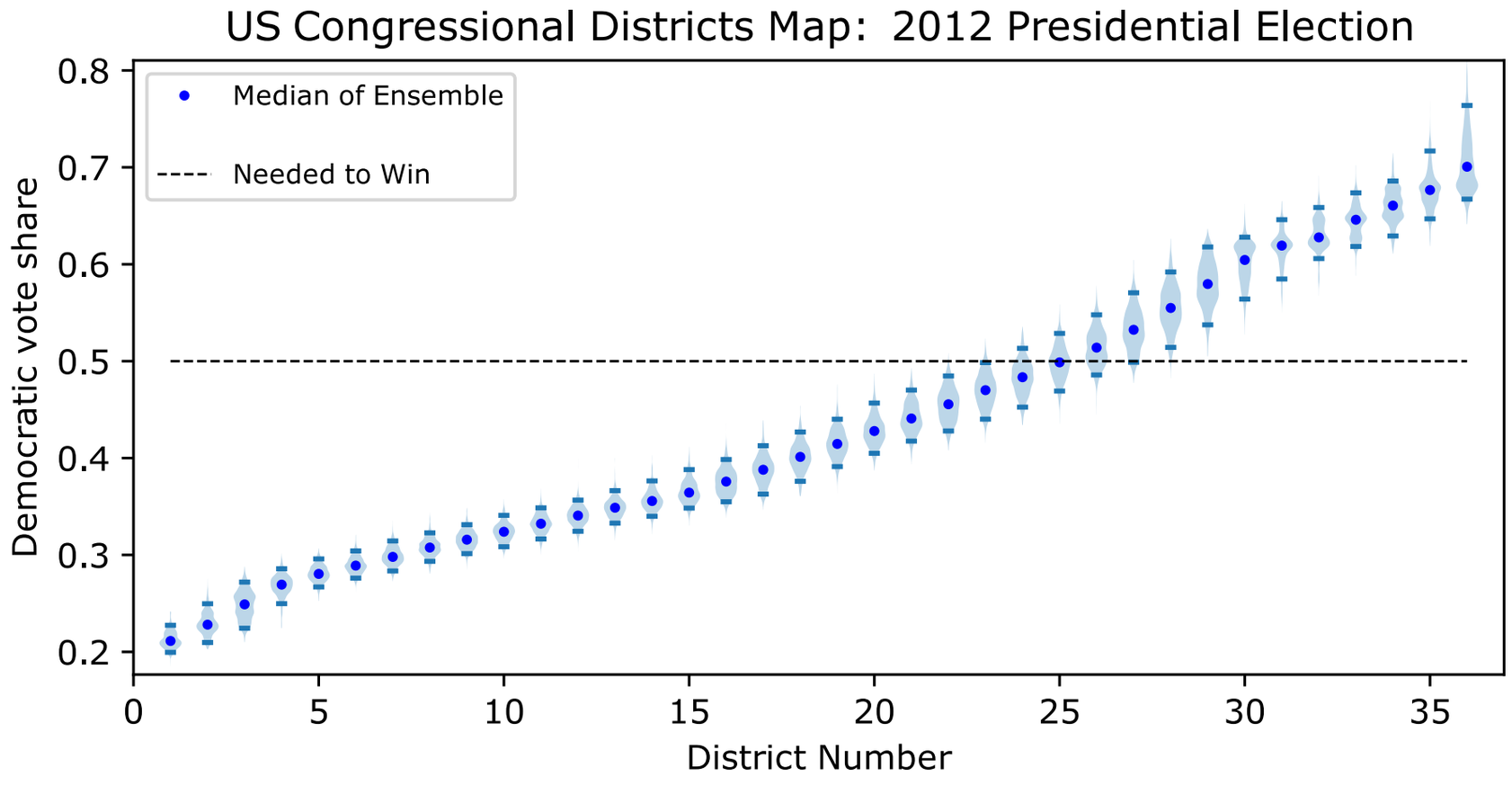
"MUM_TX: Make your maps fair, and we'll get out of your hair!"



What can we learn from an ensemble about voting behavior?

Vote Share Curves

- For each map, order districts by increasing vote share
- **Blue:** ensemble

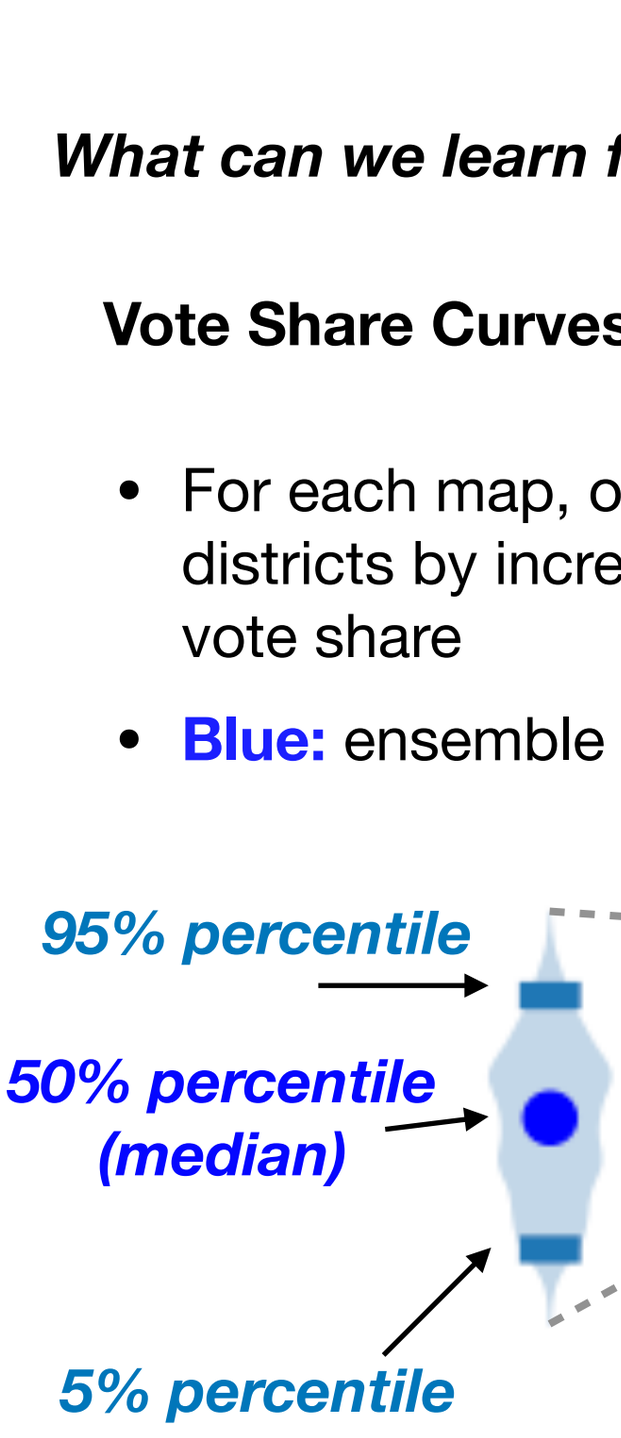


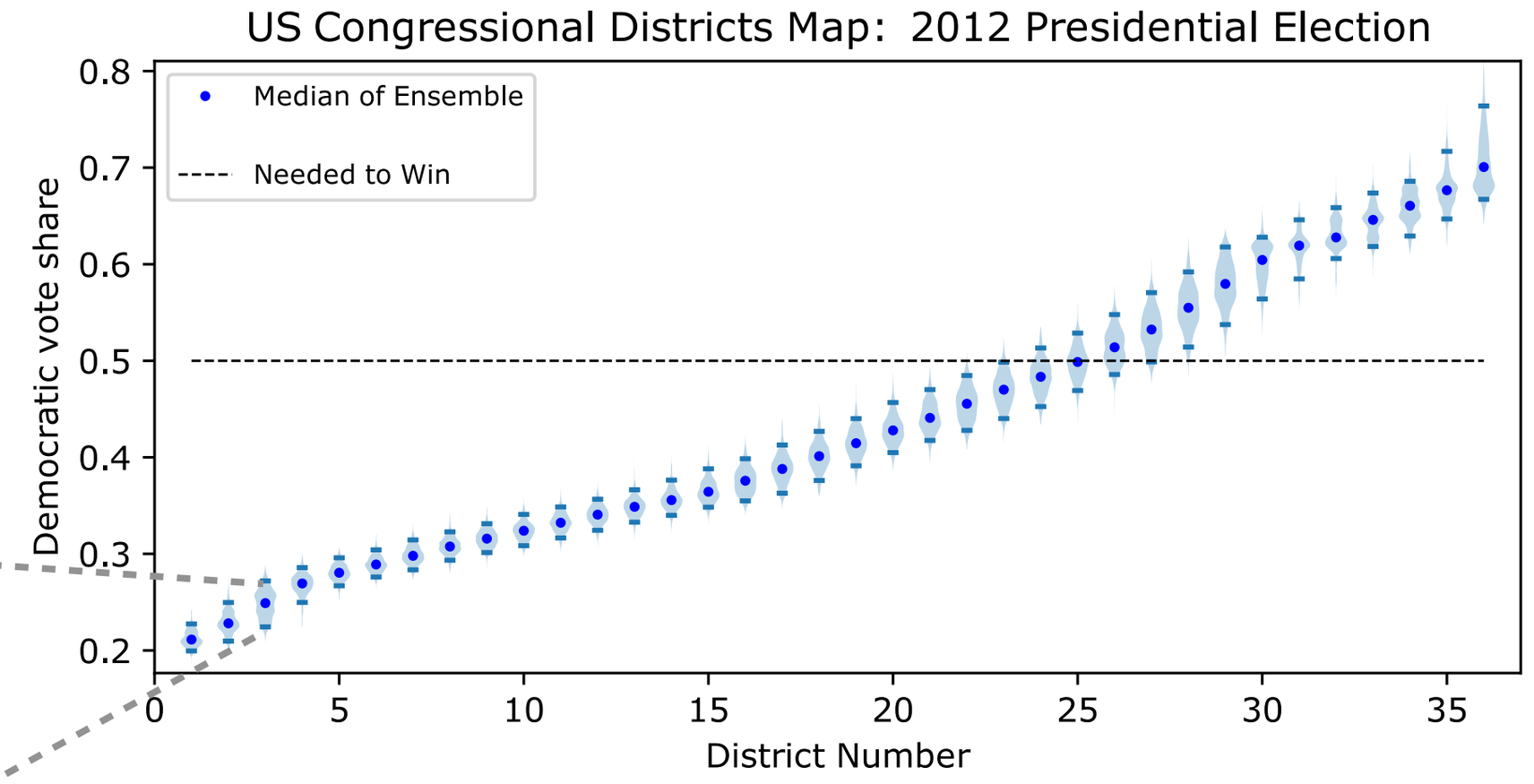
US Congressional Districts (36)
2010 Census Data

What can we learn from an ensemble about voting behavior?

Vote Share Curves

- For each map, order districts by increasing vote share
- **Blue:** ensemble





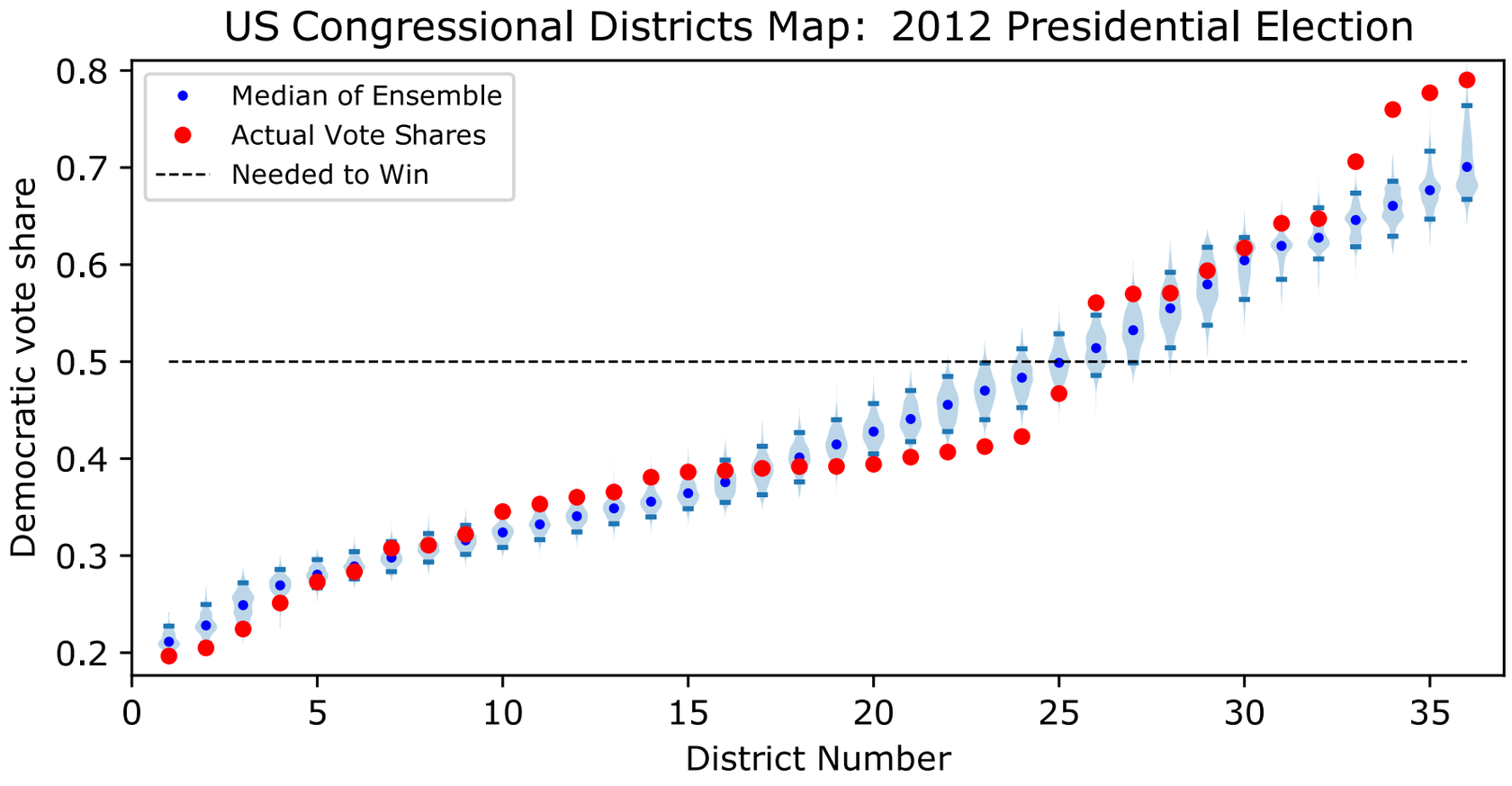
Violin plots give us a convenient way to illustrate the entire distribution of a statistic (here, the *3rd smallest Democratic* vote share)

US Congressional Districts (36)
2010 Census Data

What can we learn from an ensemble about voting behavior?

Vote Share Curves

- For each map, order districts by increasing vote share
- **Blue:** ensemble
- **Red:** enacted plan (or any comparison plan of interest)

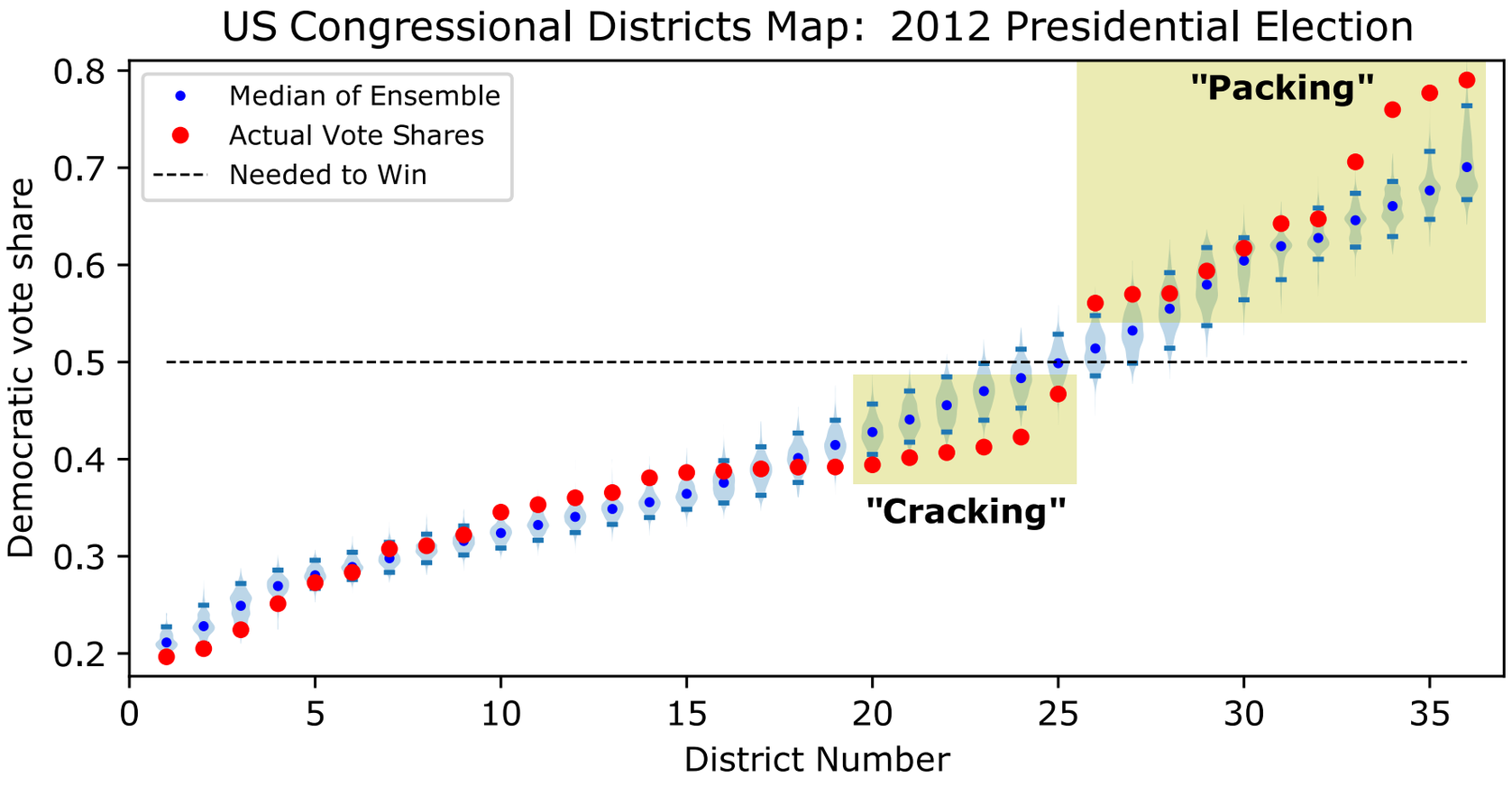


US Congressional Districts (36)
2010 Census Data

Vote share curves, when compared with an unbiased sample, can show distinct signatures of partisan gerrymandering.

Vote Share Curves

- For each map, order districts by increasing vote share
- **Blue:** ensemble
- **Red:** enacted plan (or any comparison plan of interest)



- ***Cracking:*** spreading the opposing party's voters across multiple districts
- ***Packing:*** concentrating the opposing party's voters into a few safe districts

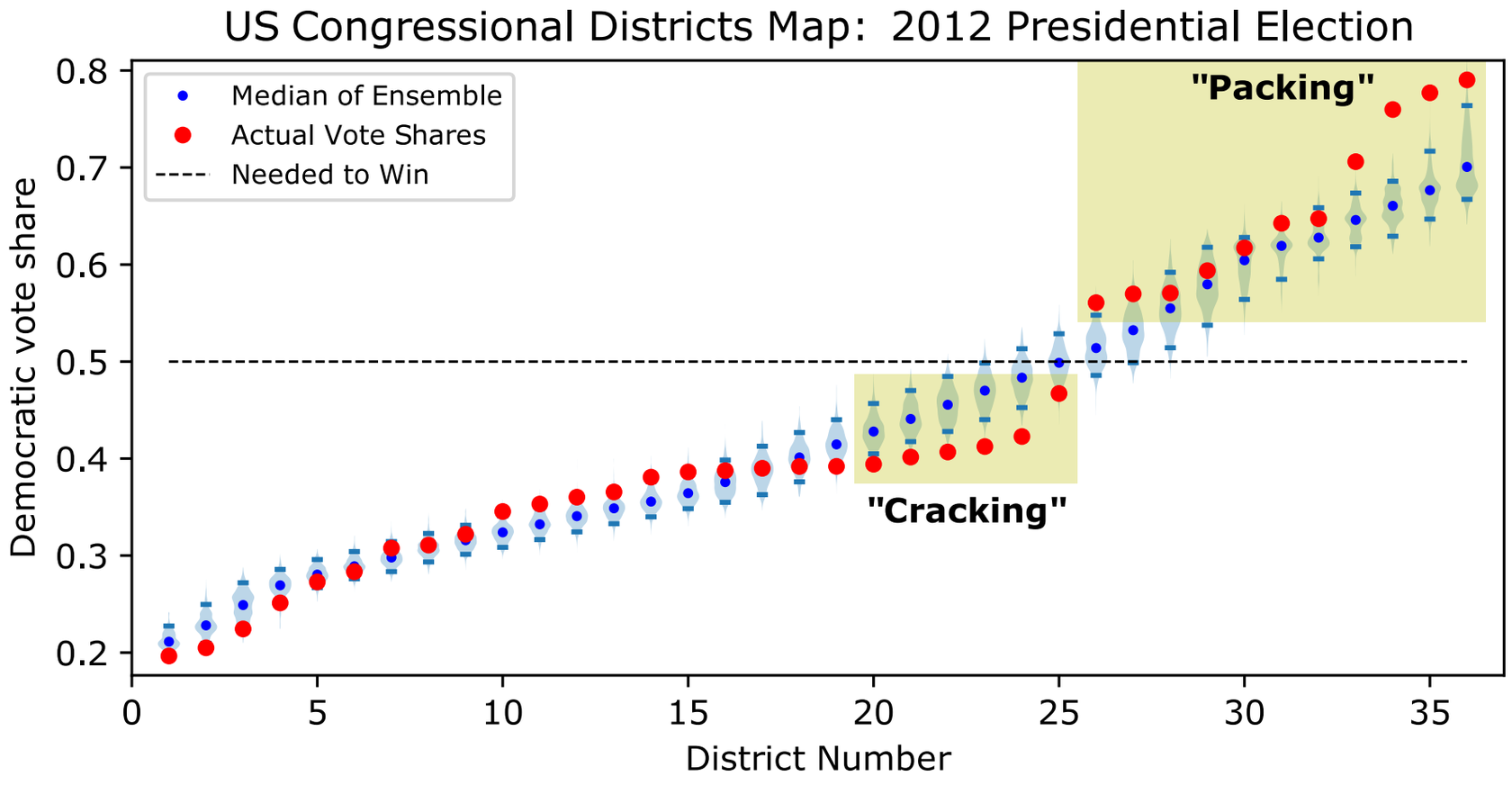
US Congressional Districts (36)
2010 Census Data

Vote share curves, when compared with an unbiased sample, can show distinct signatures of partisan gerrymandering.

Two ways we can quantify this partisan skew:

- Number of seats won at 50% vote share
  - Democrats: 12
  - Republicans: 24
  - ***Difference: -12***

- Vote share needed for majority (18 seats)
  - Democrats: 55%
  - Republicans: 45%
  - ***Difference: 10%***



US Congressional Districts (36)
2010 Census Data

Partisan Bias and the Mean-Median score are widely accepted by political scientists and are easy to compute

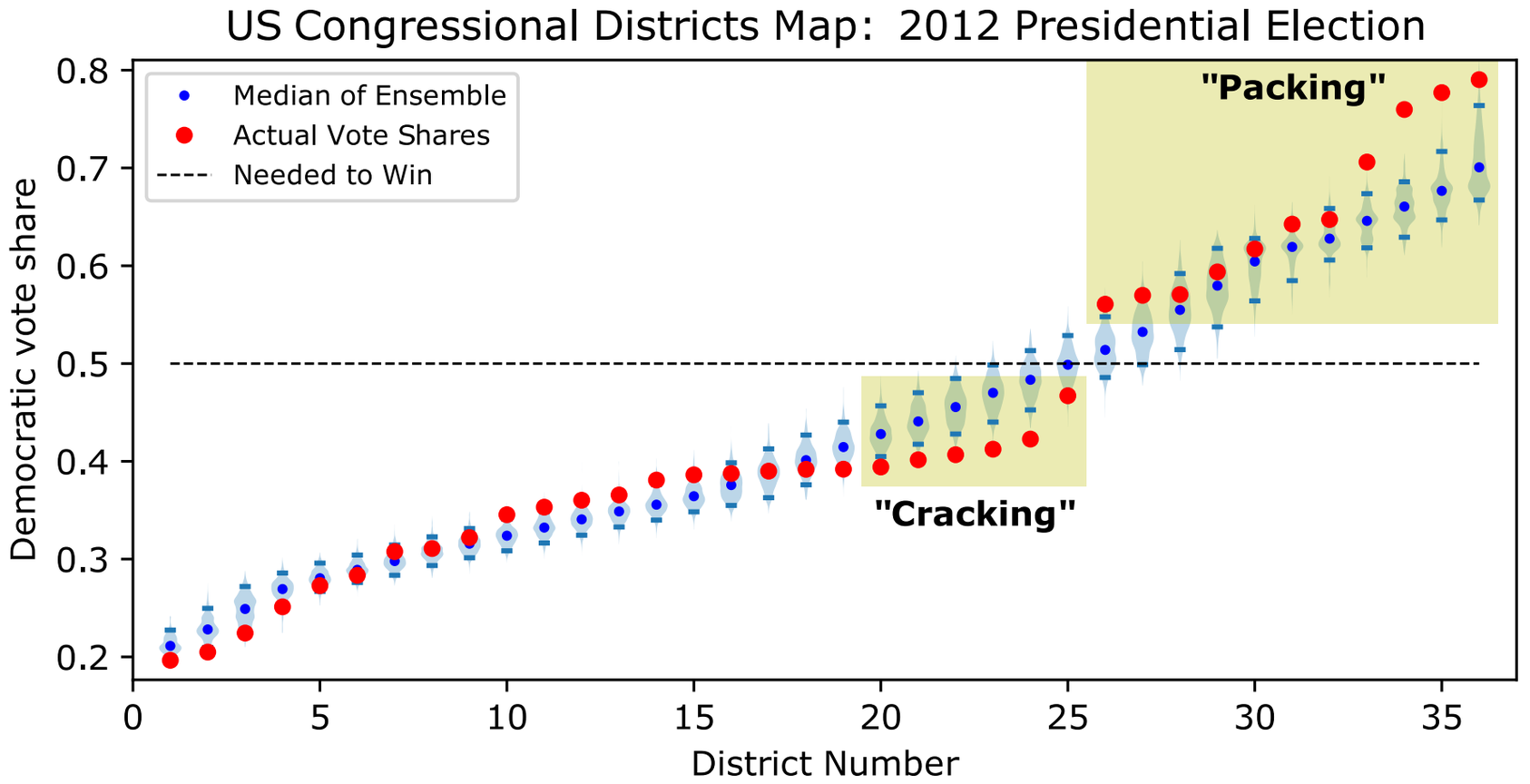
Two ways we can quantify this partisan skew:

- Number of seats won at 50% vote share
  - Democrats: 12
  - Republicans: 24
  - ***Difference: -12***

Partisan Bias

- Vote share needed for majority (18 seats)
  - Democrats: 55%
  - Republicans: 45%
  - ***Difference: 10%***

Mean-Median Score



US Congressional Districts (36)

2010 Census Data

Comparing with the ensemble shows the enacted plan is a clear outlier

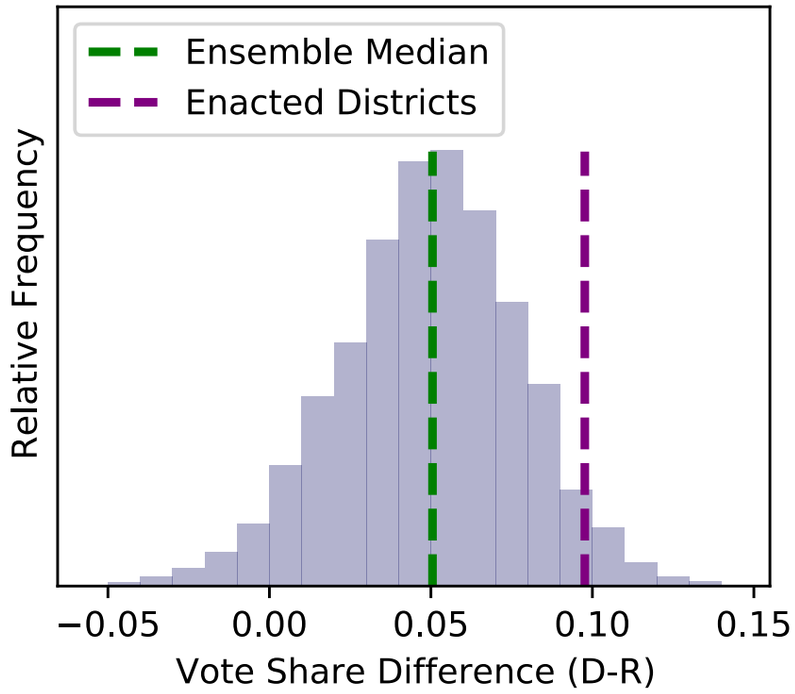
Two ways we can quantify this partisan skew:

- Number of seats won at 50% vote share
  - Democrats: 12
  - Republicans: 24
  - ***Difference: -12***

Partisan Bias

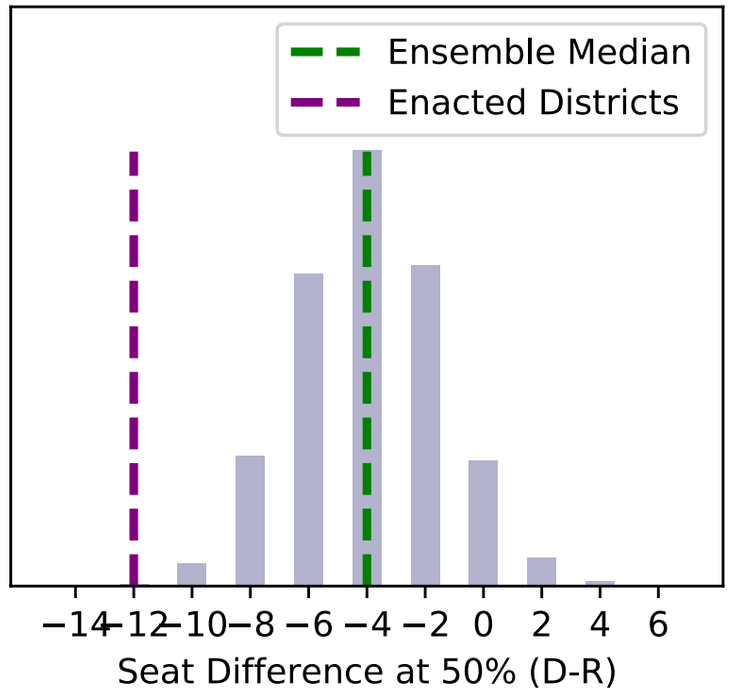
- Vote share needed for majority (18 seats)
  - Democrats: 55%
  - Republicans: 45%
  - ***Difference: 10%***

Mean-Median Score



Average: 5% (vs 10%)

Only 1 in 25 plans show this level of disparity



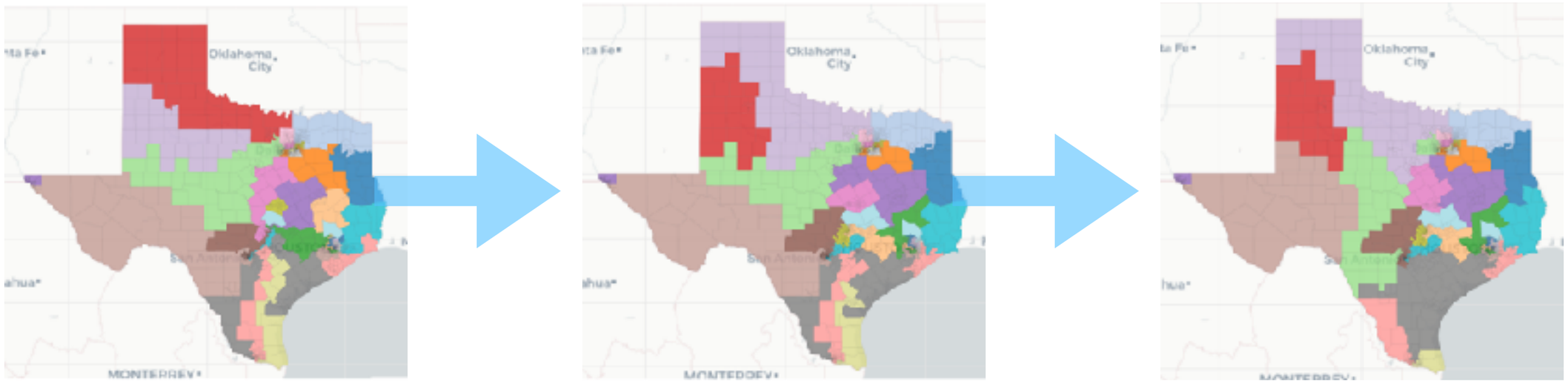
Average: -4 (vs. -12)

Only 1 in 800 plans show this level of disparity

US Congressional Districts (36)
2010 Census Data

What about 2020?

- We don’t have an “enacted" plan with 38 districts
- However, we can compute an ensemble of 38-district, population-balanced plans to anticipate what a typical plan should look like

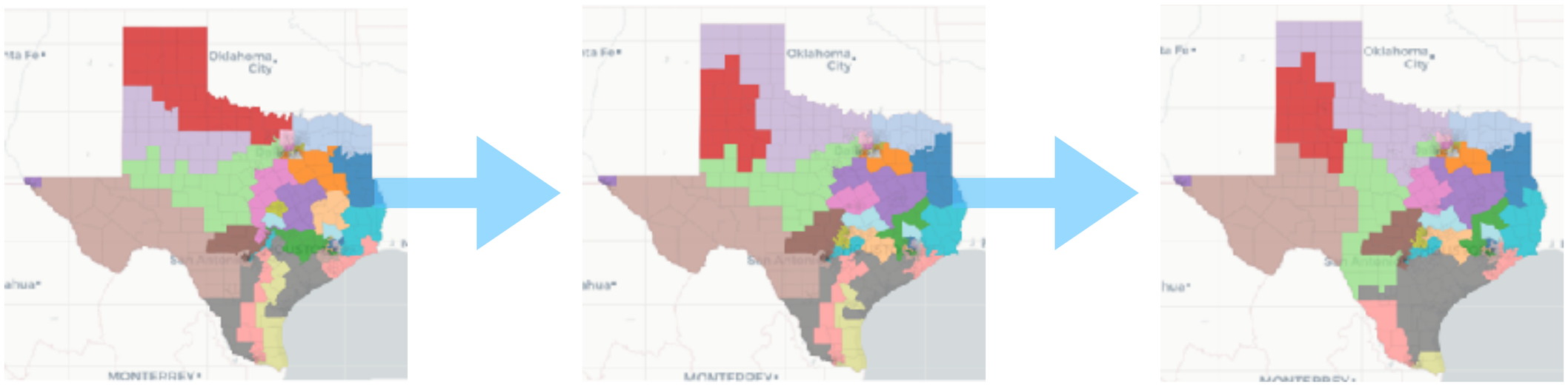


What about 2020?

- We don’t have an “enacted" plan with 38 districts
- However, we can compute an ensemble of 38-district, population-balanced plans to anticipate what a typical plan should look like

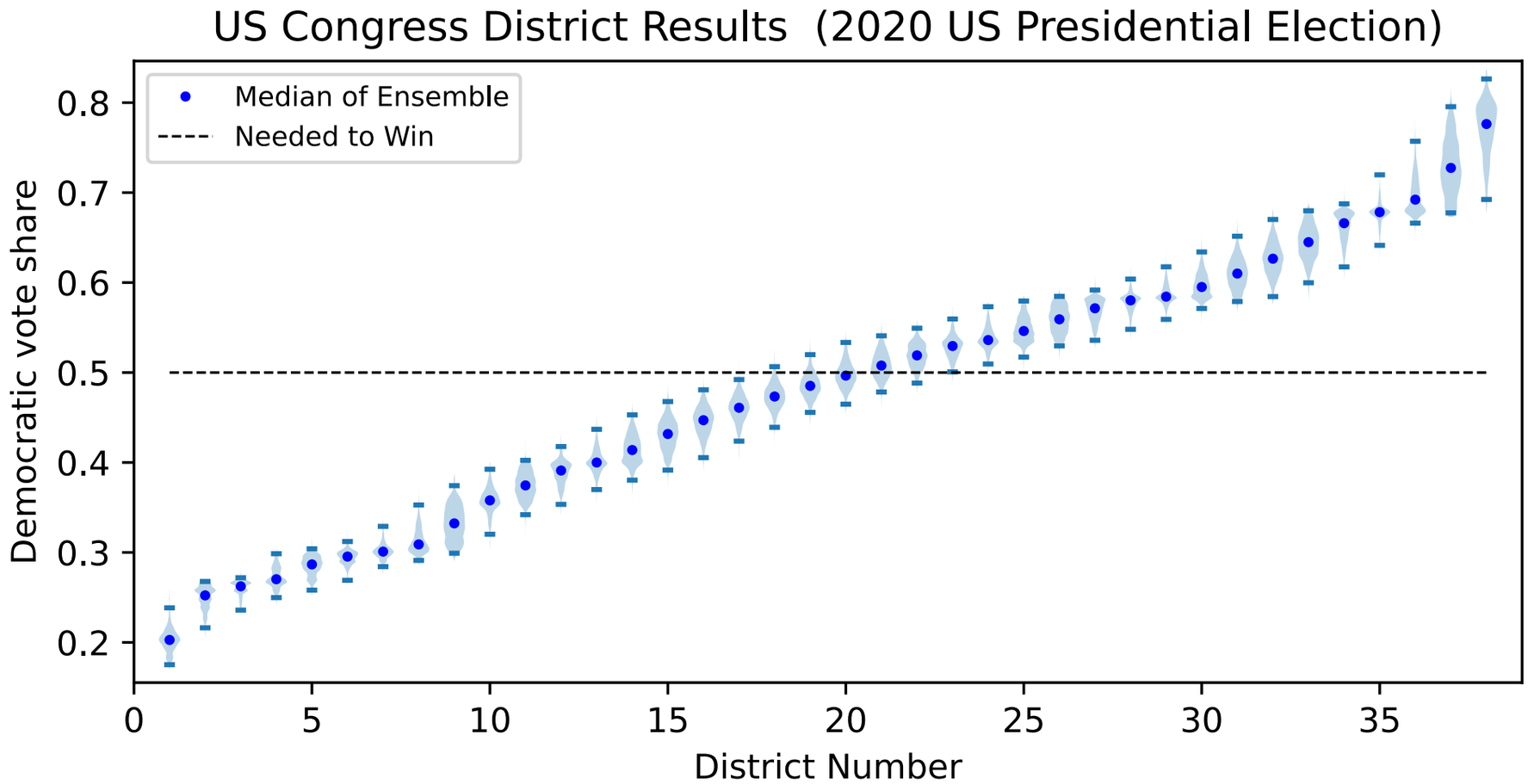
How we created our ensemble

- We first created a valid seed plan with 38 districts (vs. 36) by creating "mini-districts" in the current plan and driving it to population balance with MCMC.
- We then created an ensemble of US Congressional District plans (400,000 total, 100,000 plans included in graphics).
- We used population data from the 2020 Census and vote data from the Texas Legislative Council (2020 US Senate and Pres. elections)



What about 2020?

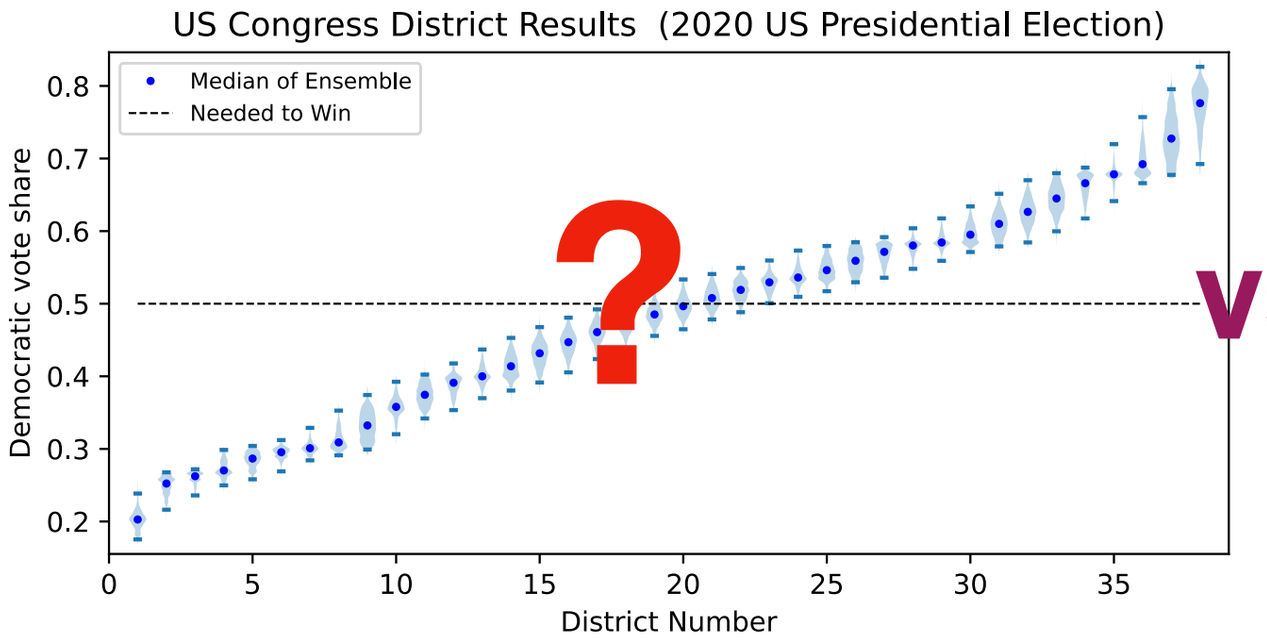
- We computed an ensemble of 38-district, population-balanced plans to anticipate what a typical plan should look like



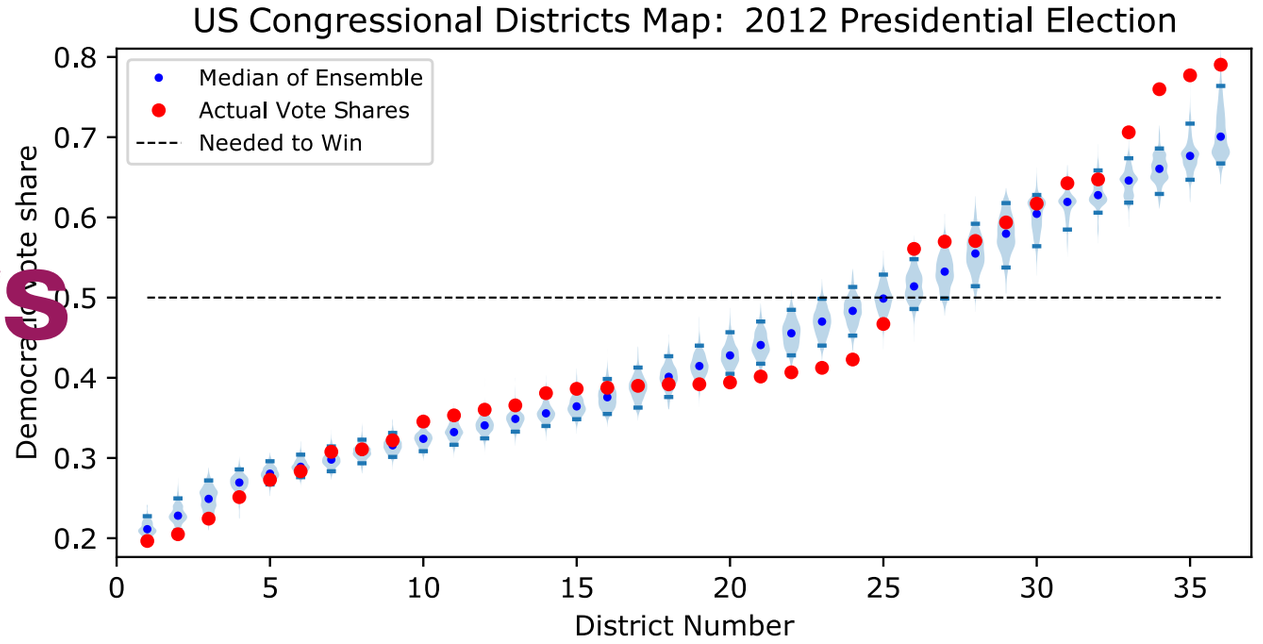
US Congressional Districts (38)
2020 Census Data

What about 2020?

- We computed an ensemble of 38-district, population-balanced plans to anticipate what a typical plan should look like
- However, we don't yet have an "enacted" plan with 38 districts
- ***When PLANC### is released, its statistics can be quickly computed and superimposed for comparison.***



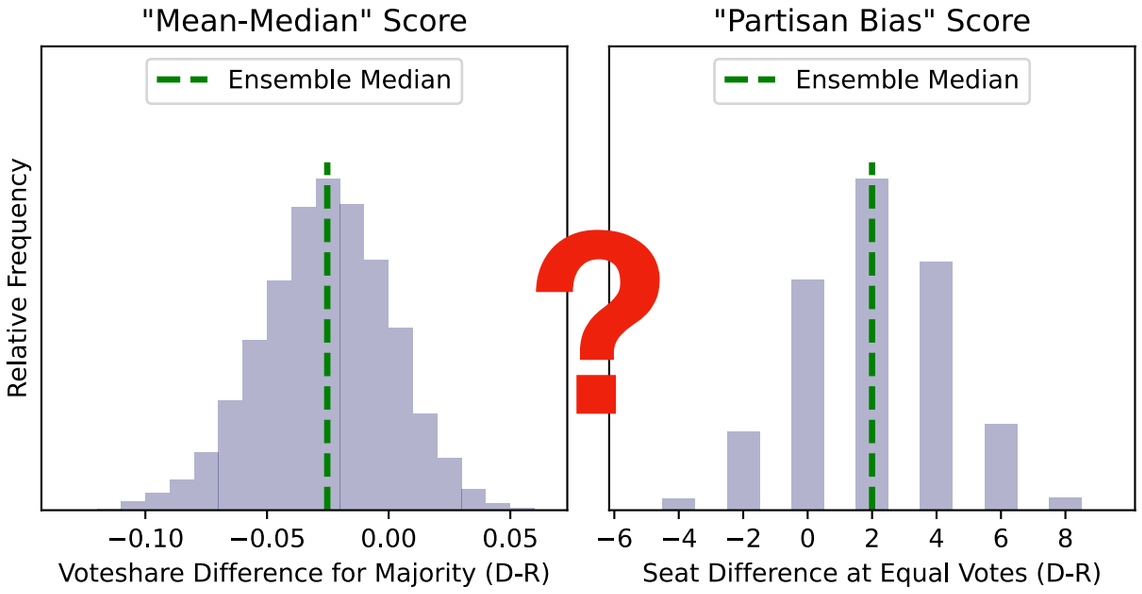
vs



US Congressional Districts (38)
2020 Census Data

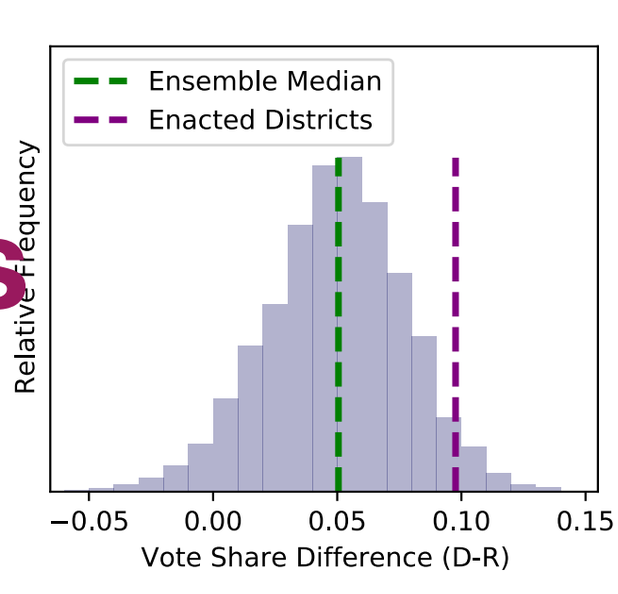
What about 2020?

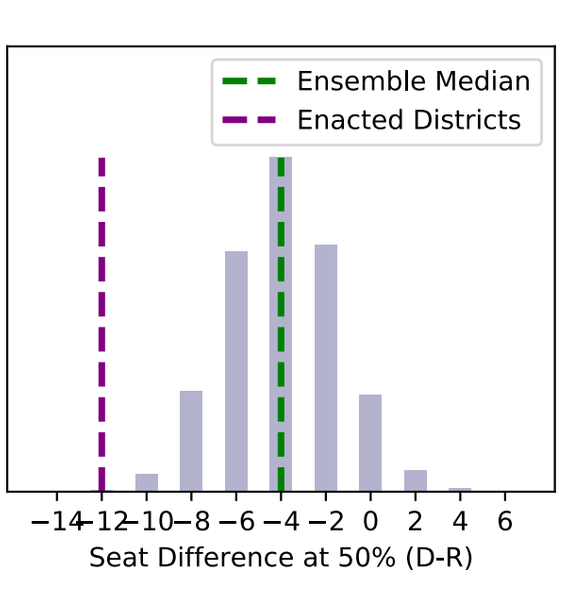
- We computed an ensemble of 38-district, population-balanced plans to anticipate what a typical plan should look like
- However, we don’t yet have an “enacted" plan with 38 districts
- ***When PLANC### is released, its statistics can be quickly computed and superimposed for comparison.***



?

VS

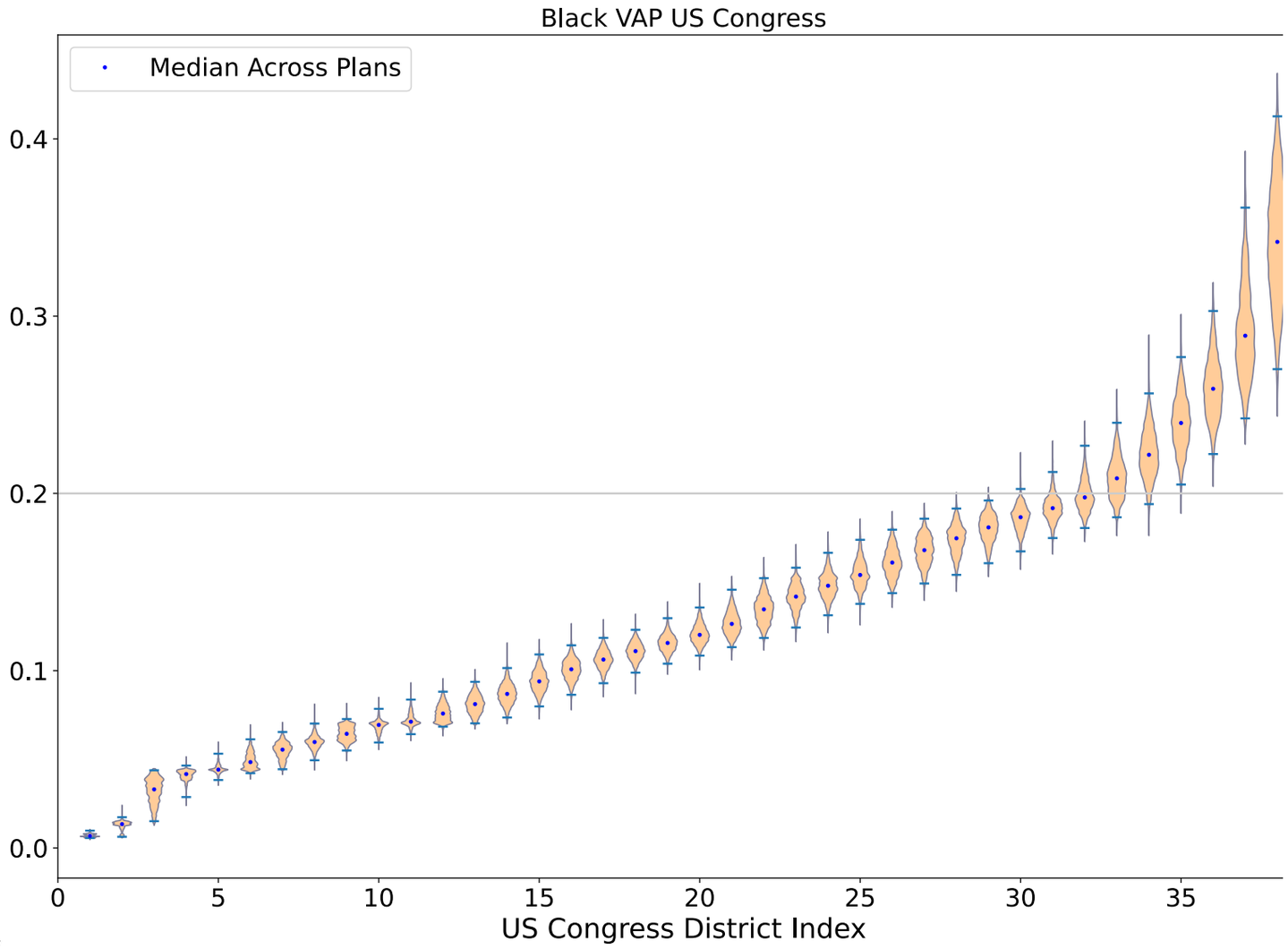




US Congressional Districts (38)
2020 Census Data

Complying with the Voting Rights Act

- For each map, order districts by ***BVAP***, HVAP, or BHVAP
- **Violin plots:** ensemble

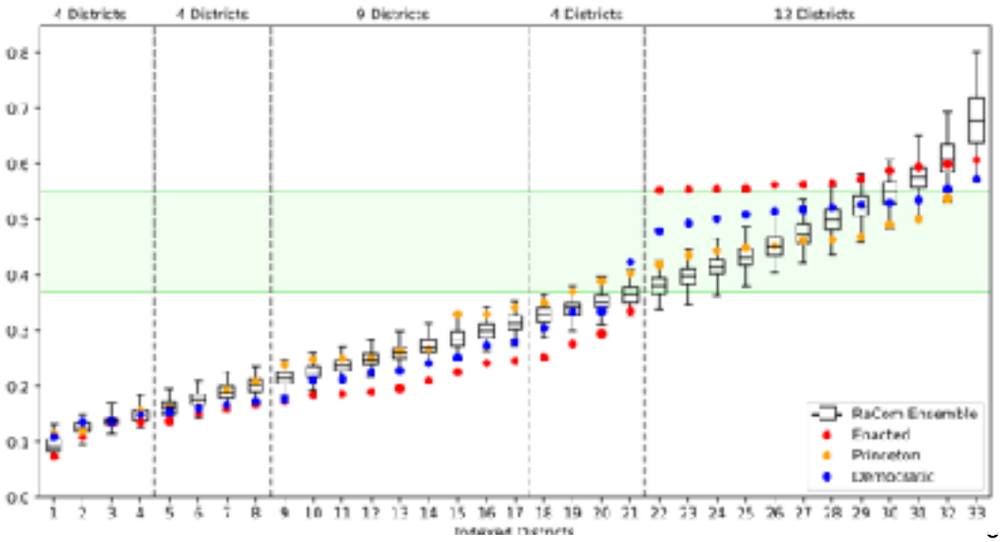


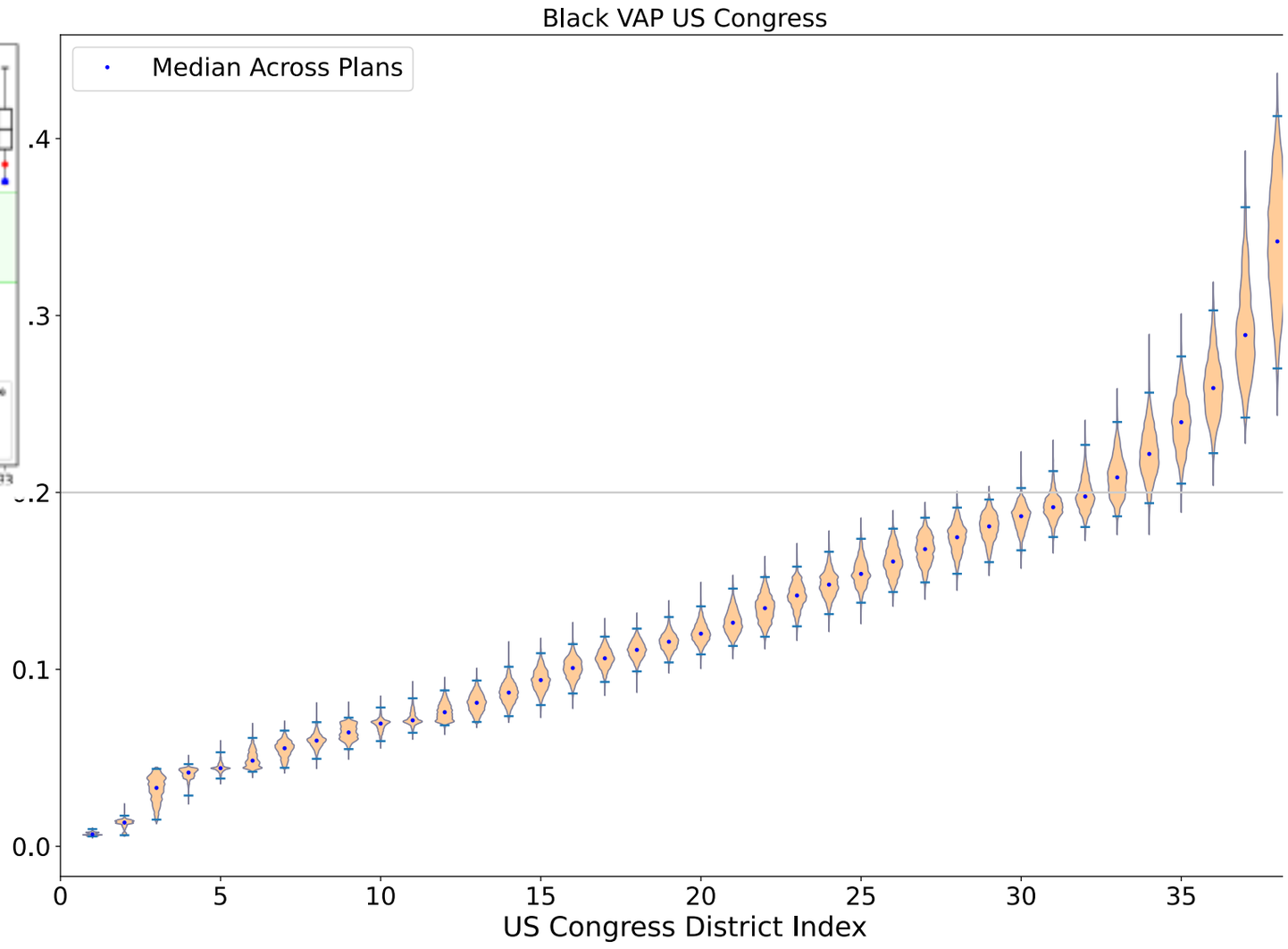
Note: it is straightforward to compute any other group or coalition of interest

HVAP= Hispanic Voting Age Population (VAP)
BVAP= Black VAP
BHVAP = Black OR Hispanic VAP

US Congressional Districts (38)
2020 Census Data

Complying with the Voting Rights Act





How could we use this information? When we have a proposed plan, outliers may show evidence of racial gerrymandering.

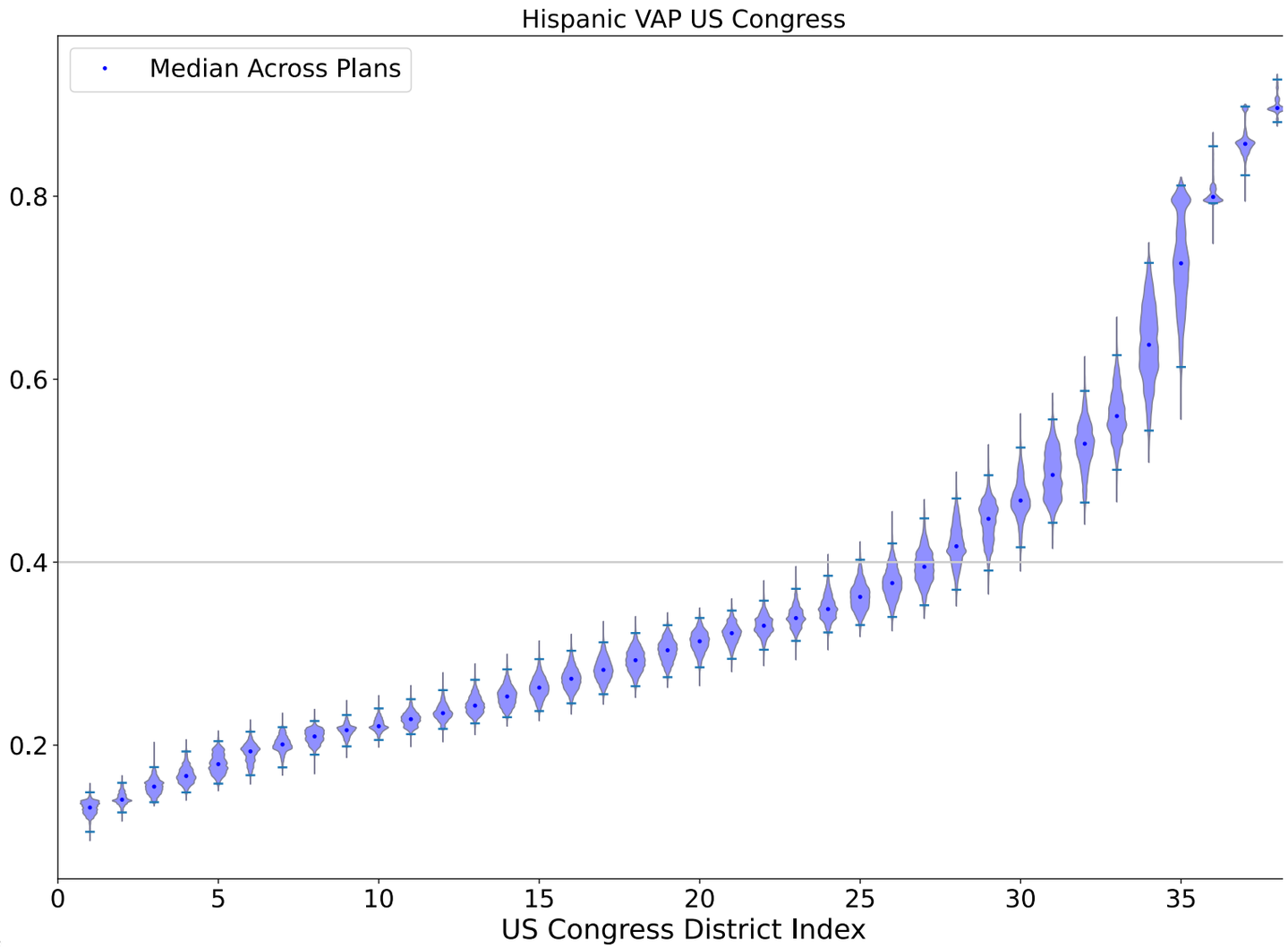
(Above: BVAP from Virginia House of Delegates report, MGGG, 2018)
https://mggg.org/VA-report.pdf

US Congressional Districts (38)
2020 Census Data

HVAP= Hispanic Voting Age Population (VAP)
BVAP= Black VAP
BHVAP = Black OR Hispanic VAP

Complying with the Voting Rights Act

- For each map, order districts by BVAP, ***HVAP***, or BHVAP
- **Violin plots:** ensemble



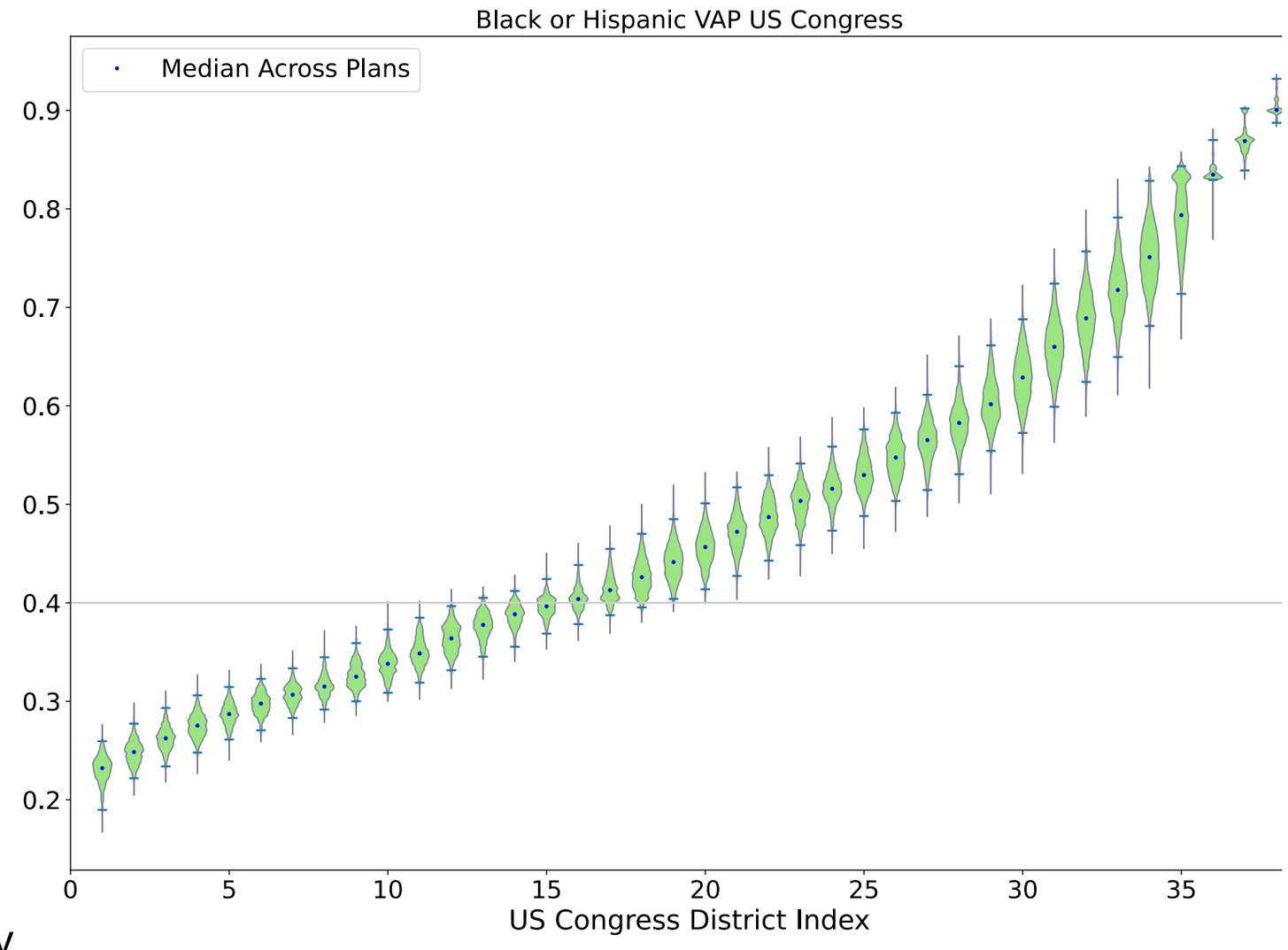
Note: it is straightforward to compute any other group or coalition of interest

HVAP= Hispanic Voting Age Population (VAP)
BVAP= Black VAP
BHVAP = Black OR Hispanic VAP

US Congressional Districts (38)
2020 Census Data

Complying with the Voting Rights Act

- For each map, order districts by HVAP, BVAP, or ***BHVAP***
- **Violin plots:** ensemble



Note: it is straightforward to compute any other group or coalition of interest

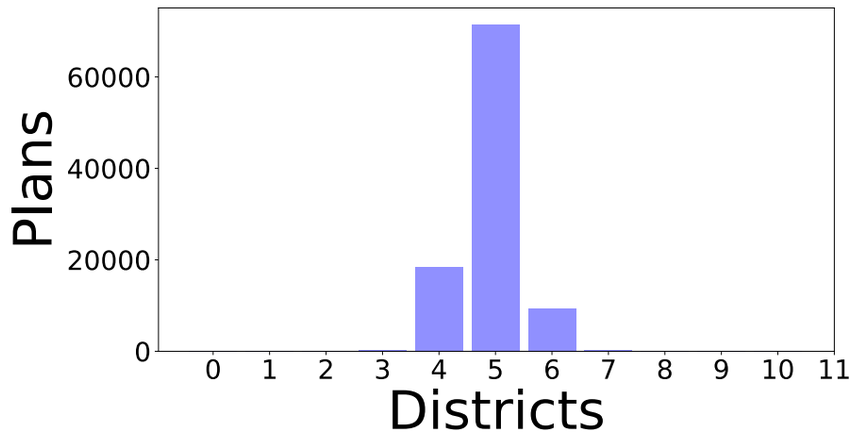
HVAP= Hispanic Voting Age Population (VAP)
BVAP= Black VAP
BHVAP = Black OR Hispanic VAP

US Congressional Districts (38)
2020 Census Data

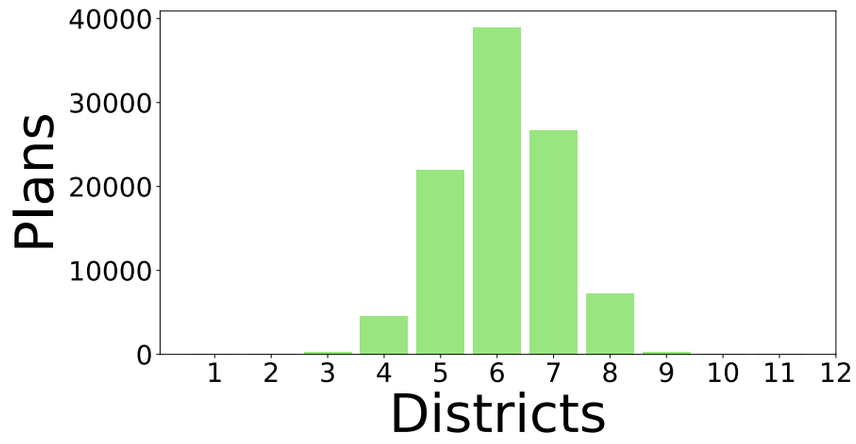
Complying with the Voting Rights Act

- How many districts have a critical threshold of group/ coalition of interest?

USCD HVAP>60%



USCD BHVAP>70%



USCD BVAP>30%

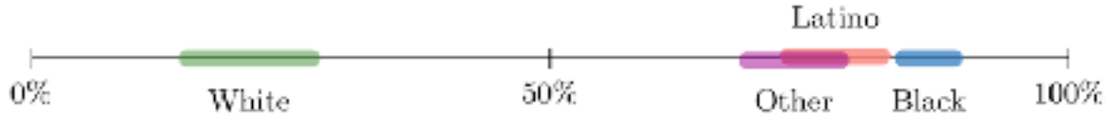
US Congressional Districts (38)

2020 Census Data

Note: it is straightforward to compute any other group or coalition of interest

Complying with the Voting Rights Act

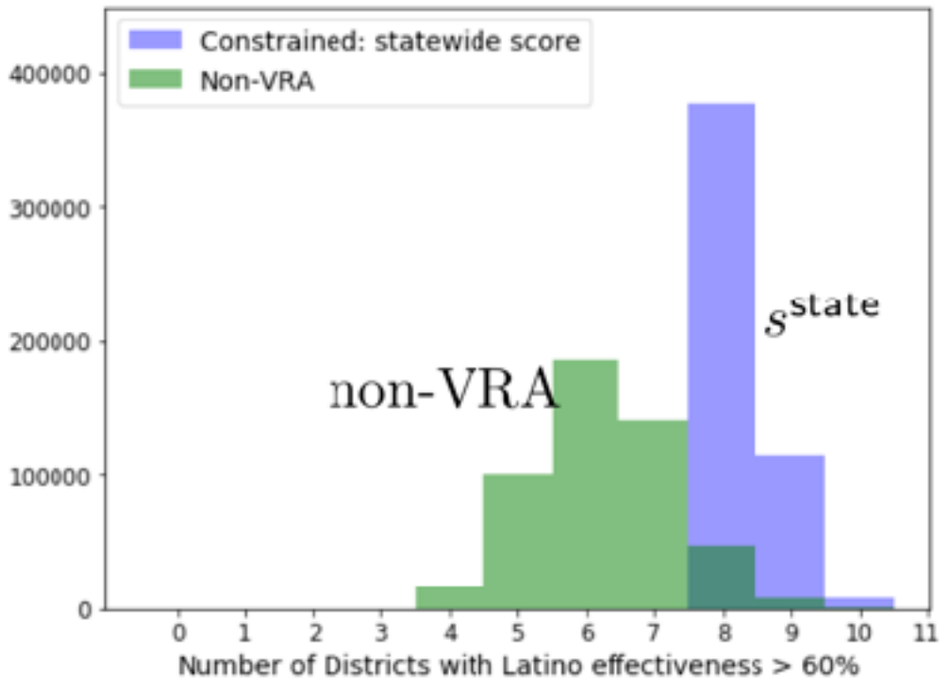
- Is demographic data the best way to ensure compliance with the Voting Rights Act?
- Maybe not! In fact, *a priori* demographic targets may violate the Equal Protection Clause.
- A new tool from the MGGG (Becker et al. 2021, https://mggg.org/vra) assigns ***effectiveness scores*** to districts based on whether minority groups have a fair chance to elect a representative of their choice
- We are incorporating this ***VRA-conscious algorithm*** into our pipeline using publicly-available computer code from MGGG.



Above: estimate of % of Democratic votes across recent statewide elections by ethnic group, showing racial polarization in voting preferences.

Below: number of districts that are effective for the Latino population, in an ensemble that uses the VRA-conscious criteria (blue) vs. one that does not (green).

(Both from Becker et al. 2021)



US Congressional Districts (38)
2020 Census Data

Ensemble sampling is reliable and replicable

- The algorithms are peer-reviewed and implemented in open source software (**GerryChain**)
- It has been used to develop plans in other states and as evidence in court cases
  - "Mathematicians' Brief" in Rucho vs. Common Cause, 2019 **(Right)**
  - League of Women Voters of Mich. v. Benson, 2019 (MI)
  - Ohio A. Philip Randolph Institute v. Householder, 2019 (OH)
  - League of Women Voters v. Commonwealth, 2018 (PA)
  - Common Cause v. Lewis, 2019 (NC)

Ensemble sampling is fast

- Computations for each set of results shown earlier took < 3 hours on a 2013 MacBook Pro

Nos. 18-422, 18-726

IN THE
Supreme Court of the United States

ROBERT A. RUCHO, ET AL.,
Appellants,
v.
COMMON CAUSE, ET AL.,
Appellees.

On Appeal from the United States District Court for the Middle District of North Carolina

LINDA H. LAMONE, ET AL.,
Appellants,
v.
O. JOHN BENISEK, ET AL.,
Appellees.

On Appeal from the United States District Court for the District of Maryland

AMICUS BRIEF OF MATHEMATICIANS, LAW PROFESSORS, AND STUDENTS IN SUPPORT OF APPELLEES AND AFFIRMANCE

Case Study: Redistricting in Pennsylvania

- **June 2017:** League of Women Voters challenges PA congressional map
- **November 2017:** Wes Pegden (Carnegie Mellon Univ mathematician) develops MCMC (**M**arkov **C**hain **M**onte **C**arlo) techniques that evaluate enacted map against ensemble of many alternate maps. His expert witness testimony is pivotal to the court's decision to strike down the PA map.
- **February 2018:** Moon Duchin (Tufts Univ mathematician) hired by PA Gov. Tom Wolfe to guide redistricting efforts
- **Summer 2018-present:** Mathematicians across US work to improve Pegden's MCMC techniques and make it more widely available

Summary

Mathematicians were key both to evaluating the enacted PA map AND guiding the redistricting. MCMC methods were already highly effective in 2017, and we've significantly improved them since.

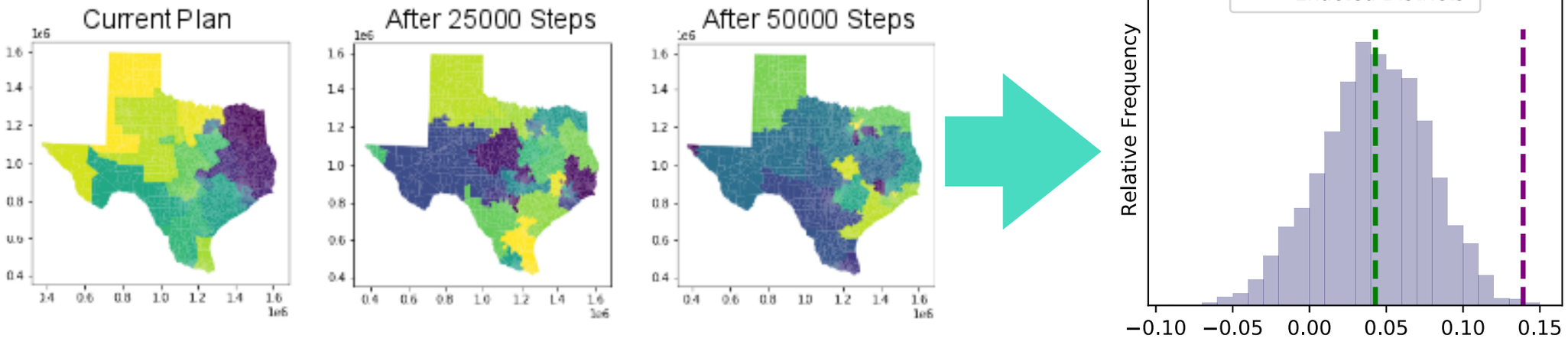
- https://ballotpedia.org/League_of_Women_Voters_of_Pennsylvania_v._the_Commonwealth_of_Pennsylvania
- https://www.governor.pa.gov/newsroom/governor-wolf-enlist-non-partisan-mathematician-evaluate-fairness-redistricting-maps/
- https://www.governor.pa.gov/wp-content/uploads/2018/02/md-report.pdf

Ensemble sampling can help you draw fair maps!

Conclusion

- Ensemble sampling allows us to create an unbiased baseline to understand what a fair plan should look like
- Ensemble sampling can be used to examine voting behavior, demographic data, and other statistics
- We have ensembles with ***2020 Census data*** and are **ready to analyze any proposed plan** for comparison.
- Ensemble sampling is fast, reliable, and validated.

"MUM_TX: Make your maps fair, and we'll get out of your hair!"



References

- Data and software
  - 2020 Population data, geodata from Census Bureau
  - 2020 election results: Texas Legislative Council
  - Software from MGGG (GerryChain): https://gerrychain.readthedocs.io/en/latest/
  - Precinct-level election geodata from MGGG (pre-2020): https://github.com/mggg-states
- Legal cases: see earlier slide
- Media coverage
  - I. Lapowsy, “The Geeks Who Put a Stop to Pennsylvania's Partisan Gerrymandering”, Wired, February 2018, https://www.wired.com/story/pennsylvania-partisan-gerrymandering-experts/
  - S. Roberts, https://www.technologyreview.com/2021/08/12/1031567/mathematicians-algorithms-stop-gerrymandering/
- Expert Reports
  - M. Duchin, *Outlier analysis for Pennsylvania congressional redistricting*, available at https://mggg.org/uploads/md-report.pdf
  - J. Mattingly, *Expert Report on the North Carolina State Legislature*, available at https:// sites.duke.edu/quantifyinggerrymandering/files/2019/09/Report.pdf
  - All expert reports prepared by MGGG: https://mggg.org/reports
- Academic centers
  - MGGG Redistricting Lab (Tufts): https://mggg.org
  - Quantifying Gerrymandering (Duke): https://sites.duke.edu/quantifyinggerrymandering/

Contacts

MathForUnbiasedMapsTX@list.smu.edu

https://www.smu.edu/Dedman/Research/Institutes-and-Centers/DCII/Scholarship/Research-Cluster-on-Political-Decision-Making

Andrea K. Barreiro, abarreiro@smu.edu,
http://faculty.smu.edu/abarreiro

***** Appendices*****

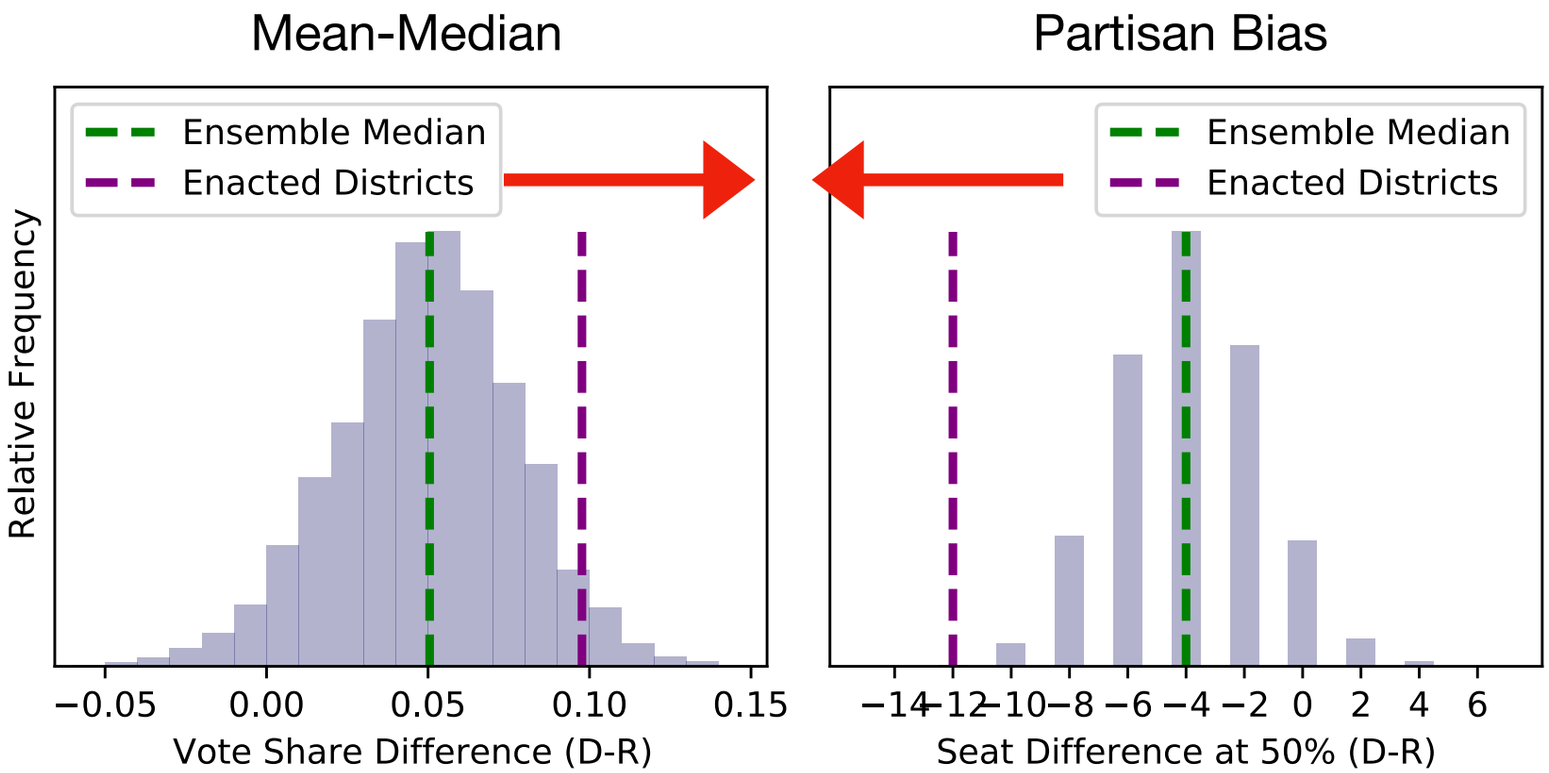
Do we want to include anything here?

For example, we could get our comments about “partisan geography” on the record.

Partisan geography: then and now

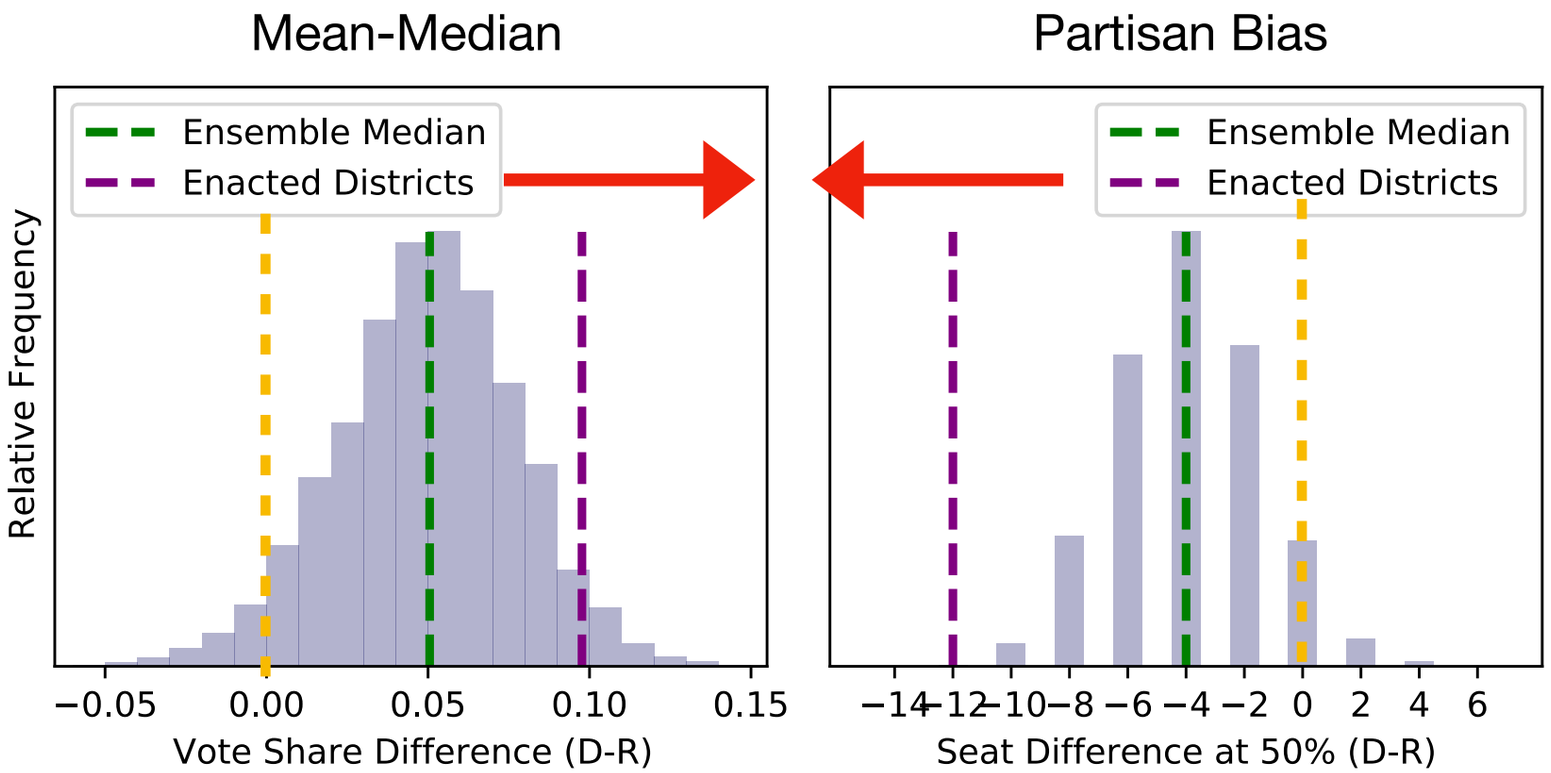
2010 vs. 2020

Ensembles from 2010 Census show a small Republican advantage



US Congressional Districts (36)
2010 Census Data
2012 Presidential Election

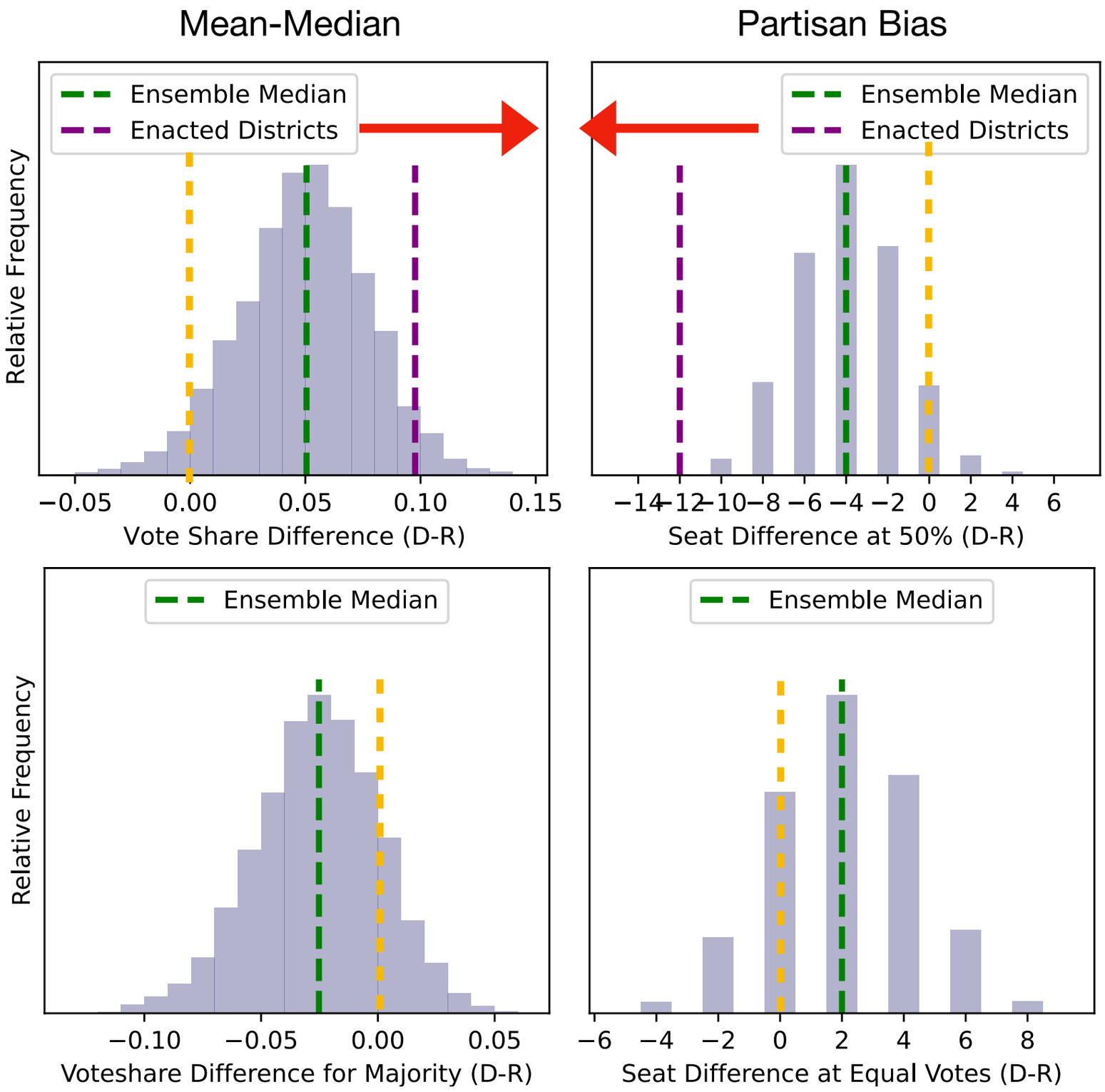
Ensembles from 2010 Census show a small Republican advantage



Party neutral (zero)

US Congressional Districts (36)
2010 Census Data
2012 Presidential Election

This advantage is NO LONGER PRESENT in 2020 Ensembles

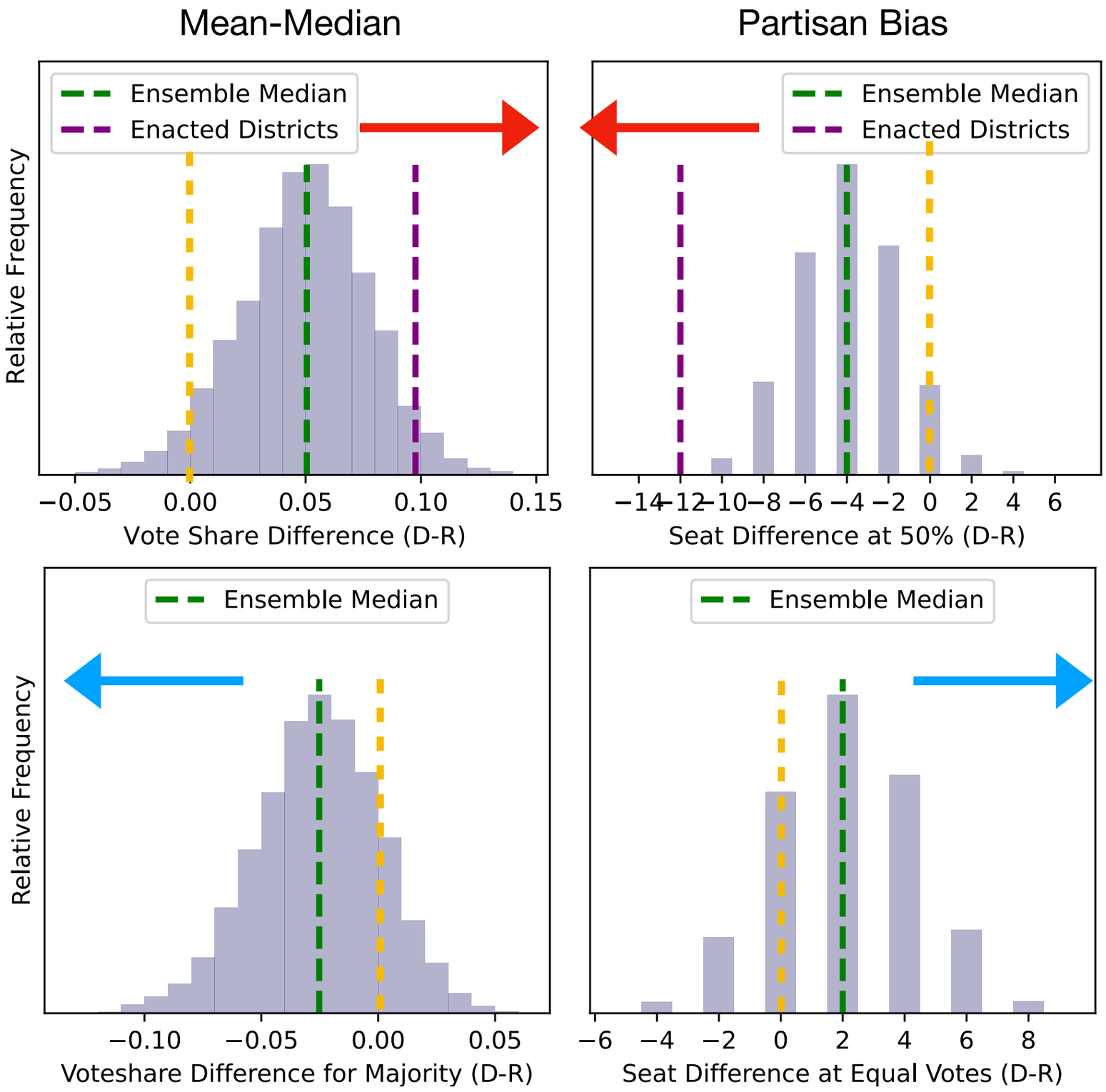


Party neutral (zero)

US Congressional Districts (36)
2010 Census Data
2012 Presidential Election

US Congressional Districts (38)
2020 Census Data
2020 Presidential Election

This advantage is NO LONGER PRESENT in 2020 Ensembles



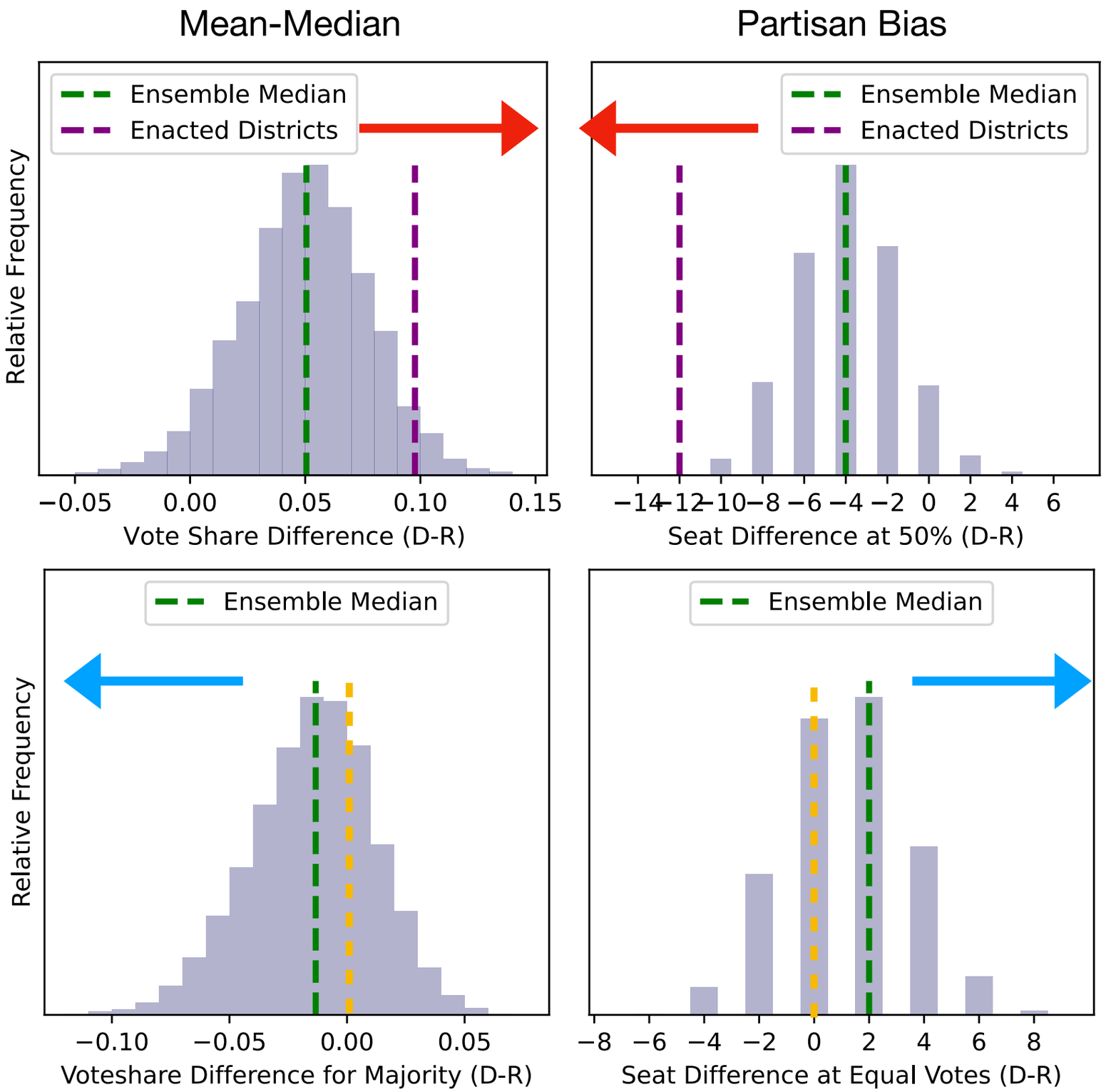
Party neutral (zero)

US Congressional Districts (36)
2010 Census Data
2012 Presidential Election

In fact, there is a small *Democratic* advantage

US Congressional Districts (38)
2020 Census Data
2020 Presidential Election

Similar results for 2020 Senate Election



Party neutral (zero)

US Congressional Districts (36)
2010 Census Data
2012 Presidential Election

2020 ensemble shows small *Democratic* advantage

US Congressional Districts (38)
2020 Census Data
2020 US Senate Election

Voting, ethnic data for USCD ensemble

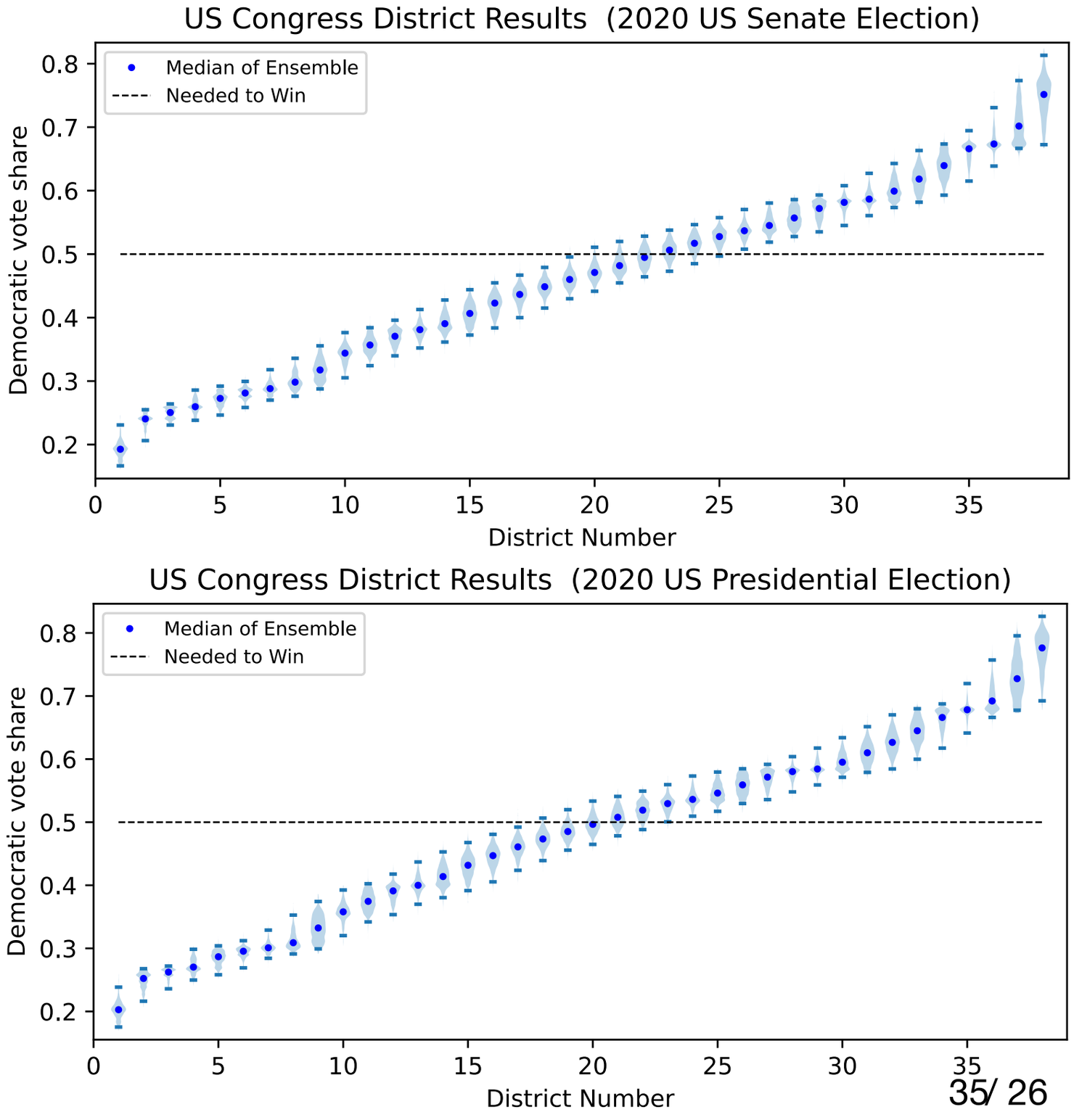
2020 Census data

Vote share curves, when compared with an unbiased sample, can show distinct signatures of partisan gerrymandering.

Vote Share Curves

- For each map, order districts by increasing vote share
- **Blue:** ensemble
- **Red:** enacted plan (or any comparison plan of interest)

Two 2020 Elections:
US Senate (top)
Presidential (bottom)



US Congressional Districts (38)
2020 Census Data

Vote share curves, when compared with an unbiased sample, can show distinct signatures of partisan gerrymandering.

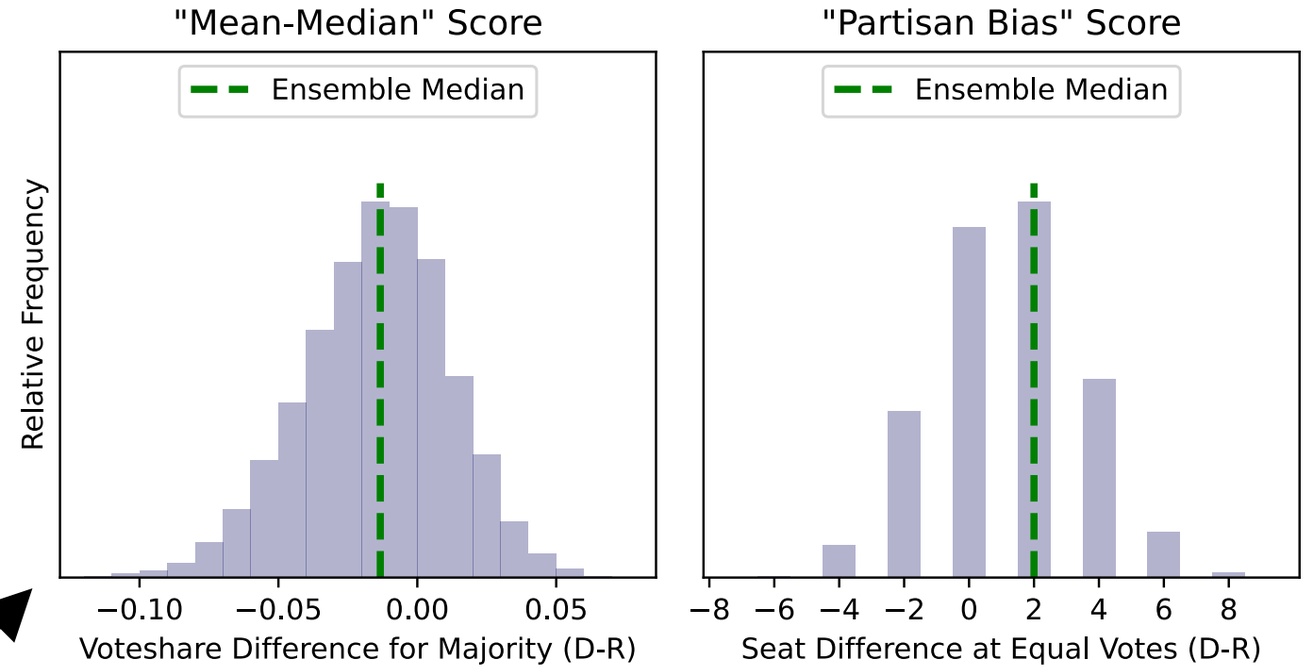
Two ways we can quantify this partisan skew:

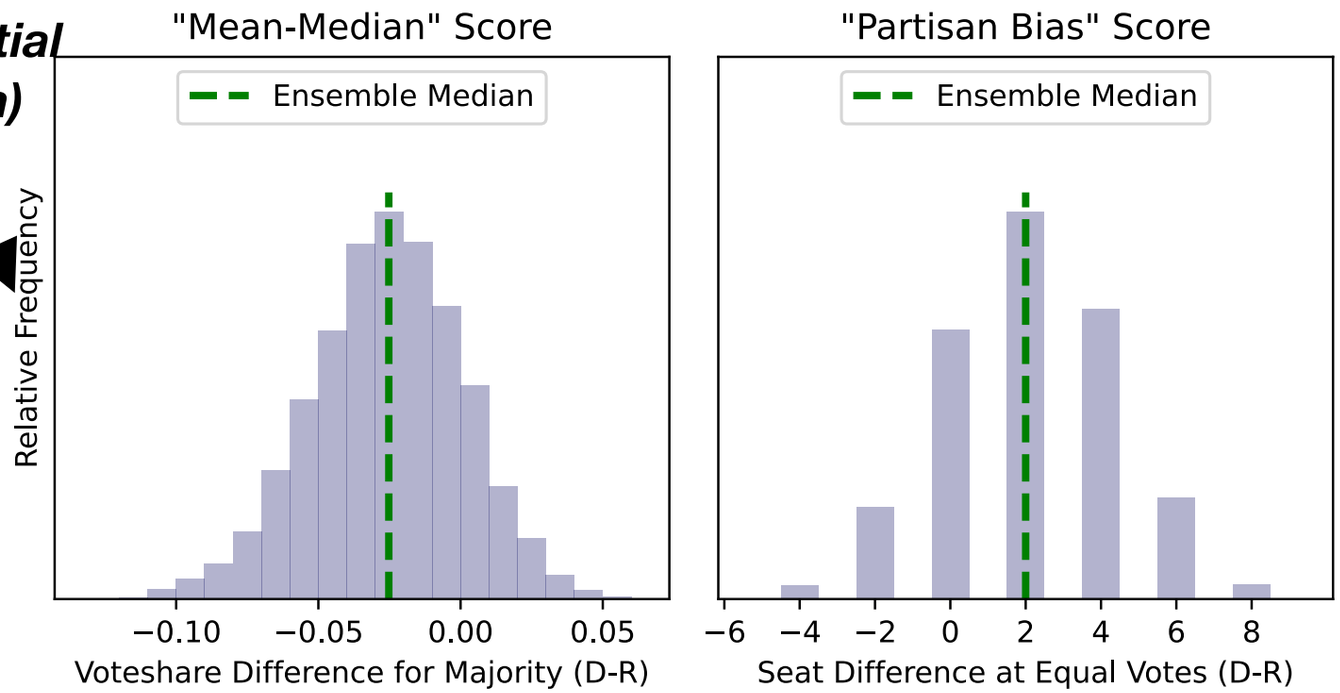
- Number of seats won at 50% vote share
  - ***Partisan Bias***
- Vote share needed for majority (16 seats)

Mean-Median Score

US Senate (top)

Presidential (bottom)





US Congressional Districts (38)
2020 Census Data

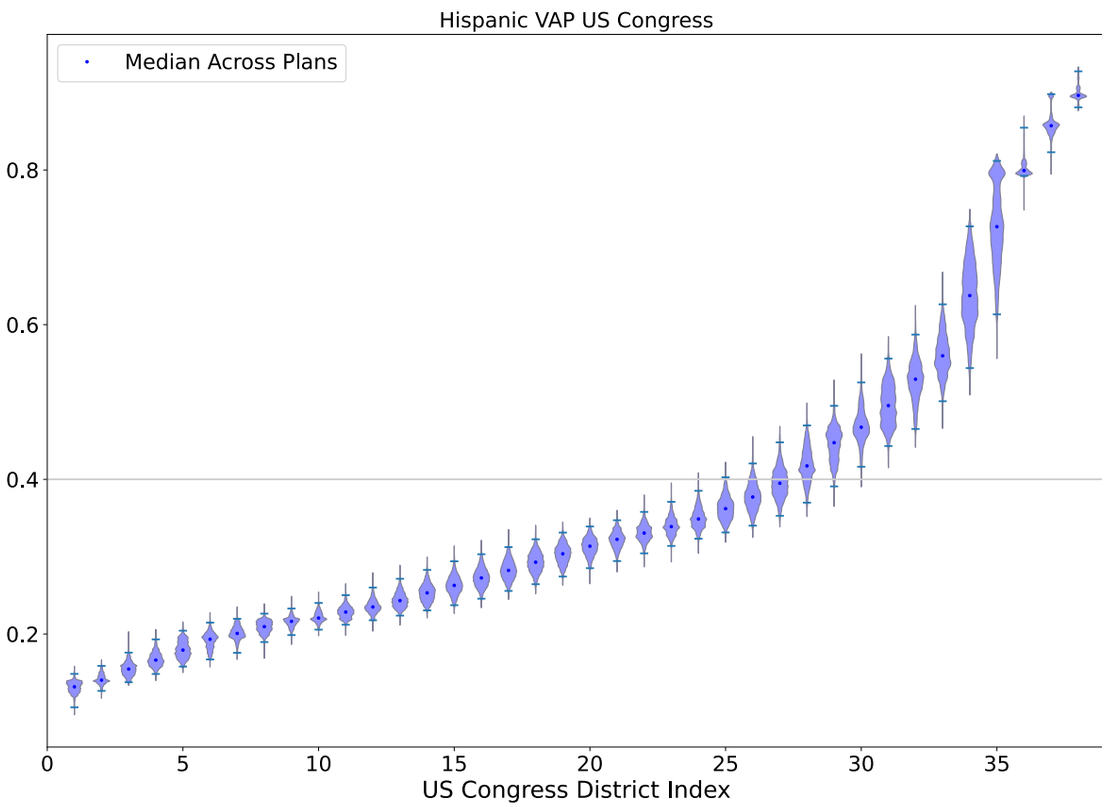
Math For Unbiased Maps TX (MUM_TX)

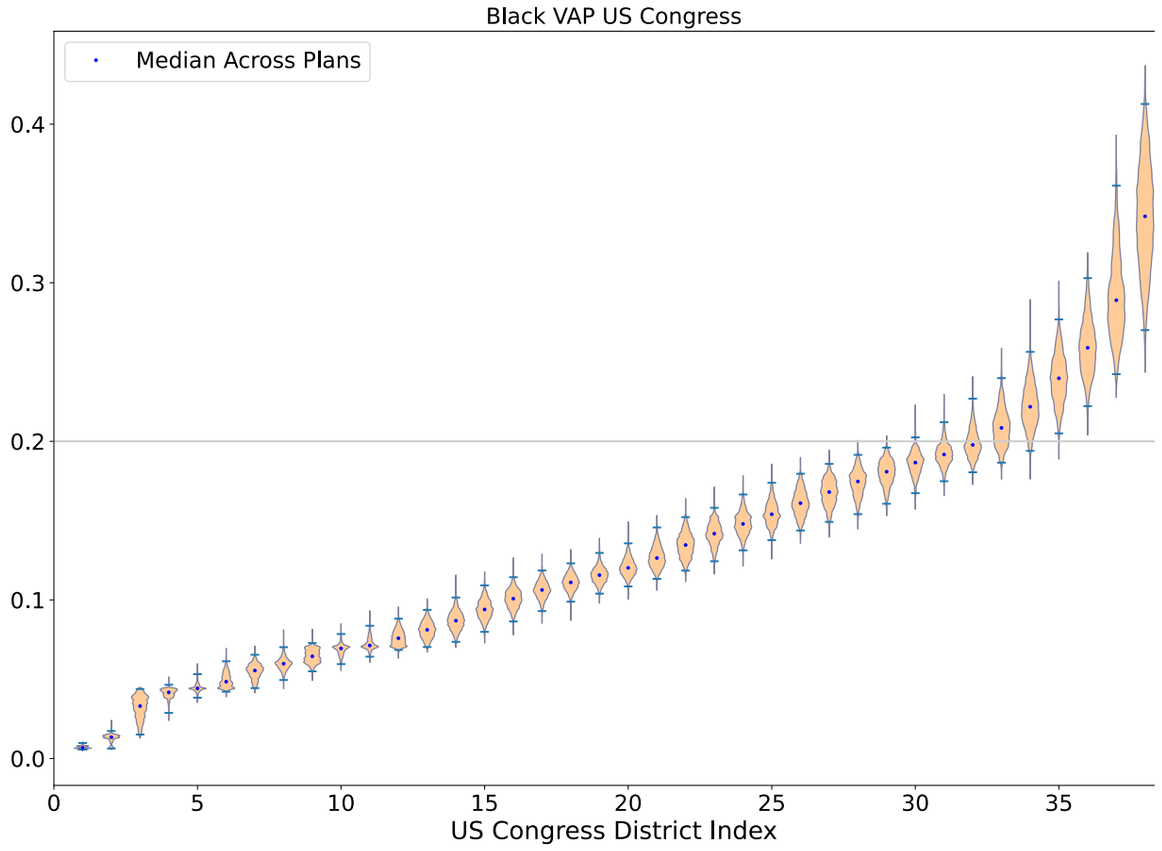
September 18, 2021

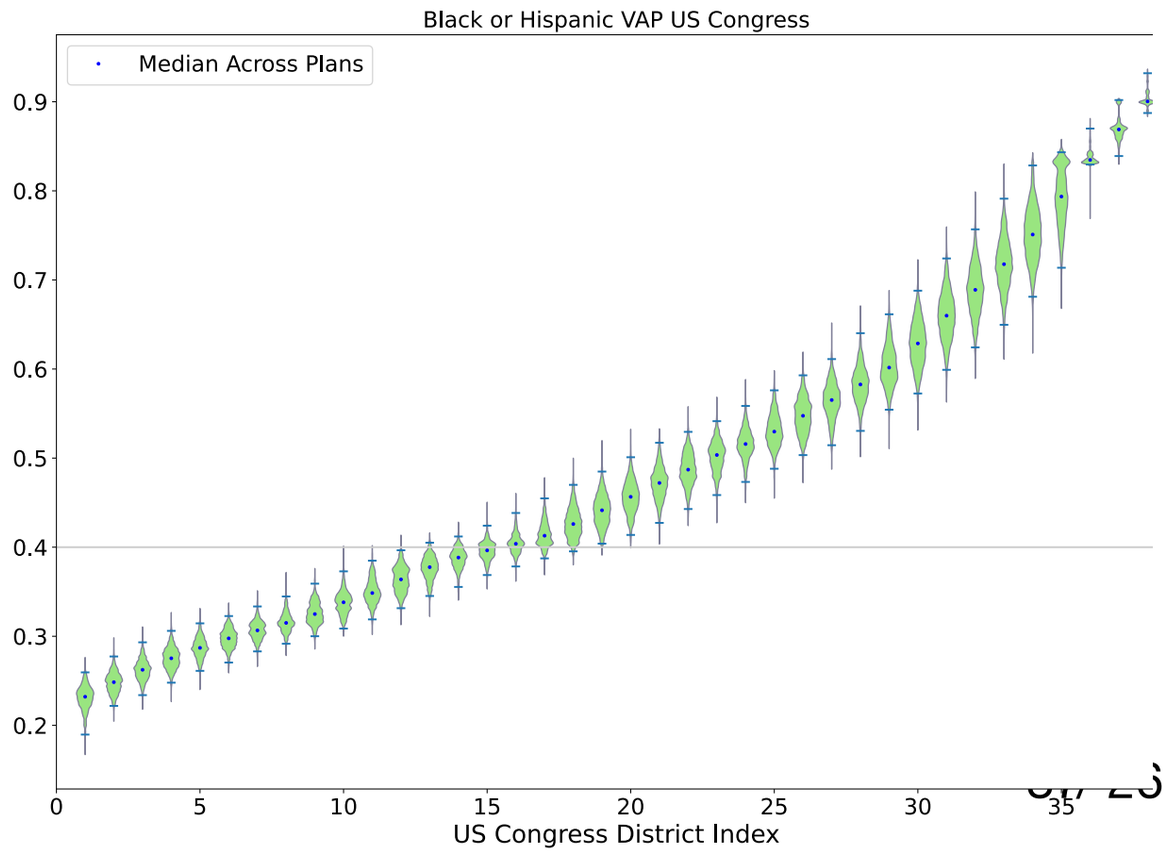
Complying with the Voting Rights Act

- For each map, order districts by ***HVAP***, ***BVAP***, or ***BHVAP***
- **Violin plots:** ensemble

Note: it is straightforward to compute any other group or coalition of interest







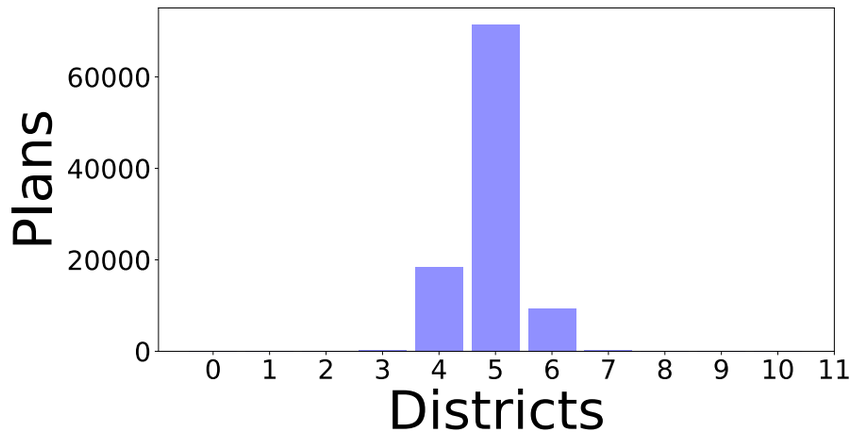
US Congressional Districts (38)

2020 Census Data

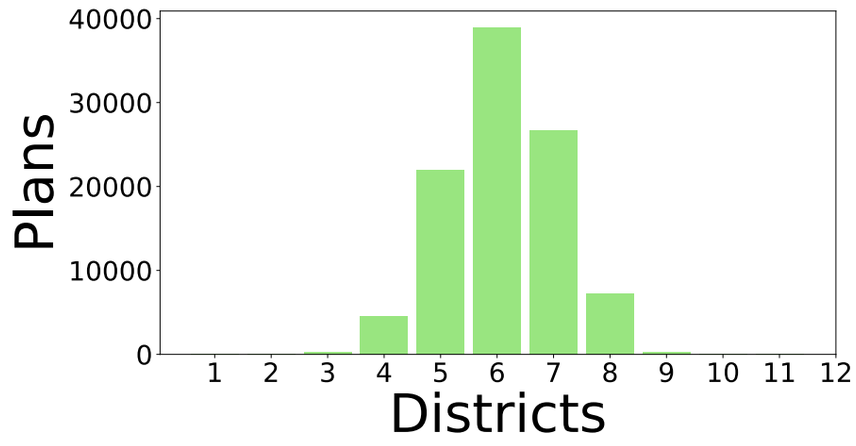
Complying with the Voting Rights Act

- How many districts have a critical threshold of group/ coalition of interest?

USCD HVAP>60%



USCD BHVAP>70%



USCD BVAP>30%

US Congressional Districts (38)

2020 Census Data

Note: it is straightforward to compute any other group or coalition of interest

Voting, ethnic data for TX Senate ensemble

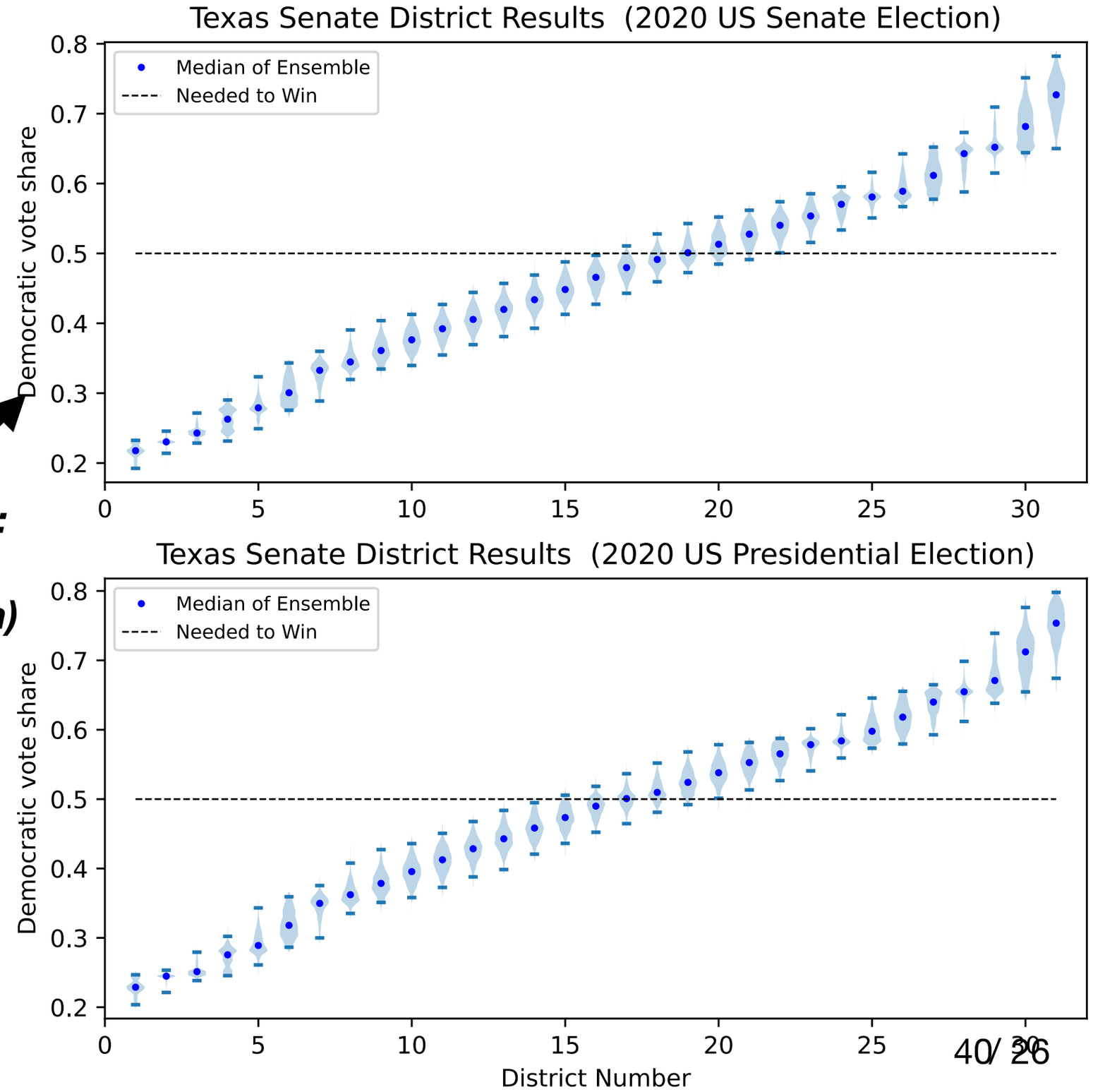
2020 Census data

Vote share curves, when compared with an unbiased sample, can show distinct signatures of partisan gerrymandering.

Vote Share Curves

- For each map, order districts by increasing vote share
- **Blue:** ensemble
- **Red:** enacted plan (or any comparison plan of interest)

Two 2020 Elections:
US Senate (top)
Presidential (bottom)



TX Senate Districts (31)
2020 Census Data

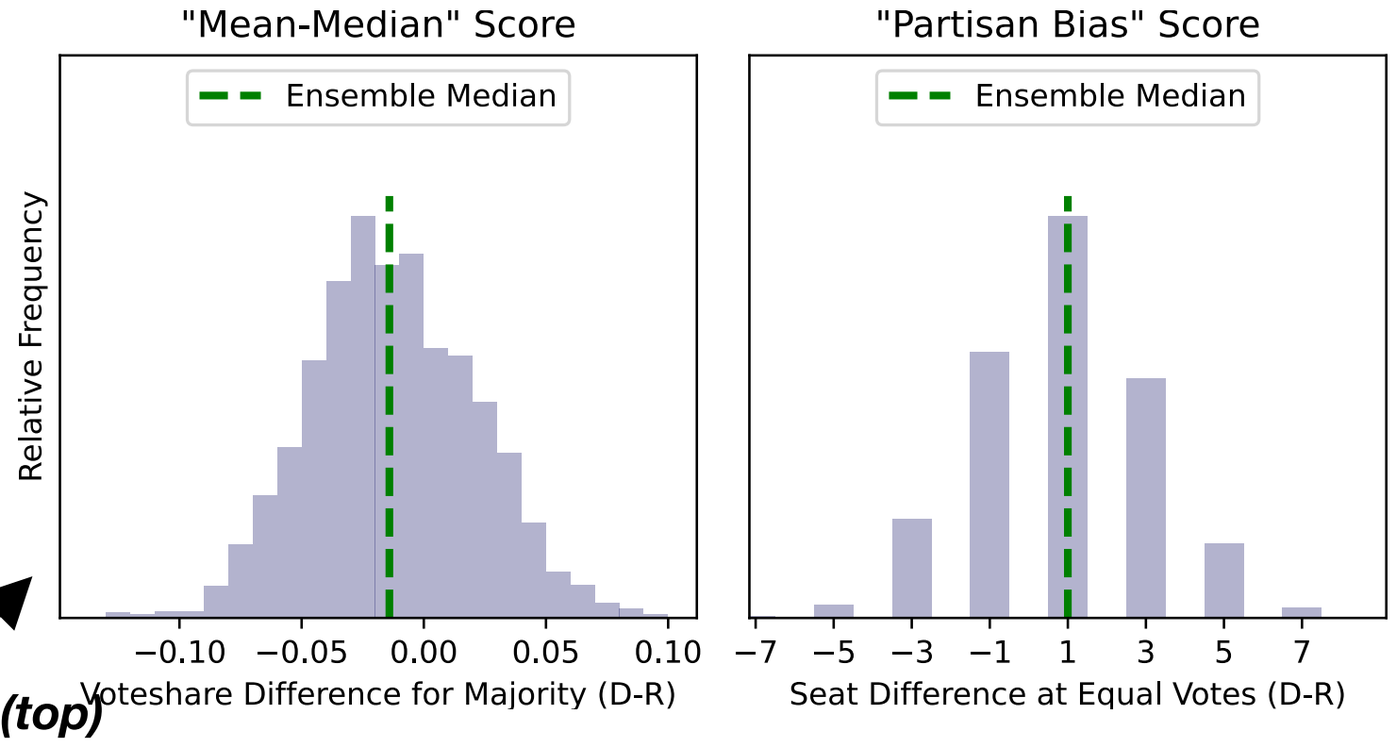
Vote share curves, when compared with an unbiased sample, can show distinct signatures of partisan gerrymandering.

Two ways we can quantify this partisan skew:

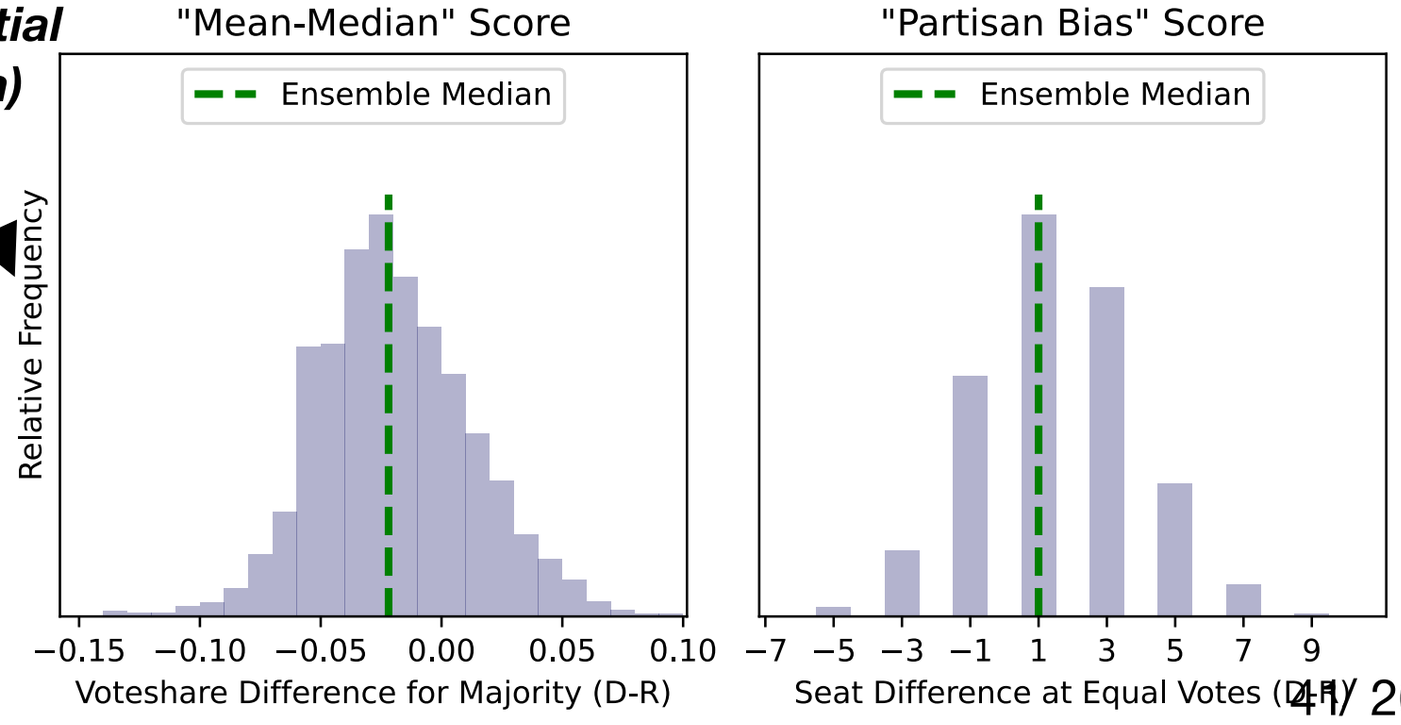
- Number of seats won at 50% vote share
  - ***Partisan Bias***
- Vote share needed for majority (16 seats)

Mean-Median Score

US Senate (top)



Presidential (bottom)

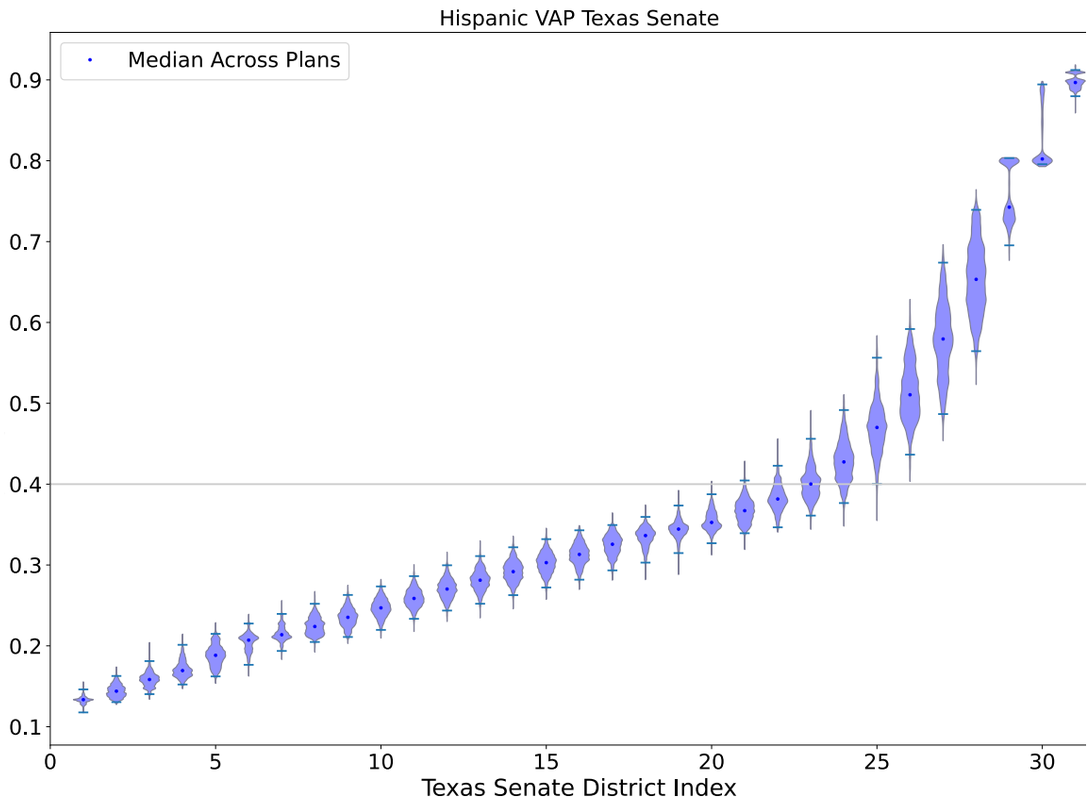


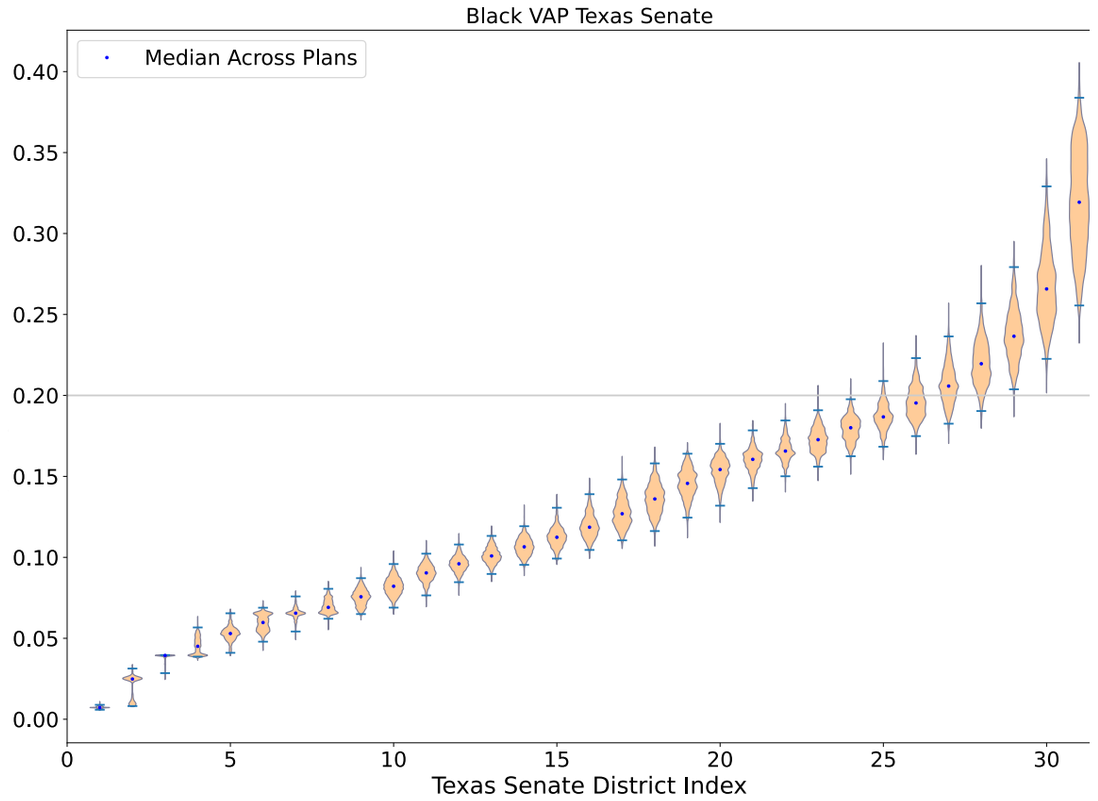
TX Senate Districts (31)
2020 Census Data

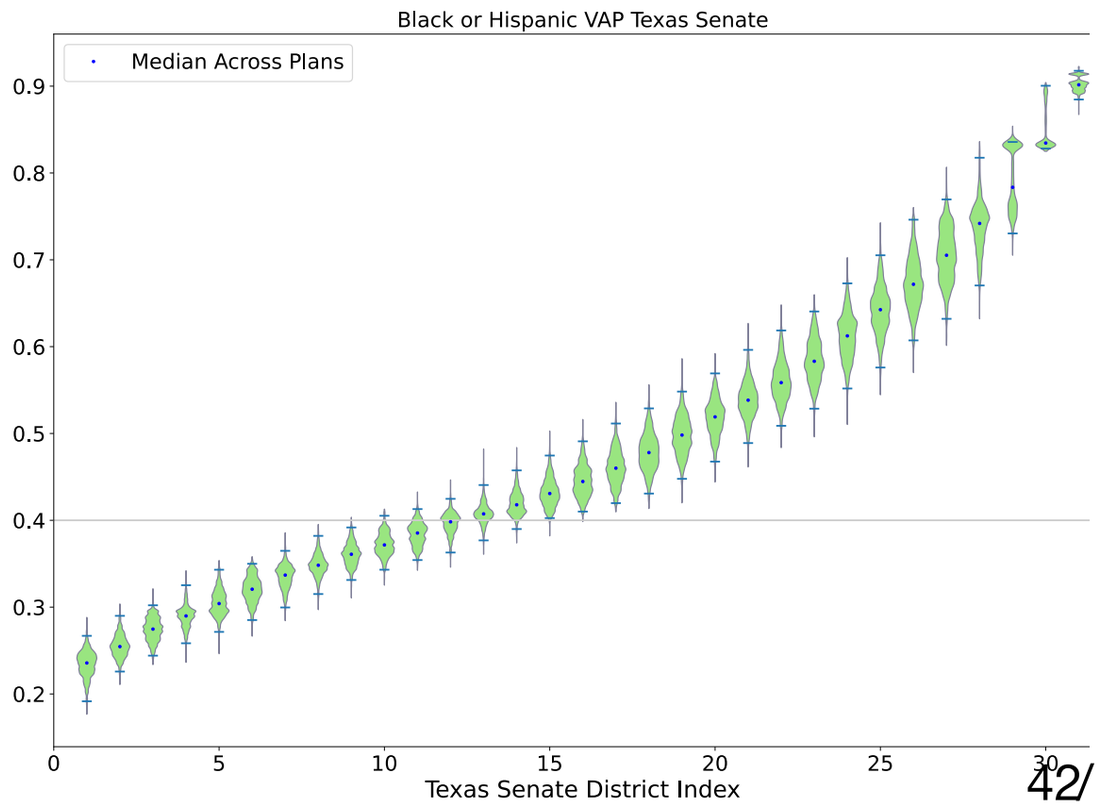
Complying with the Voting Rights Act

- For each map, order districts by ***HVAP***, ***BVAP***, or ***BHVAP***
- **Violin plots:** ensemble

Note: it is straightforward to compute any other group or coalition of interest





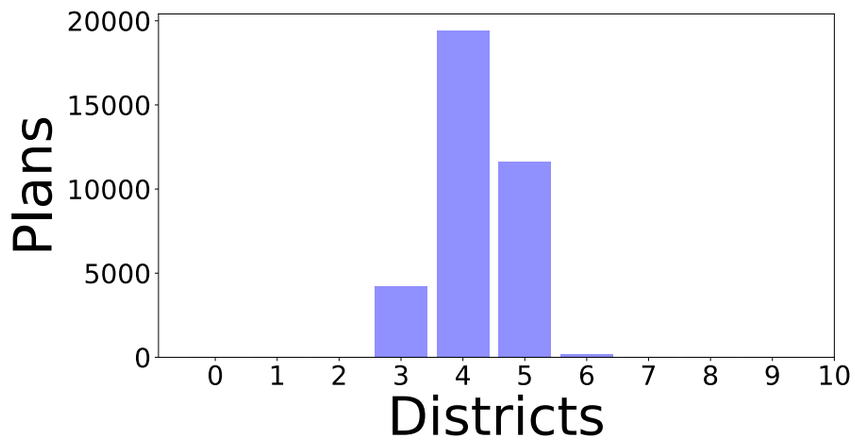


TX Senate Districts (31)
2020 Census Data

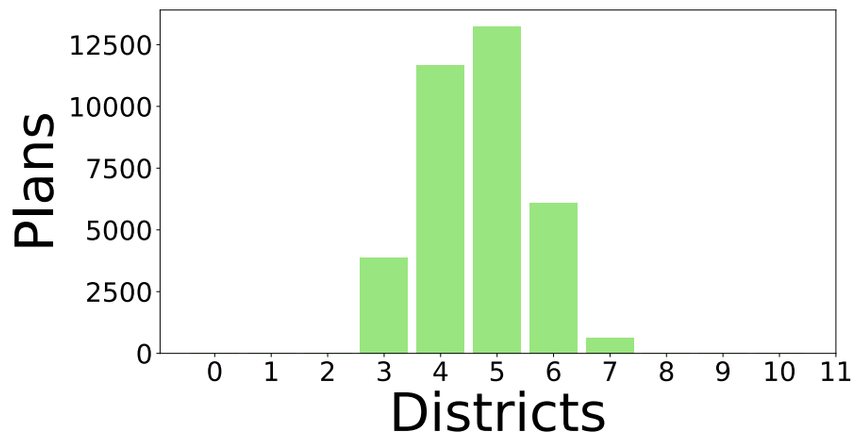
Complying with the Voting Rights Act

- How many districts have a critical threshold of group/ coalition of interest?

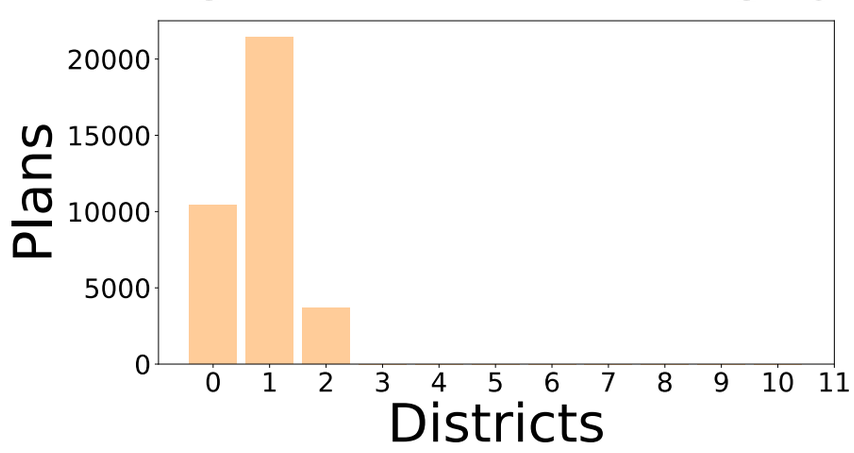
TXSN HVAP>60%



TXSN BHVAP>70%



TXSN BVAP>30%



TX Senate Districts (31)

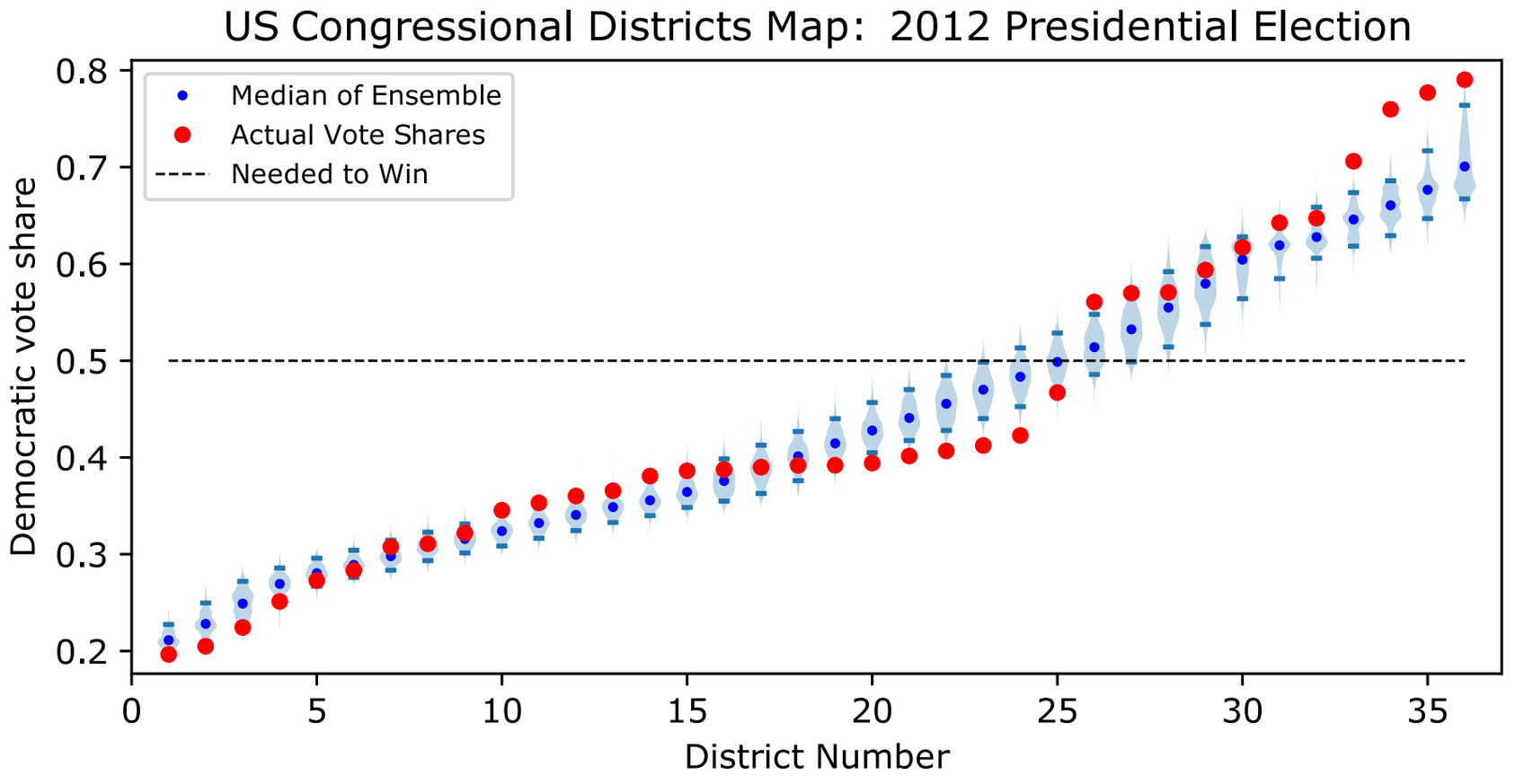
2020 Census Data

Note: it is straightforward to compute any other group or coalition of interest

Voting data for current USCD maps, 2010 Census data

Vote Share Curves

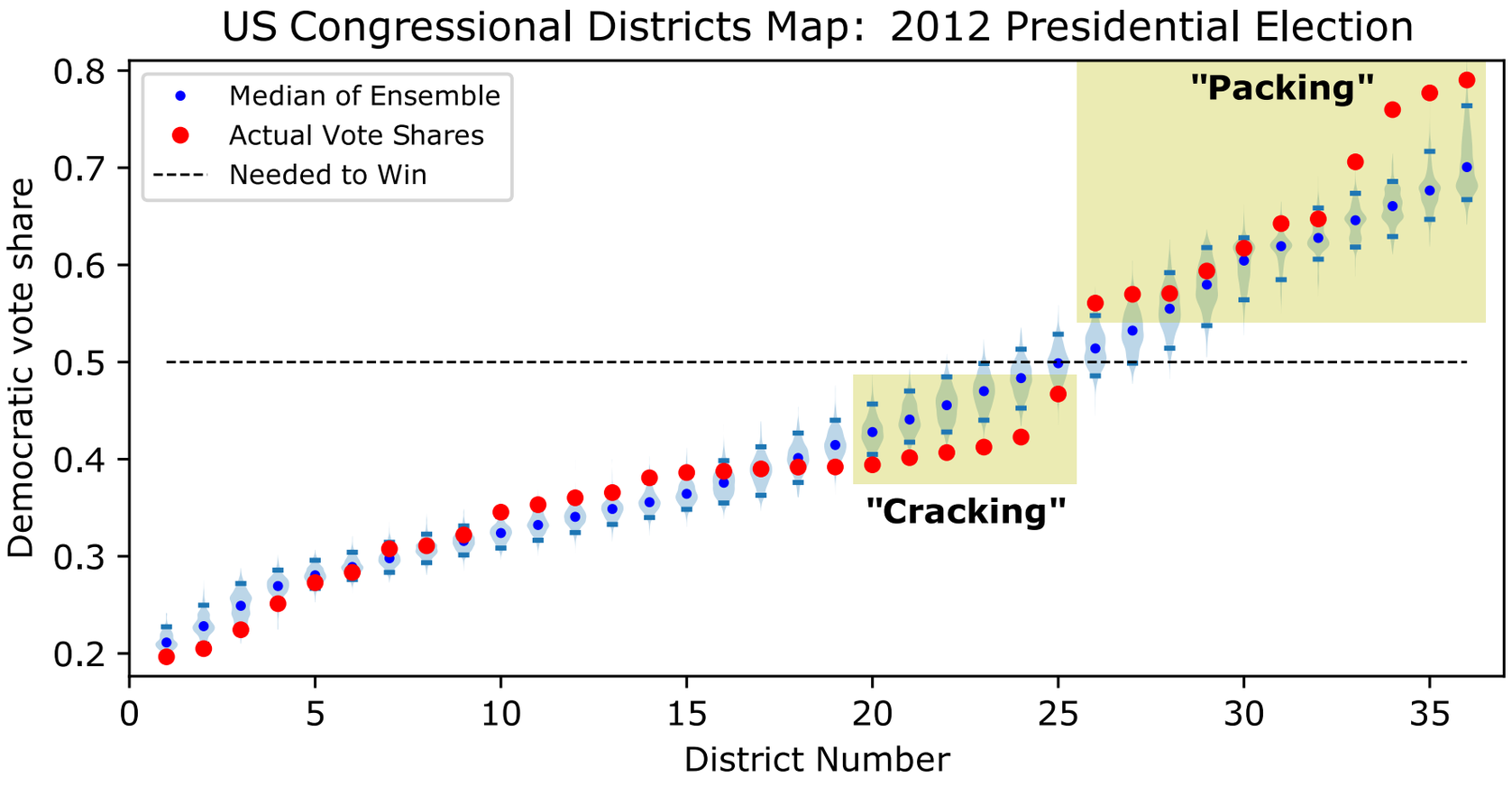
- For each map, order districts by increasing vote share
- **Blue:** ensemble
- **Red:** enacted plan (or any comparison plan of interest)



US Congressional Districts (36)
2010 Census Data

Vote Share Curves

- For each map, order districts by increasing vote share
- **Blue:** ensemble
- **Red:** enacted plan (or any comparison plan of interest)

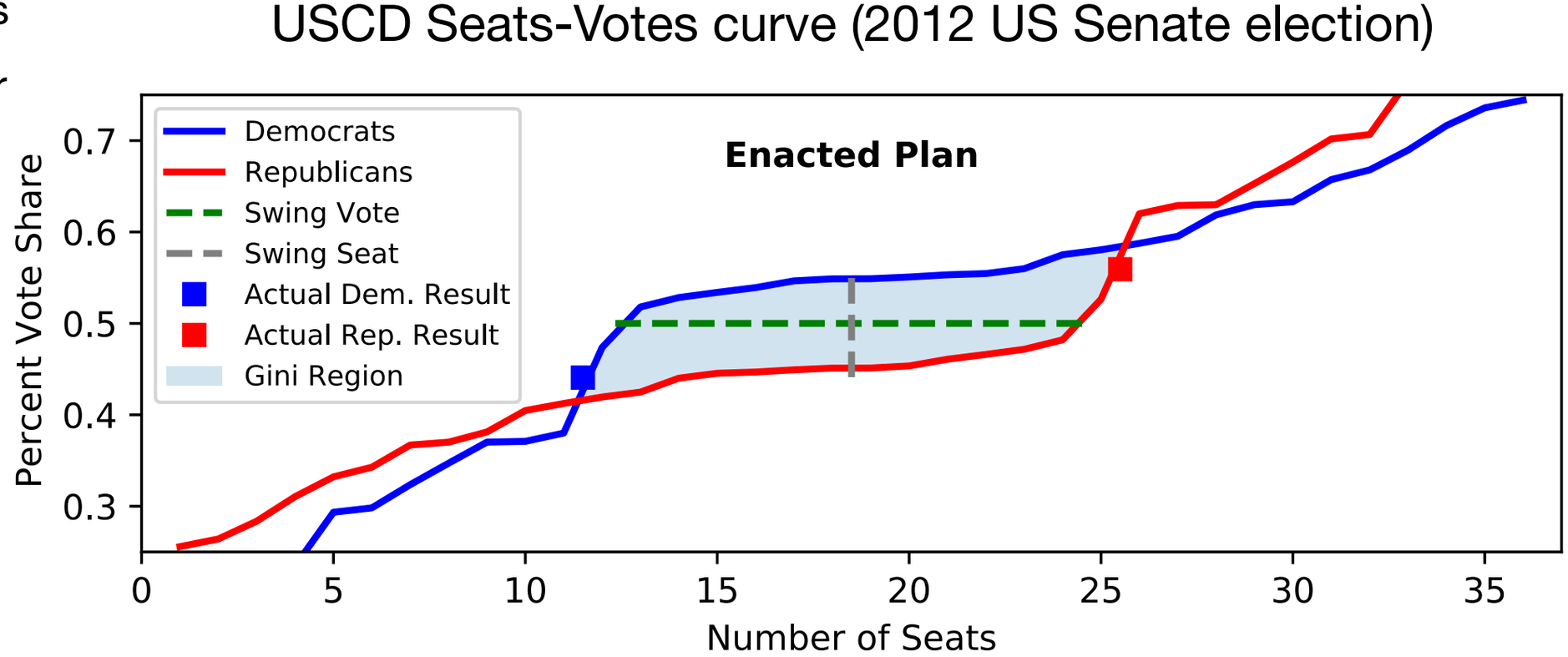


- ***Cracking:*** spreading the opposing party's voters across multiple districts
- ***Packing:*** concentrating the opposing party's voters into a few safe districts

US Congressional Districts (36)
2010 Census Data

Seats-Votes Curve

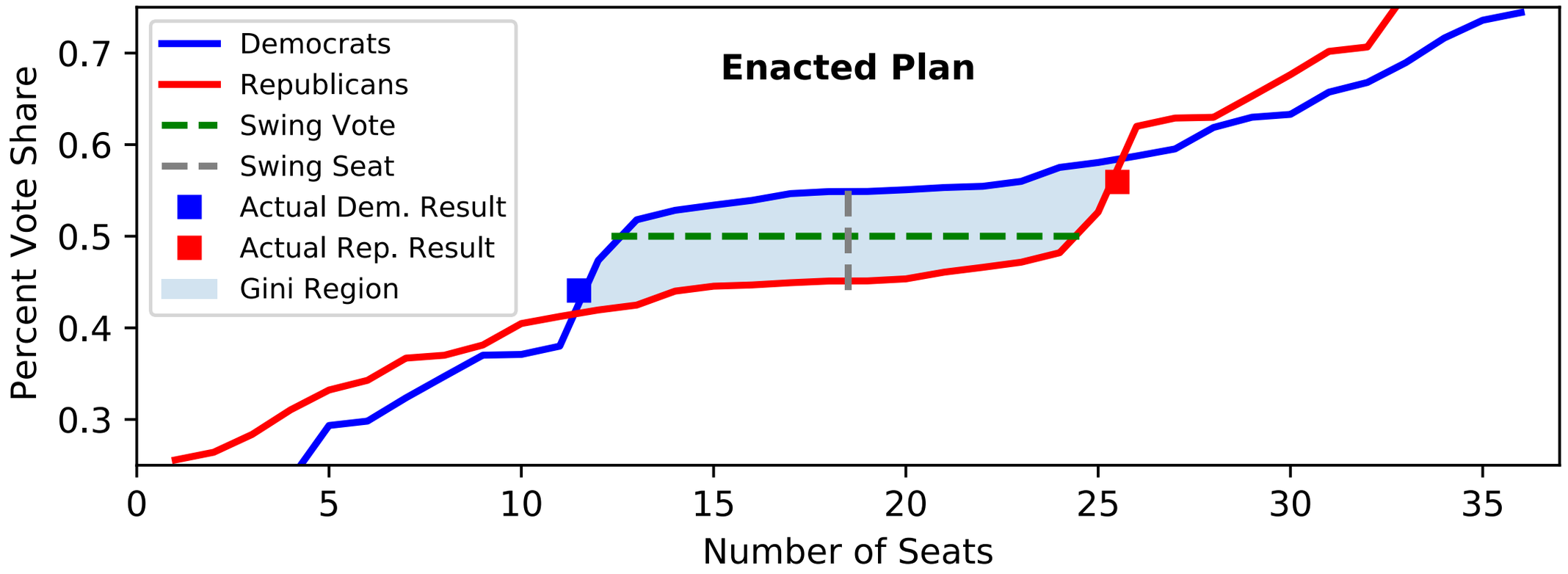
- Estimated statewide vote share needed to win a given # of seats
- This assumes a linear shift in voting preferences



US Congressional Districts (36)
2010 Census Data

Above: Seats-votes curve for the US Congressional Districts, computed using vote data from the 2012 Senate election

What can we learn from the Seats-Votes Curve?



Partisan Bias

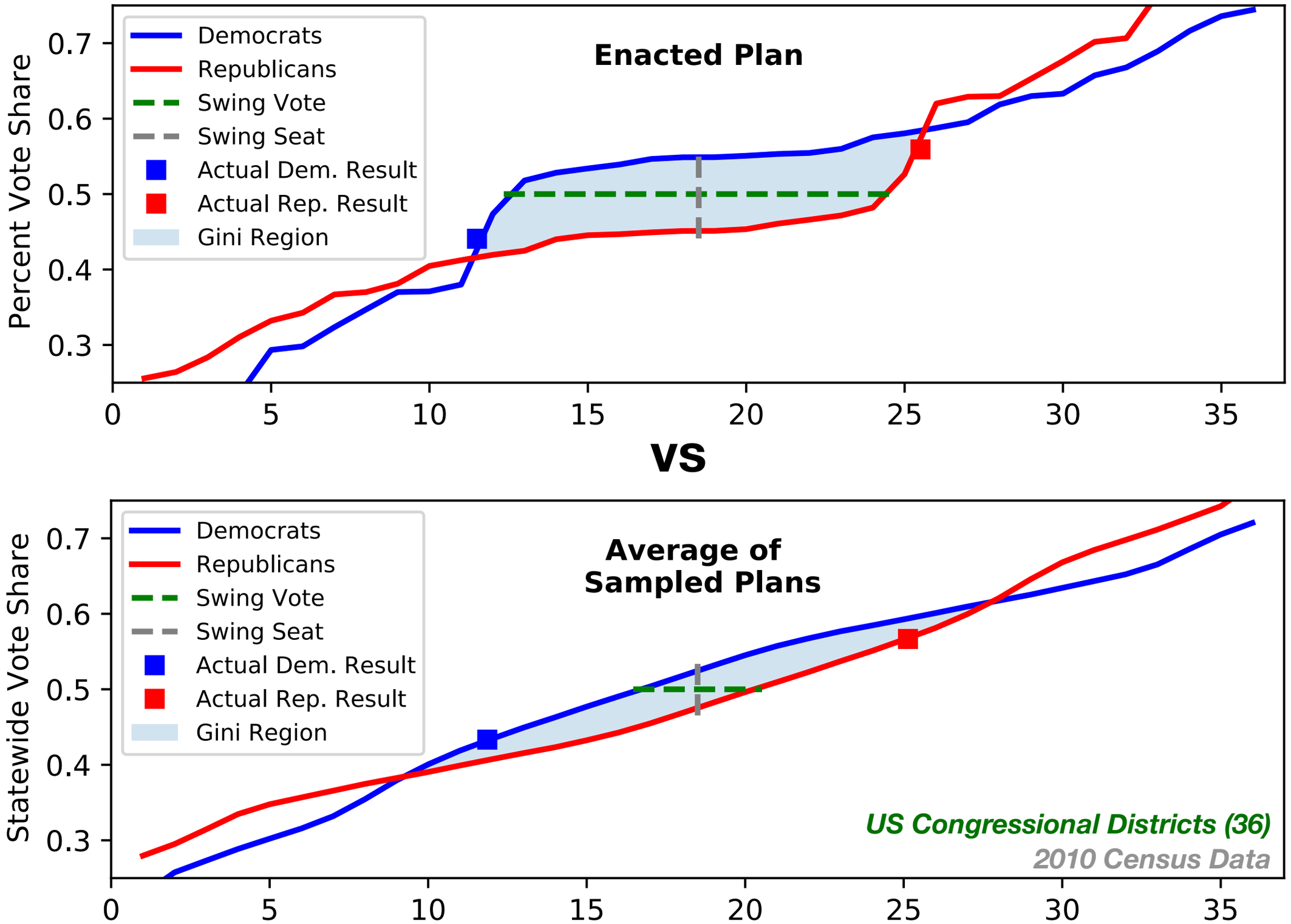
- Number of seats won at 50% vote share
  - Democrats: 12
  - Republicans: 24
  - ***Difference: -12***

Commonly used metrics in political science literature

Mean-Median Score

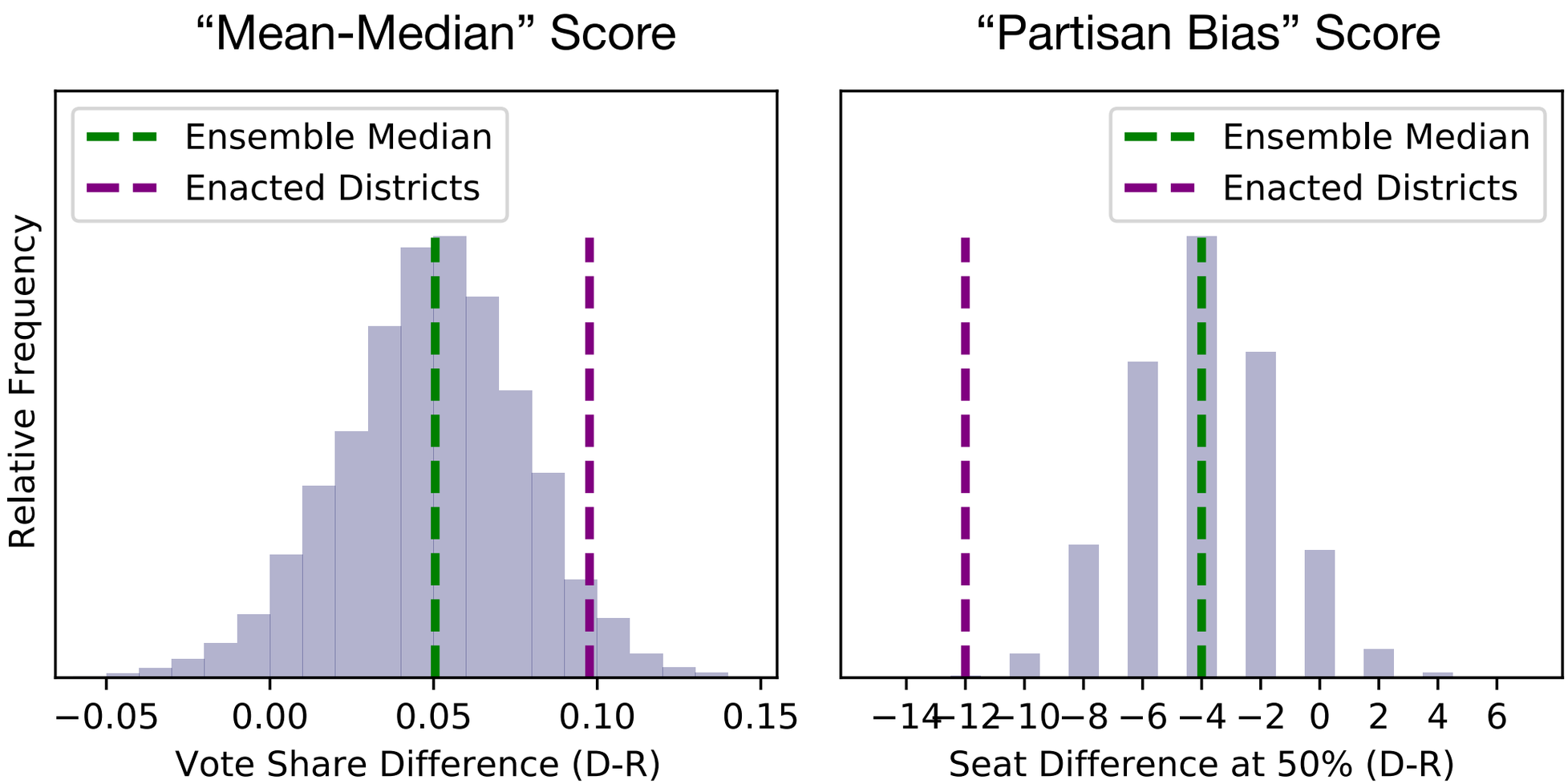
- Vote share needed for majority (18 seats)
  - Democrats: 55%
  - Republicans: 45%
  - ***Difference: 10%***

Is this outcome representative of an ensemble of plans?
No!



Is this outcome representative of an ensemble of plans?
No!

US Congressional Districts (36)
2010 Census Data



Average: 5%

Actual plan: 10%

Only 1 in 25 plans show this level of disparity

Average: -4

Actual plan: -12

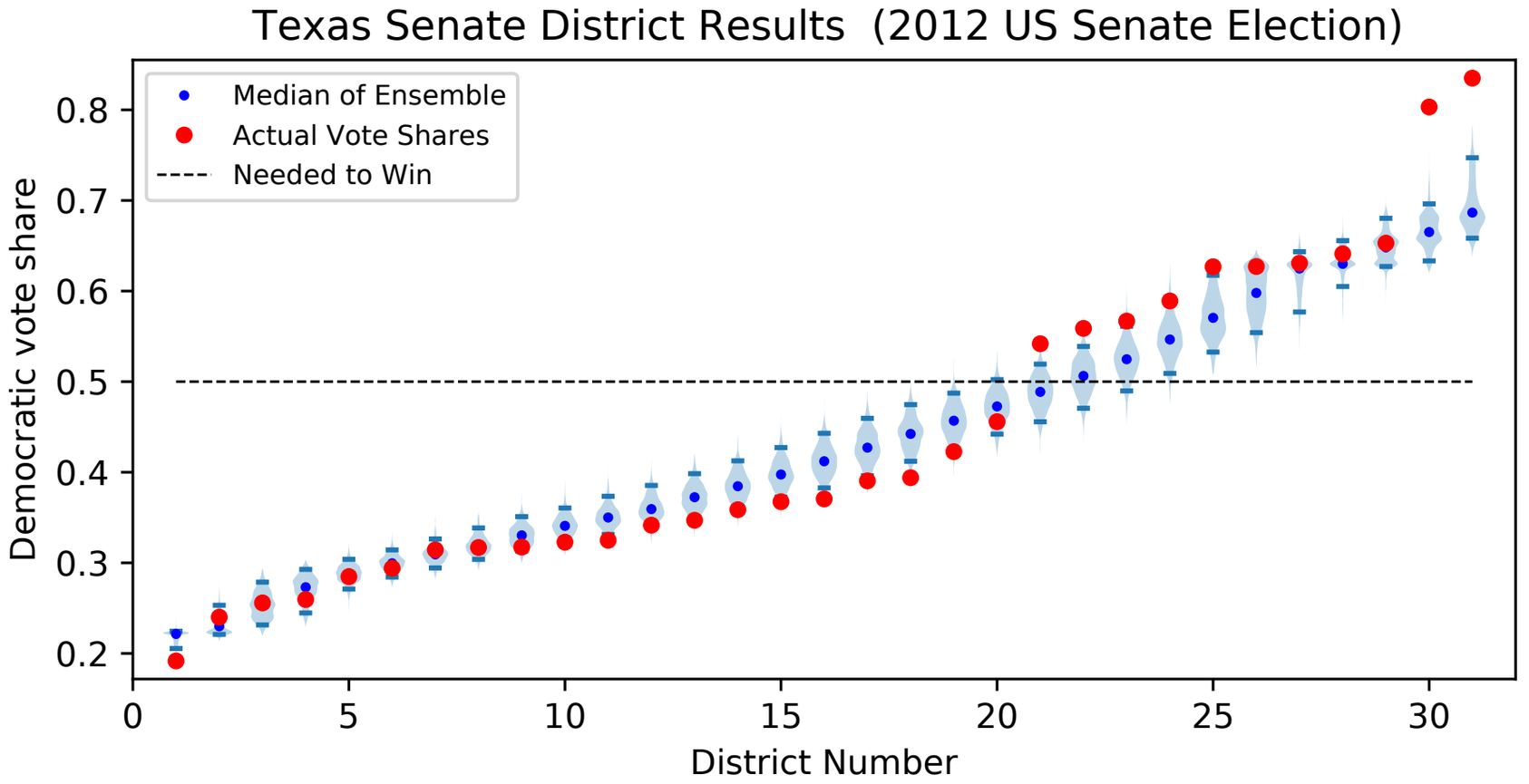
Only 1 in 800 plans show this level of disparity

*Vote% Diff expressed as fraction of 1; i.e. 0.1=10%

Voting data for current TX Senate maps, 2010 Census data

Vote Share Curves

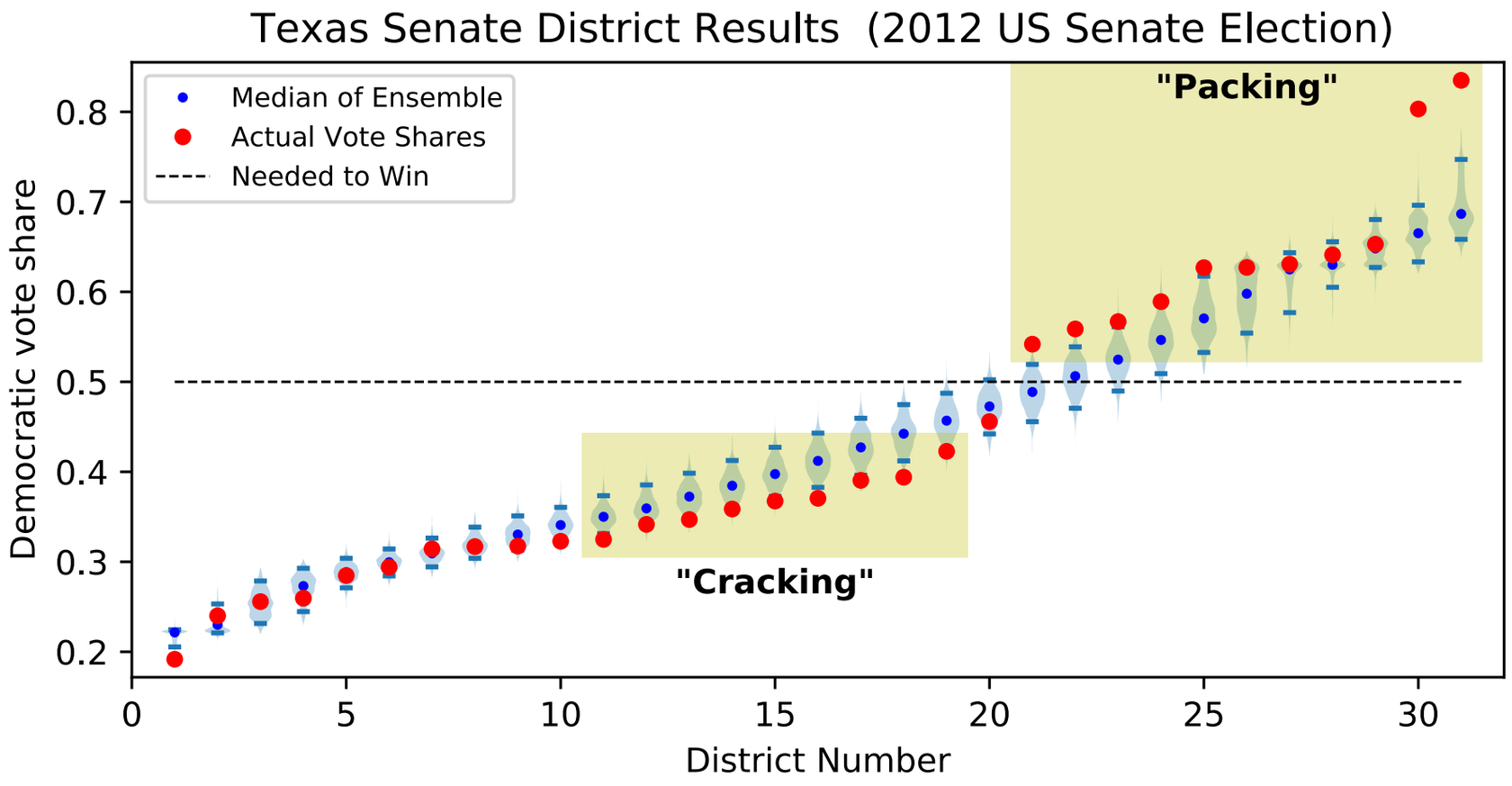
- For each map, order districts by increasing vote share
- **Blue:** ensemble
- **Red:** enacted plan (or any comparison plan of interest)



TX Senate Districts (31)
2010 Census Data

Vote Share Curves

- For each map, order districts by increasing vote share
- **Blue:** ensemble
- **Red:** enacted plan (or any comparison plan of interest)

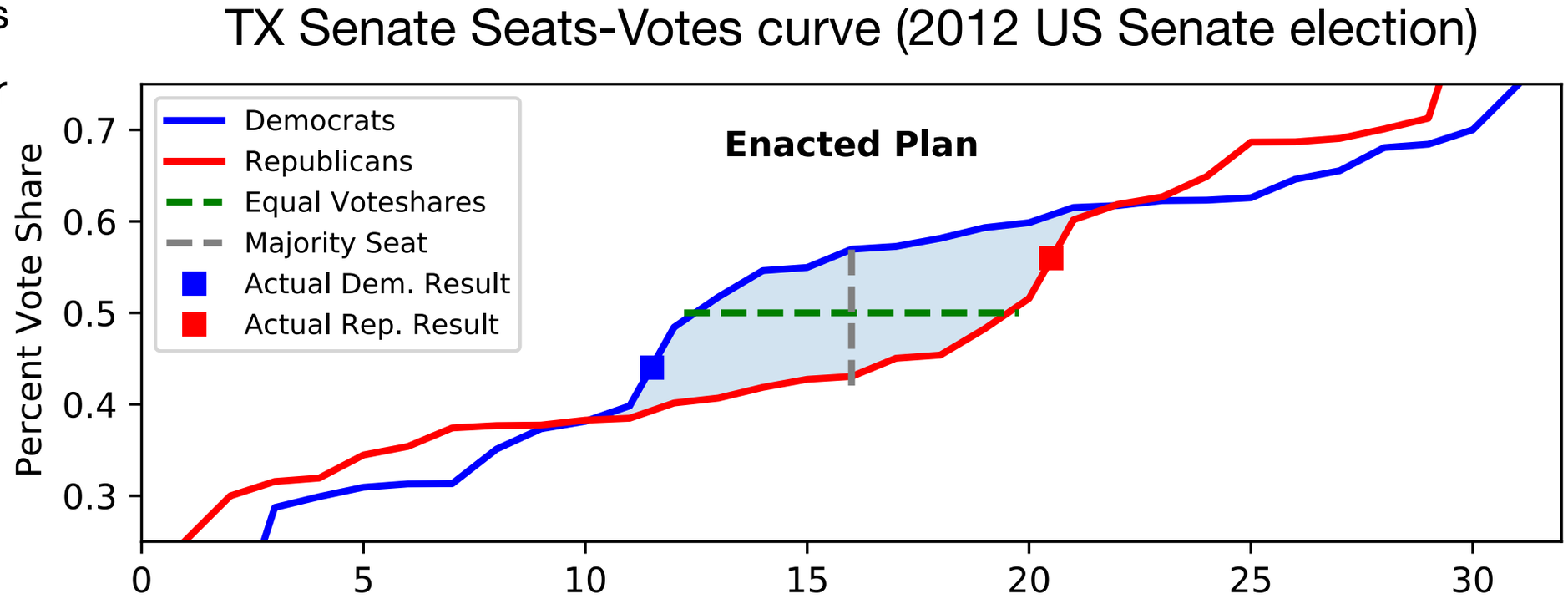


- ***Cracking:*** spreading the opposing party's voters across multiple districts
- ***Packing:*** concentrating the opposing party's voters into a few safe districts

TX Senate Districts (31)
2010 Census Data

Seats-Votes Curve

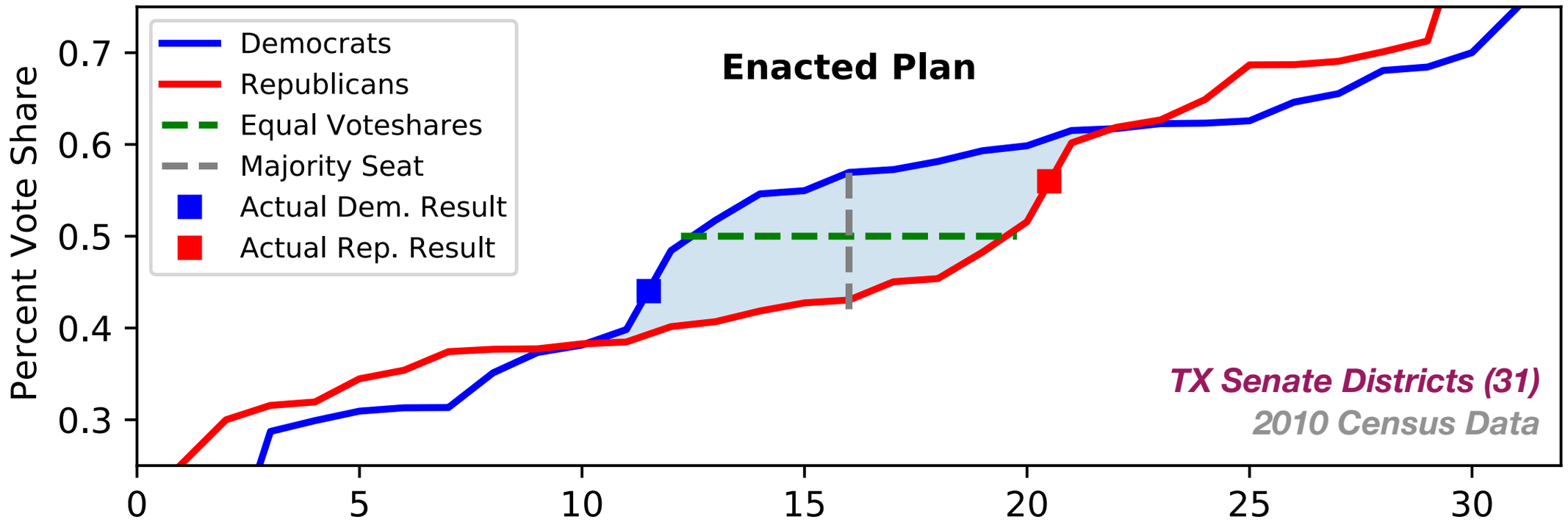
- Estimated statewide vote share needed to win a given # of seats
- This assumes a linear shift in voting preferences



Above: Seats-votes curve for the TX State Senate plan, computed using vote data from the 2012 Senate election

TX Senate Districts (31)
2010 Census Data

What can we learn from the Seats-Votes Curve?



Partisan Bias

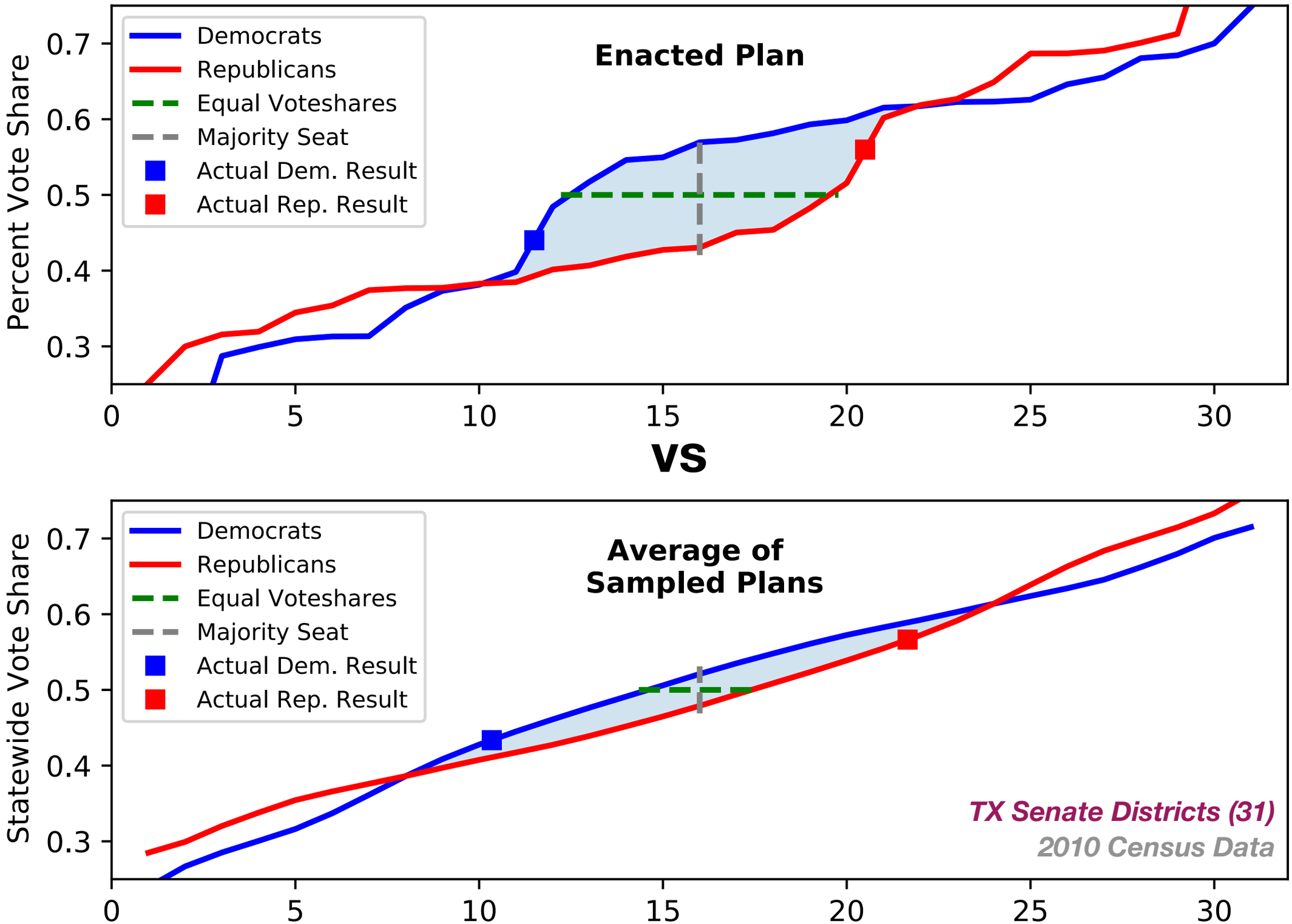
- Number of seats won at 50% vote share
  - Democrats: 12
  - Republicans: 19
  - ***Difference: -7***

Commonly used metrics in political science literature

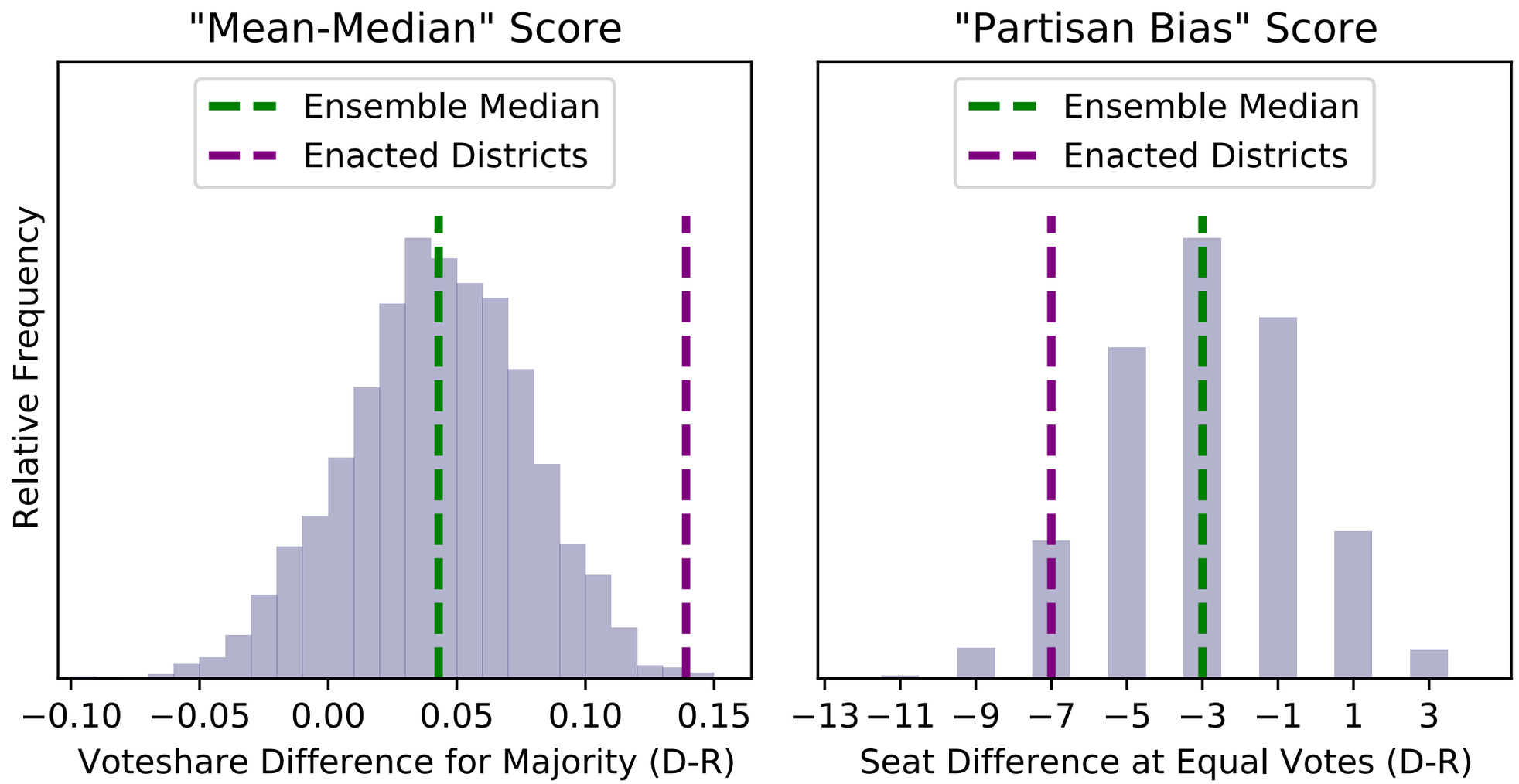
Mean-Median Score

- Vote share needed for majority (16 seats)
  - Democrats: 57%
  - Republicans: 43%
  - ***Difference: 14%***

Is this outcome representative of an ensemble of plans?
No!



Is this outcome representative of an ensemble of plans?
No!



Average: 4%

Actual plan: 14%

Only 1 in 500 plans show this level of disparity

Average: -3

Actual plan: -7

1 in 14 plans show this level of disparity

*Vote% Diff expressed as fraction of 1; i.e. 0.1=10%