

Use of Simulation Algorithms for Legislative Redistricting Analysis and Evaluation

Kosuke Imai

Departments of Government and Statistics
Harvard University

SMaSH: Symposium for Mathematical Sciences at Harvard
May 17, 2022

Joint work with Christopher Kenny, Cory McCartan, and others

Self-introduction

- Affiliation: Department of Government and Department of Statistics
- Major research areas: applied statistics
 - 1 Causal inference and policy evaluation
 - 2 Computational social science
- Recent projects:
 - Evaluation of the Indian national health insurance program
 - Evaluation of the pretrial risk assessment scores in the U.S. criminal justice system
 - Evaluation of legislative redistricting in the U.S. and Japan

What Is Legislative Redistricting?

- **One person, one vote:** a basic principle of representative democracy
- Population change \rightsquigarrow reapportionment + redistricting
- Congressional reapportionment and redistricting in America
 - happens after every decennial Census
 - reapportionment of 435 House of Representatives across 50 states
 - redistricting within each state
 - every state has different redistricting criteria and rules
- Other redistricting
 - state legislative districts
 - local electoral districts
 - school districts and attendance zones

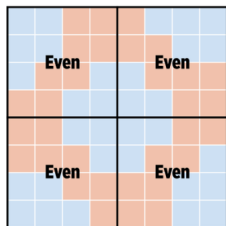
Gerrymander = Gerry + Salamander

- Massachusetts Governor Elbridge Gerry's 1812 map

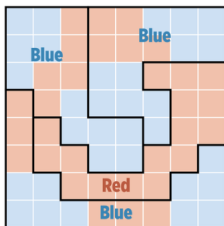


Today's Gerrymandering

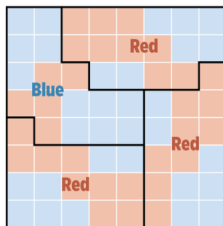
- Basic strategies — **packing and cracking** — have not changed



Even distribution
2 red, 2 blue



Packing
1 red, 3 blue



Cracking
3 red, 1 blue

- What have changed:
 - 1 availability of granular data and mapping software (e.g., Maptitude)
 - 2 recent court decisions
 - racial gerrymandering: *Shelby County v. Holder* (2013)
 - partisan gerrymandering: *Rucho v. Common Cause* (2019)

Algorithm-Assisted Redistricting Methodology (ALARM)



Developing methodology and tools to analyze legislative redistricting.

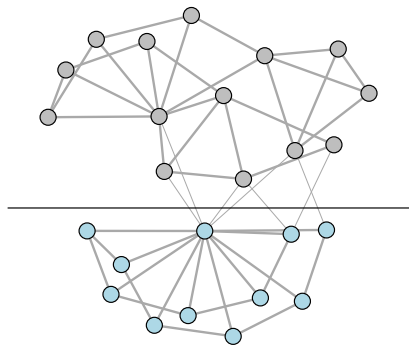
- What we do:
 - Develop redistricting simulation algorithms and other related methods
 - Apply these algorithms and methods to evaluate redistricting plans
 - Build open-source software packages: [redist](#) and [redistmetrics](#)
- A technological solution to detecting gerrymandering
- Empowering researchers, policy makers, data journalists, and citizen data scientists

Why Use Simulation Algorithm for Redistricting Evaluation?

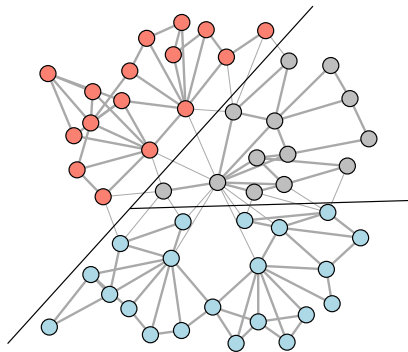
- Traditional redistricting evaluation
 - ① compute various fairness metrics
 - ② compare them across states and over time
- Confounded by differences in political geography and redistricting rules
- Simulation-based redistricting evaluation
 - ① generate many **alternative plans** under a set of redistricting criteria
 - ② compare them with a proposed plan to evaluate its properties
- Benefits of simulation approach
 - ① can control for **state-specific** political geography and redistricting rules
 - ② **transparency** and ability to isolate a relevant factor
 - ③ mathematical properties \rightsquigarrow **representative sample** of alternative plans

Redistricting as a Balanced Graph Partition Problem

25 Precincts, Two Districts



50 Precincts, Three Districts

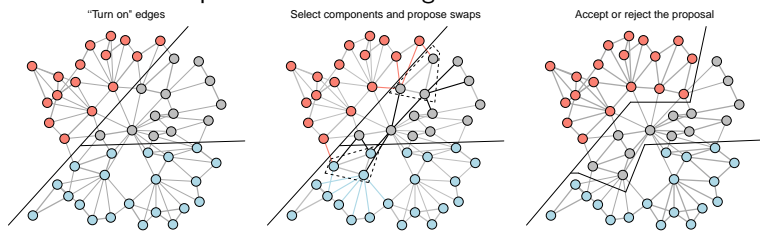


- Efficient **enumeration algorithm** exists based on zero-suppressed binary decision diagram (ZDD; Fifield *et al.* 2020)
- Only applicable for a very small redistricting problem

Markov chain Monte Carlo (MCMC) Algorithms

1 Flip algorithms (Fifield *et al.*, 2020; Mattingly & Vaughn, 2014; Chikina *et al.* 2017)

- start with a valid plan and then reassign units on district boundaries



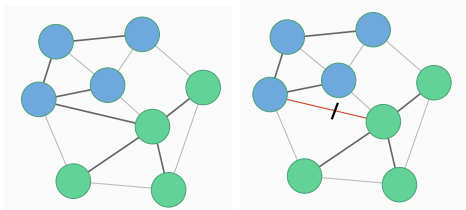
- target distribution

$$\pi(\xi) \propto \underbrace{\exp(-J(\xi))}_{\text{custom constraints}} \times \underbrace{1_{\xi \text{ connected}}}_{\text{contiguity requirement}} \times \underbrace{1_{\text{dev}(\xi) \leq D}}_{\text{population balance}}$$

- incremental changes; applicable for local exploration
- does not scale; compactness needs to be specified in $J(\cdot)$

2 Merge-split algorithms (DeFord *et al.*, 2021; Carter *et al.* 2019)

- randomly choose a pair of adjacent districts, merge them, and split them into two new districts using uniform spanning trees



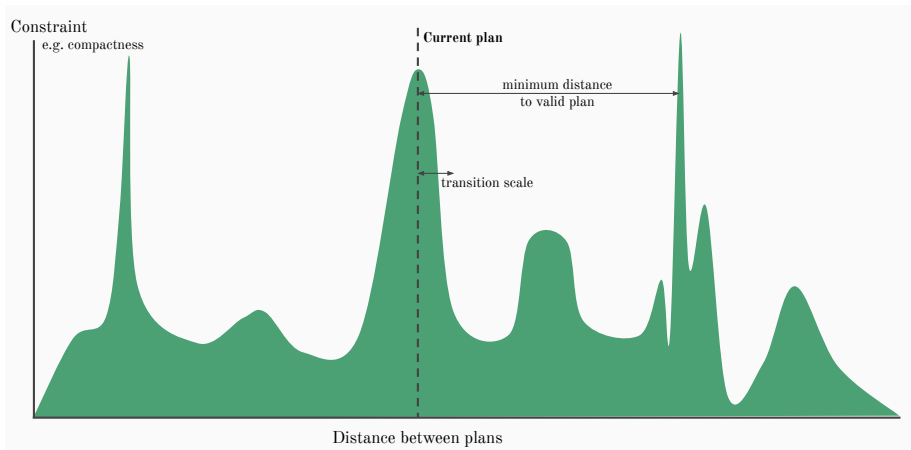
- target distribution

$$\pi(\xi) \propto \tau(\xi)^p \exp(-J(\xi)) \times \mathbf{1}_{\xi \text{ connected}} \times \mathbf{1}_{\text{dev}(\xi) \leq D}$$

where $\tau(\xi)$ counts the product of the number of spanning trees in each district of the plan ξ

- much improved mixing

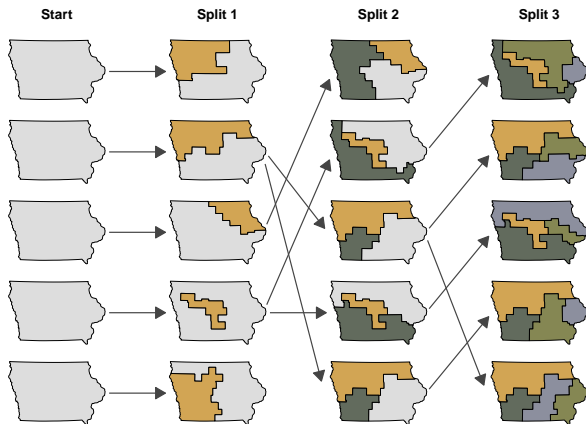
Challenge of MCMC Algorithms



- simulated annealing, parallel tempering \rightsquigarrow difficult to apply in practice

Sequential Monte Carlo (SMC) Algorithm (McCartan and Imai, 2020)

- Start with a blank state but in parallel, use the spanning approach to sample a district at a time, resample at each step with weights

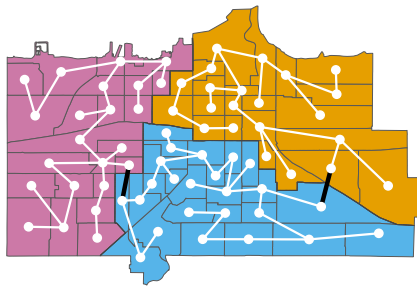
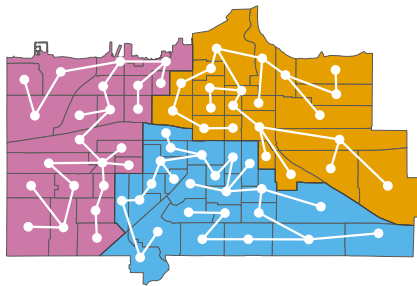
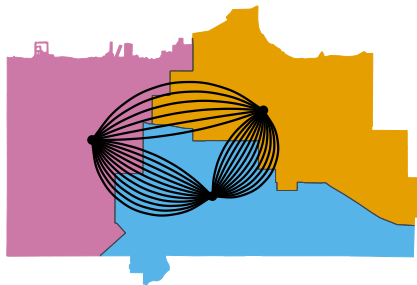
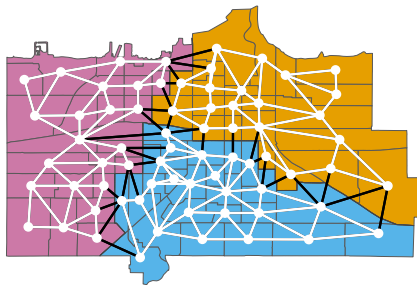


- sampled plans are nearly independent

The SMC Algorithm in Summary

- Algorithm 1: Splitting procedure
 - 1 Generate a uniform spanning tree (Wilson's algorithm)
 - 2 Sort edges by population deviation
 - 3 Sample one edge from top k_i edges and remove it
 - 4 Check population bounds
- Algorithm 2: SMC with partial rejection control
 - 1 Generate S initial copies of map; set all weights to 1
 - 2 For $i \in \{1, 2, \dots, n - 1\}$:
 - a. Until there are S successes
 - i. Sample a map according to the weights
 - ii. Use Algorithm 1 to split off a new district from each sampled map
 - iii. Reject if population bounds are not met
 - b. Calculate new weights based on splitting probability
 - 3 Output complete plans and compute final weights
- Limitation: hard to incorporate plan-wide or region-specific constraints

Avoiding County Splits



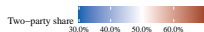
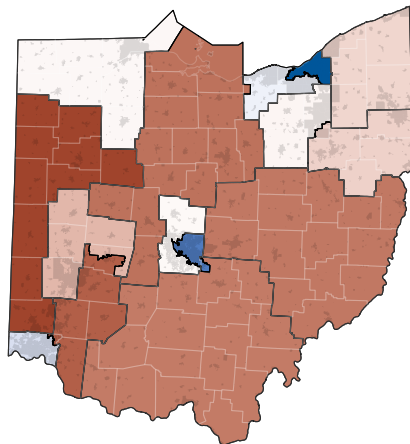
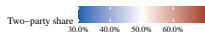
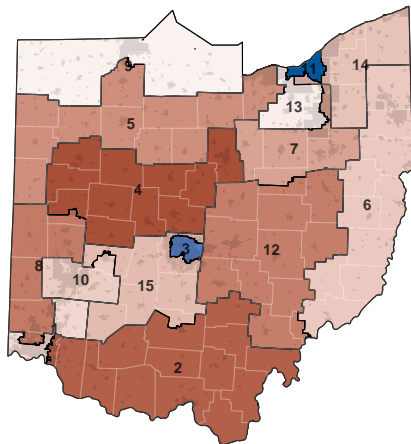
Application: Ohio Congressional Redistricting

- Currently 16 districts: 4 Democrats and 12 Republicans
 - 2020 President: Biden 45%, Trump 53%
 - 2018 Senate: Brown 53%, Renacci 47%
- After 2020 Census, the number of seats is reduced to 15 districts
- 2018 Ohio voters passed the constitutional amendment
 - bipartisan support leads to a 10 year map
 - if that fails, it becomes a 4 year map
- Redistricting
 - State Senate and House approved the initial map
 - No bipartisan support \rightsquigarrow 4 year map
 - November 20: Governor DeWine signed the map

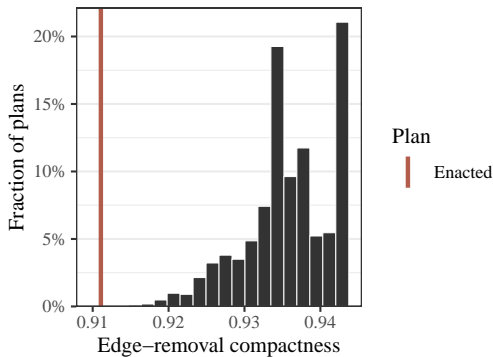
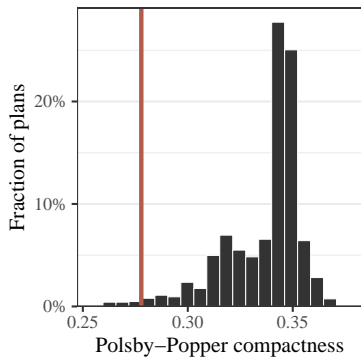
League of Women Voters of Ohio *et al.* v. Ohio Redistricting Commission, *et al.*

- I served as an expert witness for Relators
- Simulation analysis based on Sequential Monte Carlo algorithm
 - 5,000 alternative plans
 - contiguous and compact districts
 - compliant with the Voting Rights Act (Cleveland)
 - several complicated splitting constraints
 - Section 2(B)(5): out of Ohio's 88 counties,
 - at least 65 counties should not be split
 - no more than 18 counties can be split no more than once
 - no more than 5 counties can be split no more than twice

The Enacted and Example Simulated Plans



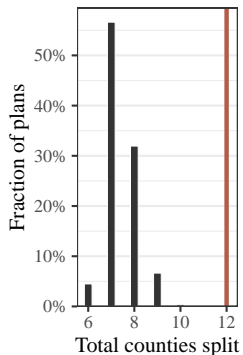
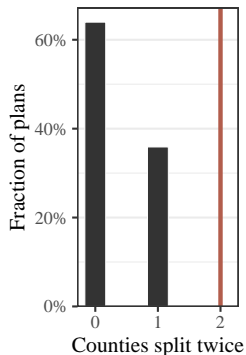
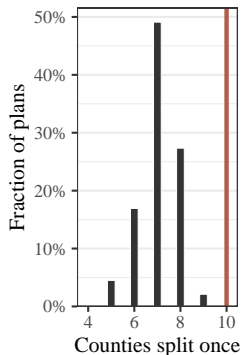
Compactness



- Polsby-Popper: the ratio of the district area to the area of a circle with the same perimeter
- Edge-removal: highly correlated with $\log \tau(G) - \log \tau(\xi)$

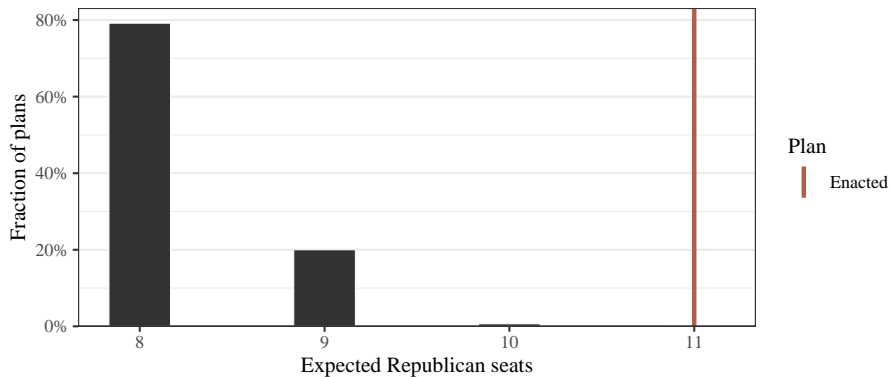
$$\tau(\xi)^{\rho} \propto \exp(-C\rho\text{rem}(\xi))$$

Administrative Boundary Splits



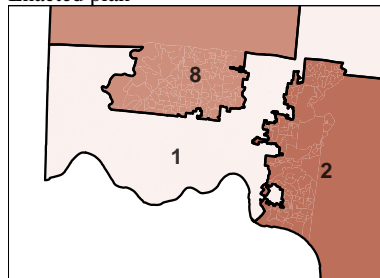
Plan
Enacted

Expected Number of Republican Seats

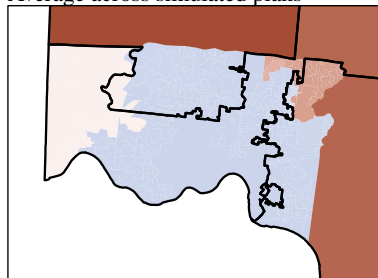


Hamilton County: Cincinnati Area

Enacted plan



Average across simulated plans

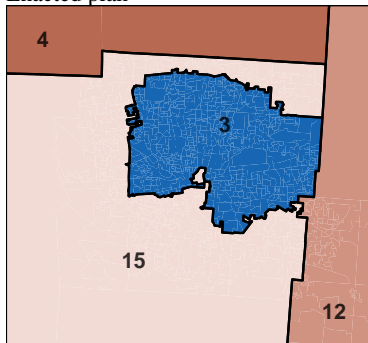


Two-party
vote share

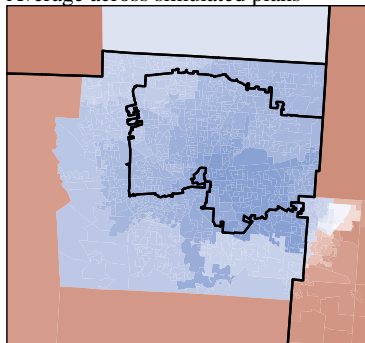


Franklin County: Columbus Area

Enacted plan



Average across simulated plans

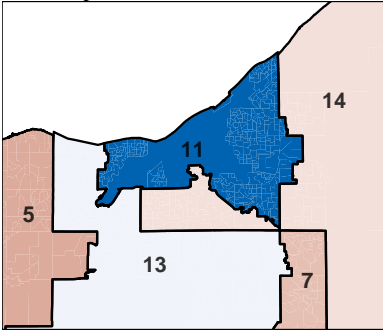


Two-party
vote share

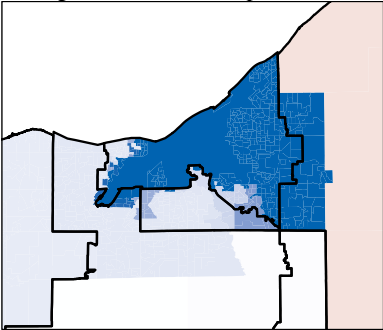


Cuyahoga County: Cleveland Area

Enacted plan



Average across simulated plans



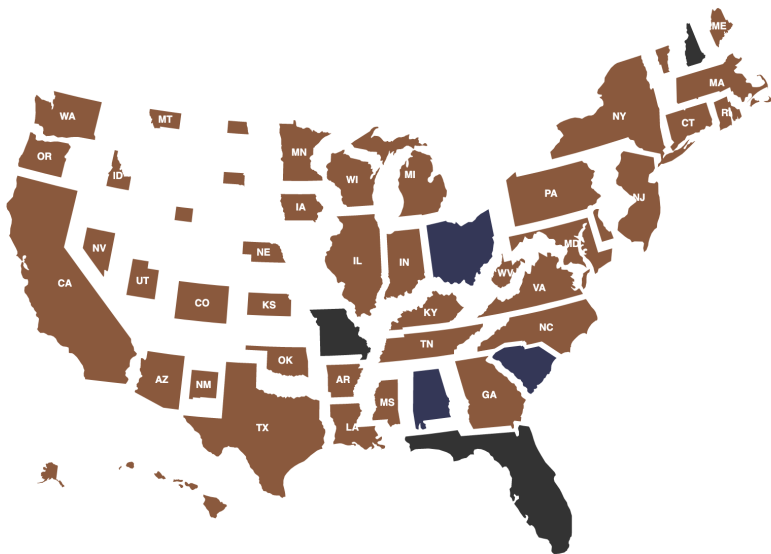
Ohio Supreme Court Strikes Down the Enacted Map



The Court Opinion

Id. at Section 1(C)(3)(a). The above evidence, particularly Dr. Imai's conclusion that the enacted plan will result in, on average, 2.8 more Republican seats than are warranted, shows that the General Assembly's decision to shift what could have been—under a neutral application of Article XIX—Democratic-leaning areas into competitive districts, i.e., districts that give the Republican Party's candidates a better chance of winning than they would otherwise have had in a more compactly drawn district, resulted in a plan that unduly favors the Republican Party and unduly disfavors the Democratic Party.

50 State Redistricting Simulations Project



States colored **red** have enacted a congressional map and been fully analyzed, states colored **blue** have enacted a plan but haven't yet been analyzed, and states colored **gray** haven't enacted a plan yet.

Concluding Remarks

- Redistricting matters
 - fair representation and policy outcomes
 - competitiveness of districts and responsiveness
 - political polarization
- How should we stop gerrymandering?
 - independent commission (e.g., Michigan)
 - use of algorithms to detect gerrymandering
- Roles of experts
 - legislative process
 - court testimony
- Open problems
 - large-scale redistricting problems (e.g., state legislatures)
 - redistricting plans based on Census blocks
 - algorithm-generated redistricting plan proposals