Using Algorithms to Detect Gerrymandering and Improve Legislative Redistricting: Cases from the United States and Japan

Kosuke Imai

Harvard University

Equity and Algorithms Series, Digital Scholarship Support Group
Japan Forum, Reischauer Institute of Japanese Studies, Harvard University
November 3, 2023

Joint work with Christopher Kenny, Cory McCartan, Tyler Simko, Shiro Kuriwaki, George Garcia III, Sho Miyazaki, Kento Yamada, Kevin Wang, and Melissa Wu
Motivation

Today’s world for quantitative social science:
1. increasing availability of granular data
2. rapid methodological advancement

Social scientists can and should solve problems of the real world!

Redistricting as a major policy decision

How can we use data and algorithms to evaluate redistricting plans?
- traditional methods: comparison across states and time periods
- confounded by state-specific political geography and rules

Use of simulation algorithms
1. obtain a representative sample of redistricting plans under constraints
2. compare the enacted plan with this baseline distribution

Technological solution to detecting gerrymandering

Tool for analyzing redistricting
What we do:

1. develop efficient and flexible simulation algorithms
2. build open-source software packages for the entire workflow
3. evaluate redistricting plans in the United States and elsewhere

Goal: empower researchers, policy makers, data journalists, and citizen data scientists with powerful tools
Redistricting Basics

- Classic gerrymandering strategies: packing and cracking

  ![Even distribution](image)
  ![Packing](image)
  ![Cracking](image)

- What has changed:
  - availability of granular data
  - mapping software (e.g., Maptitude, Dave’s Redistricting app)

- US Congressional redistricting
  - racial gerrymandering: *Allen v. Milligan*
  - partisan gerrymandering: *Rucho v. Common Cause*
**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 $\leadsto$ representative sample of alternative plans
Sequential Monte Carlo (SMC) Algorithm (McCartan and Imai, 2020)

- Start with a blank state in parallel, sample a district at a time, resample with weights at each step

![Diagram showing the SMC algorithm process](image-url)
50 State Redistricting Simulations Project

Comprehensive project to simulate alternative congressional redistricting plans for all fifty states.

- tidied 2020 Census plus statewide election data from the VEST
- collect state-specific redistricting requirements
- construct algorithmic constraints based on these and traditional redistricting criteria
- 5,000 simulation plans based on SMC
- code and data are available at the Harvard Dataverse
Georgia Example

- 14 Congressional districts
- According to Georgia’s House Legislative and Congressional Reapportionment Committee, districts must:
  1. be contiguous
  2. have equal populations
  3. be geographically compact
  4. preserve county and municipality boundaries as much as possible
  5. avoid the unnecessary pairing of incumbents
- We attempted to account for everything except incumbency constraint
- Voting rights act (VRA) compliance is tricky
Widespread Partisan Gerrymandering Cancels Nationally
Map of Partisan Gerrymandering
Partisan Gerrymandering Reduces Competitiveness

The graph shows the comparison between non-partisan simulations and the enacted plan. The x-axis represents the average Democratic district vote share, while the y-axis indicates the number of seats. The graph illustrates how the enacted plan deviates from the non-partisan simulations, particularly in the distribution of seats based on the average Democratic vote share.
Application in the Court: Ohio Congressional Redistricting

- Currently 16 districts: 4 Democrats and 12 Republicans
- After 2020 Census, the number of seats is reduced to 15 districts
- 2018 Ohio voters passed the constitutional amendment


Simulation analysis
- 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
- Polsby-Popper: the ratio of the district area to the area of a circle with the same perimeter
- Edge-removal
Administrative Boundary Splits

Counties split once

Counties split twice

Total counties split

Plan

Enacted
Expected Number of Republican Seats

![Bar chart showing the expected number of Republican seats for different fractions of plans. The x-axis represents the expected Republican seats (8 to 11), and the y-axis represents the fraction of plans (0% to 80%). The chart shows a high concentration of plans with 8 expected Republican seats.]
Packing: Franklin County (Columbus Area)

Enacted plan

Average across simulated plans

Two-party vote share

30%
40%
50%
60%
Ohio Supreme Court Strikes Down the Enacted Map
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.
United States Supreme Court: Alexander v. NAACP et al.

- South Carolina racial gerrymandering case argued on Oct 11, 2023
- Served as an expert witness for the plaintiffs
- Used simulation to provide evidence that a disproportionately large number of Black voters are packed into District 6

Justice Alito: Did Dr. Imai run a simulation using the political data as well but then decide to shelve it when the results were not favorable to your client?

Ms. Aden: That is not what I believe the record reflects, Your Honor.

Justice Alito: It just never occurred to him that politics might have something to do with this?
Is There Partisan Bias in Japanese Redistricting?

- Non-partisan commission ⇛ no partisan bias?
- Potential sources of partisan bias
  - members are appointed by the prime minister and approved by the Diet
  - governors are invited to provide their opinions
- 2020 Japanese redistricting
  - redistricting in 25 prefectures out of 47
  - 10 prefectures lost a seat
  - 5 prefectures gained a seat / seats
  - 5 prefectures redrew districts without changing the number of seats
Little Partisan Bias at the Prefecture Level

Difference in the number of districts with more than 50% of normal votes won by the ruling coalition

Favors Ruling Parties → Favors Opposition Parties

Increased Seat(s) Decreased Seat No Change in Number of Seats

Prefectures, ordered by seat difference

- Chiba (14)
- Hyogo (12)
- Hokkaido (12)
- Kanagawa (20)
- Miyagi (5)
- Gifu (5)
- Niigata (5)
- Ibaraki (7)
- Ehime (3)
- Fukuoka (11)
- Saitama (16)
- Aichi (16)
- Shiga (3)
- Yamaguchi (3)
- Wakayama (2)
- Tochigi (5)
- Shizuoka (8)
- Shimane (2)
- Osaka (19)
- Okayama (4)
- Nagasaki (3)
- Hiroshima (6)
- Gunma (5)
- Fukushima (4)
- Fukuoka (11)
- Ehime (3)
- Ibaraki (7)
- Niigata (5)
- Gifu (5)
- Miyagi (5)
- Kanagawa (20)
- Hokkaido (12)
- Hyogo (12)
- Chiba (14)
Some but relatively little partisan bias at the district level.

- Tokyo (30)
- Kanagawa (20)
- Osaka (19)
- Aichi (16)
- Saitama (16)
- Chiba (14)
- Hokkaido (12)
- Hyogo (12)
- Fukuoka (11)
- Shizuoka (8)
- Ibaraki (7)
- Hiroshima (6)
- Gifu (5)
- Gunma (5)
- Miyagi (5)
- Niigata (5)
- Ehime (3)
- Nagasaki (3)
- Shiga (3)
- Okayama (4)
- Fukuoka (4)
- Osaka (5)
- Tochigi (5)
- Kagoshima (5)
- Okayama (4)
- Shizuoka (8)
- Fukuoka (11)
- Hyogo (12)
- Hokkaido (12)
- Chiba (14)
- Saitama (16)
- Aichi (16)
- Osaka (19)
- Kanagawa (20)
- Tokyo (30)

Z-score of the enacted plan on the vote share of the ruling parties, compared to the simulated plans, by ordered districts.
Miyagi Prefecture

- # of seats: 6 → 5
- # of municipality splits: 2 → 0
- # of county splits: 2 → 0
- Population deviation: 1.94 → 1.64

Lower House Electoral results
- 2017: 5 LDP, 1 independent
- 2021: 4 LDP, 2 opposition

- 5,000 simulated plans
Simulated Plans Have More Desirable Properties

Municipality Splits

County Splits

Max–min ratio

Fraction of plans

Density

Max–min ratio

Municipality Splits

County Splits

Max–min ratio
Partisan Bias

The image contains a box plot and a map. The box plot shows the distribution of the share of the ruling coalition across ordered districts, with ordered districts ranging from 1 to 5. The map is color-coded to represent the share of the ruling coalition, with Sendai marked as a specific location.
Concluding Remarks

- Redistricting matters
  - fair representation and policy outcomes
  - competitiveness of districts and responsiveness
  - political polarization
  - state and local offices, education districts, non-US contexts

- How should we stop gerrymandering?
  - independent commission (e.g., Michigan)
  - use of algorithms to detect gerrymandering

- Role of experts
  - legislative process
  - court testimony
  - work with non-partisan groups and commissions

- Open problems
  - large-scale redistricting problems (e.g., state legislatures)
  - algorithm-generated redistricting plan proposals
  - communities of interest, impact of redistricting rules