Èdouard Lucas:

The theory of recurrent sequences is an inexhaustible mine which contains all the properties of numbers; by calculating the successive terms of such sequences, decomposing them into their prime factors and seeking out by experimentation the laws of appearance and reproduction of the prime numbers, one can advance in a systematic manner the study of the properties of numbers and their application to all branches of mathematics.

Computational Approaches for Political Redistricting
Part IV: Specific State Analyses

Daryl DeFord

CSAIL – GDP Group

IAP 2019
Massachusetts Institute of Technology
January 29, 2019

MORAL:

Computational Redistricting is
NOT a solved problem!

Advertisements

1. VRDI – 6 week summer program for graduate and undergraduate students (Deadline 2/1)
   - Application: tinyurl.com/apply-vrdi-2
   - Information: gerrydata.org
2. Contact:
   - Email: ddeford at mit.edu
   - Website: mggg.org
   - Slack channel: GerryChat.slack.com
3. Research Projects
   - Math Problems: tinyurl.com/gerryprojects
   - Data Problems: tinyurl.com/GerryChainProjects
4. IAP Info:
   - Resources: people.csail.mit.edu/ddeford/CAPR
   - Today 12-1 In–depth state examples

Outline

Compactness Measures
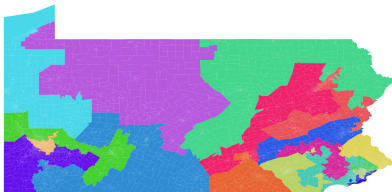
Polsby–Popper

Theorem (Isoperimetry)

Let Ω be a bounded open subset of $\mathbb{R}^2$ with finite perimeter. Then:

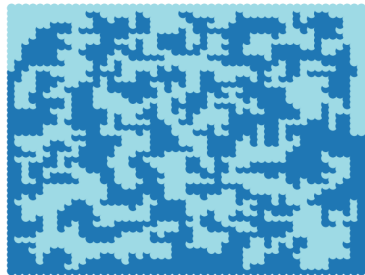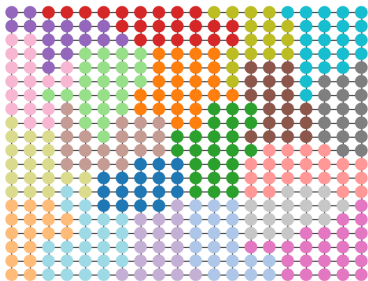$$4\pi A \leq P^2$$

Definition (Polsby–Popper)

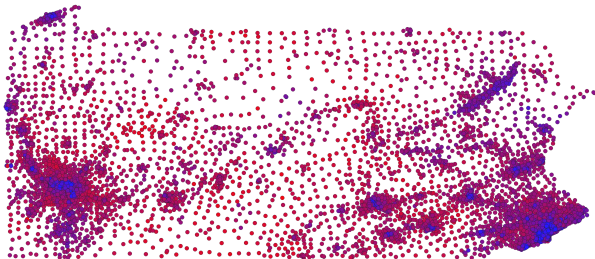The Polsby–Popper score of a district is:

$$PP(\Omega) = \frac{4\pi A}{P^2}$$

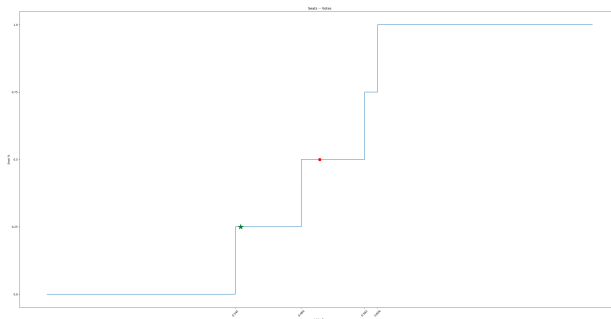# (Discrete) Total Perimeter

Partisanship Measures

Seats–Votes Curves

Figure: Dem %: [.249,.389,.273,.51]
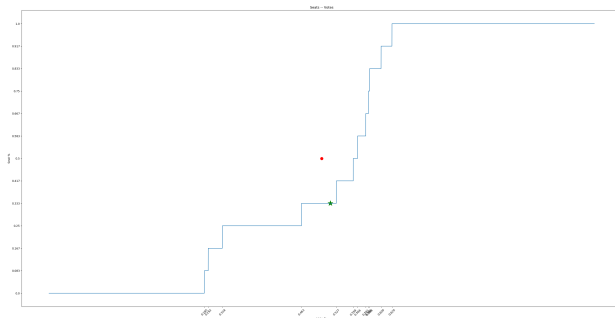
Seats–Votes Curves

Figure: Dem %: [.698,.458,.724,.43,.435,.428,.553,.489,.407,.387,.731,.45]
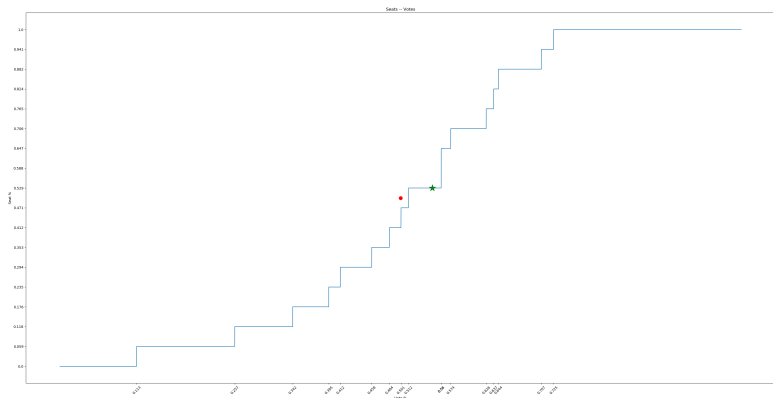
Seats–Votes Curves

Figure: Dem %:
[.487,.79,.934,.635,.652,.589,.535,.546,.403,.487,.41,.34,.705,.421,.322,.473,.563]

Partisan Metrics

Definition (Mean–Median)

Horizontal distance between $(.5, .5)$ and the seats votes curve).
Alternatively, difference between the median and mean of the votes vector.

Definition (Partisan Bias)

Vertical distance between $(.5, .5)$ and the seats votes curve

Definition (Partisan Asymmetry)

Integral of the difference between the seats votes curve and its reflection
around $(.5, .5)$.

Definition (Efficiency Gap)

$$\frac{W^A - W^B}{T}$$

With equal turnout: twice the seat margin minus the vote margin.

Partisan Examples
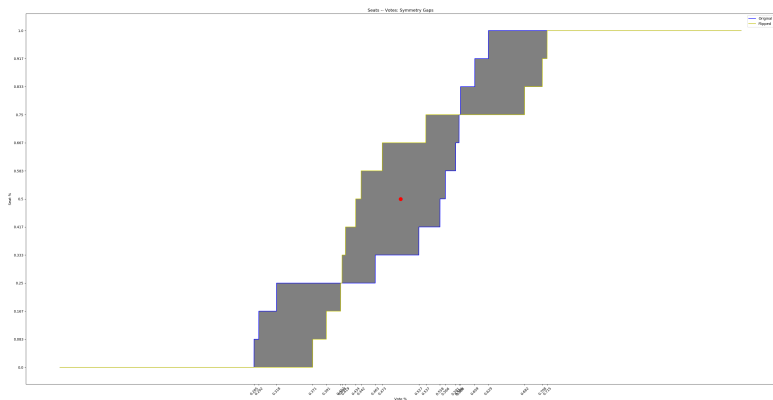
- Utah
  - Mean–Median: -.024
  - Efficiency Gap: -.039
  - Asymmetry: .048
- Pennsylvania
  - Mean–Median: .011
  - Efficiency Gap: .063
  - Asymmetry: .050
- North Carolina
  - Mean–Median: .062
  - Efficiency Gap: .198
  - Asymmetry: .093
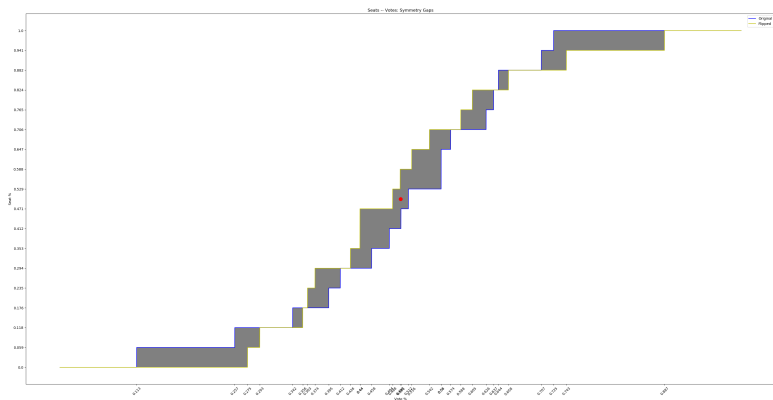
Seats–Votes Asymmetry
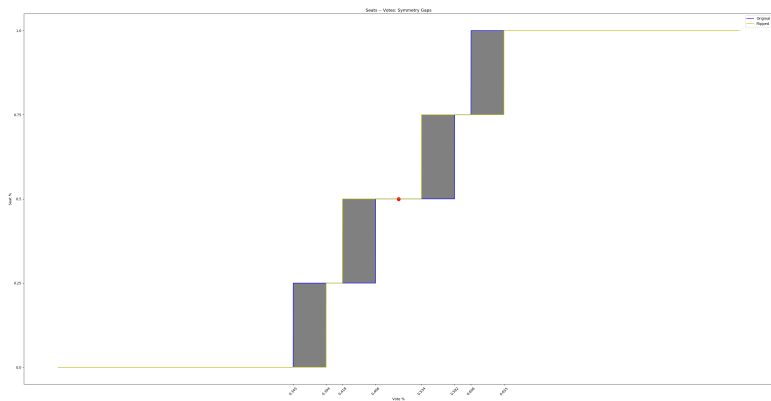
Seats–Votes Asymmetry

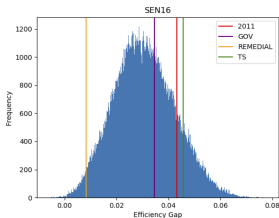# Seats–Votes Asymmetry

Court History

2011

538 GOP

538 Dem

8th Grade

538 Compact

Gov

Remedial

TS

Data Setup

- Choice of units: Precincts
- Voting Data: Weighted Senate Results
- Initial Plans: TS/GOV/Enacted
- Demographics: Census

Theoretical Tool

Theorem (CFP[1])

Let $M = X_0, X_1, X_2, \ldots$ be a reversible Markov chain with a stationary distribution π on the state space Ω and let $G : \Omega \to \mathbb{R}$ be a ranking function. If $X_0 \sim \pi$, then for any fixed k, the probability that $G(X_0)$ is an ε–outlier from among the list of values $G(X_0), G(X_1), G(X_2), \ldots, G(X_k)$ is at most $\sqrt{2\varepsilon}$.

[1] M. Chikina, A. Frieze, and W. Pegden: Assessing significance in a Markov chain without mixing, PNAS, (2017).

Experimental Setup

- In order to apply the theorem we need a reversible chain
  - Uniform over permissible
  - Metropolis–Hastings
- Lots of choices to make!

Compactness

Figure: Average of reciprocal Polsby–Popper scores

Other Constraints

- Population Balance:
  - All within 1%
  - MH – L2
- VRA Compliance
  - Preserve two Philadelphia regions
  - Geoclusters more broadly
- County Splits:
  - Entropy of Counties split Districts
  - Entropy of Districts split Counties

Other Constraints

- Population Balance:
  - All within 1%
  - MH – L2
- VRA Compliance
  - Preserve two Philadelphia regions
  - Geoclusters more broadly
- County Splits:
  - Entropy of Counties split Districts
  - Entropy of Districts split Counties
- Glue together with linear coefficients for MH weighting

Initial Results

Gerrymandering Detected!

Problem Reformulation

- Move from Local to Global MCMC steps
- Reformulate constraints
  - Weaker population bound
  - Discrete instead of continuous compactness
  - Tree steps instead of single edge flip
  - Many seeds instead of single long walk
  - More election data
  - More comparison plans

Problem Reformulation

- Move from Local to Global MCMC steps
- Reformulate constraints
  - Weaker population bound
  - Discrete instead of continuous compactness
  - Tree steps instead of single edge flip
  - Many seeds instead of single long walk
  - More election data
  - More comparison plans
- Better Geography?

Pennsylvania Landscapes
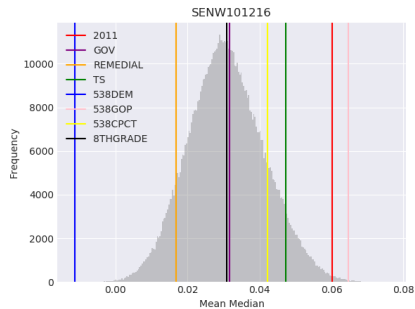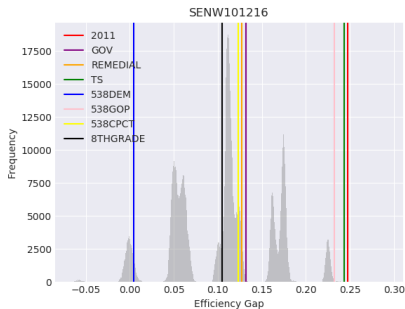
Initial Results

Gerrymandering Detected?

Legal History
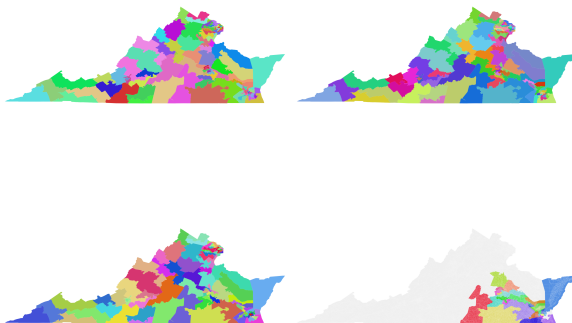
Data Collection

- Choice of units: Blocks
- Voting Data: Prorated Presidential Results
- Initial Plans: (rounded) Enacted/Rep/Dem/Princeton
- Demographics: Virginia DLS

Tree Seeds Ensemble

Research Question?

- What is best possible representation?
- What are impacts of packing?
- Is it possible to avoid packing?
- Partisan consequences of unpacking?

Proposed Plans
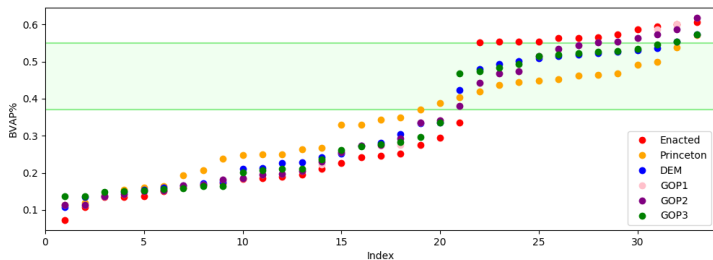
Seed Comparison
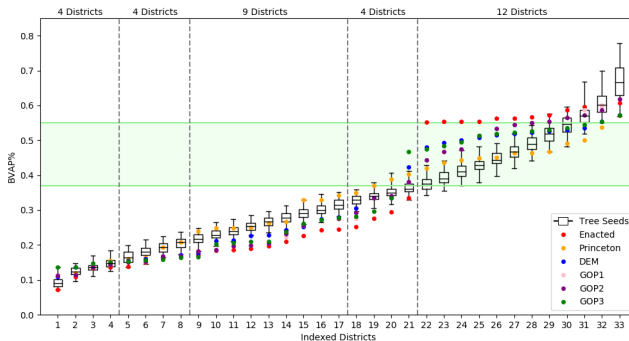
Full State Possibilities
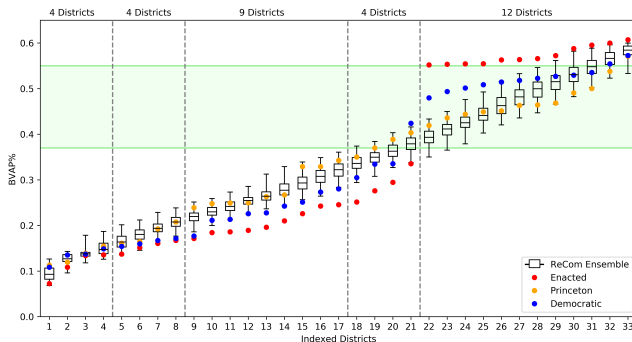
Full State Possibilities

The 37% line

BVAP% by CBC Membership

- Rep in CBC
- Rep not in CBC
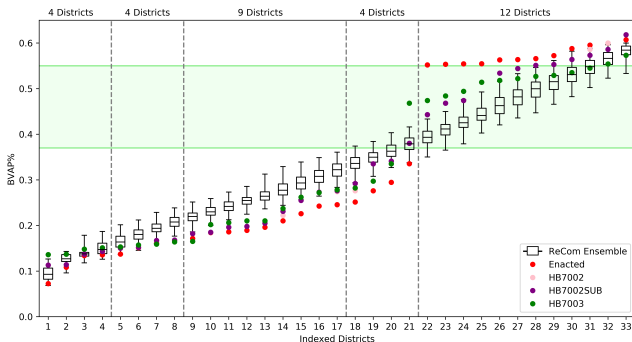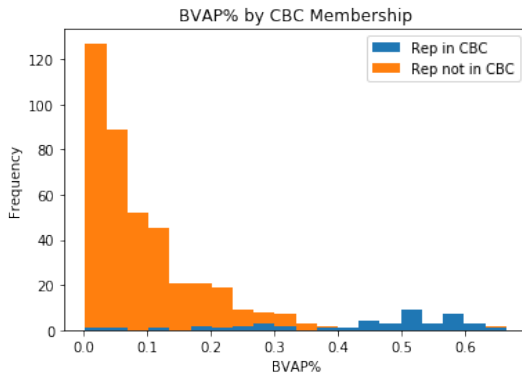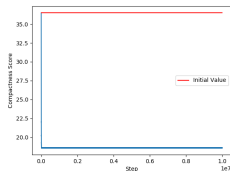
Compactness

(a) Single Edge
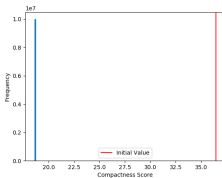
(b) Single Edge
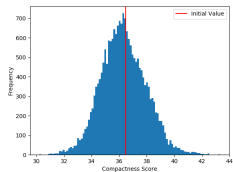
(c) Tree Walk

Compactness

Initial Results

Gerrymandering Detected?

Legal History
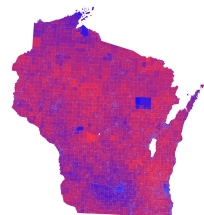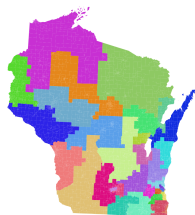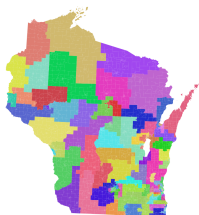
Competitiveness

Definition (Competitive)

At least $x\%$ of districts are within $y\%$ of z. Where $z \in \{.5,$ state mean, state median, plan mean, plan median $\}$

Data Collection

- Choice of units: Wards
- Voting Data: All Statewide
- Initial Plans: Enacted plans
- Demographics: Census

Senate Seats
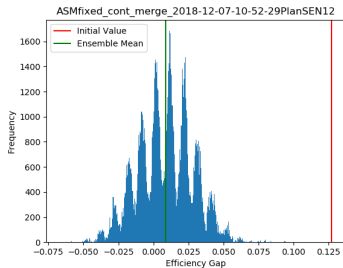
(a) SEN12
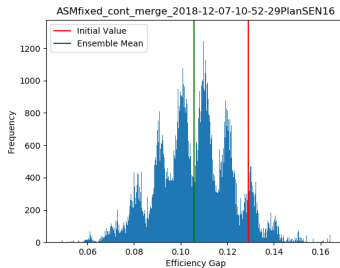
(b) SEN16

Senate Mean–Median
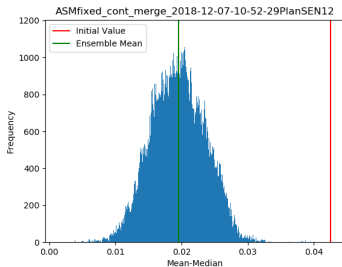
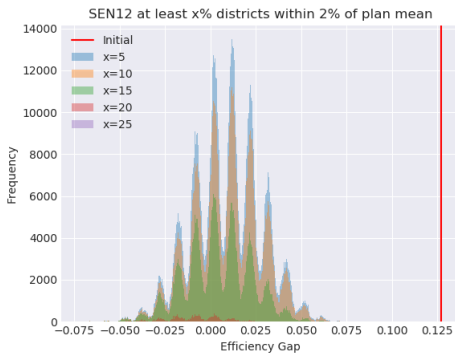(a) SEN12
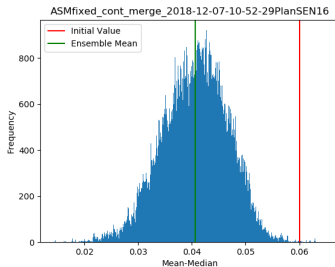
(b) SEN16
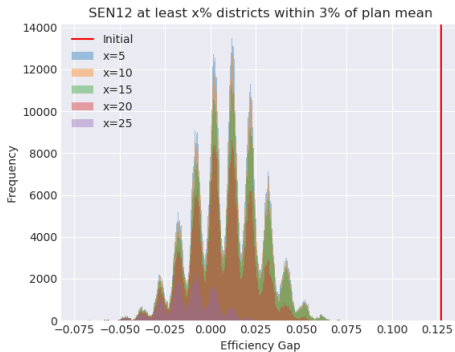
Senate Efficiency Gap
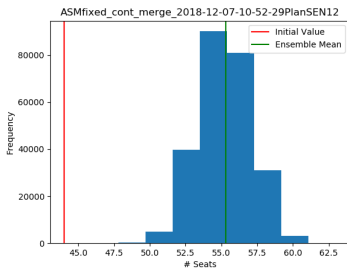
(a) SEN12
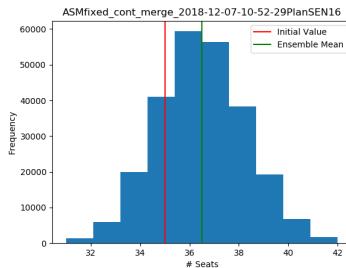
(b) SEN16

Senate 12 Box Plots
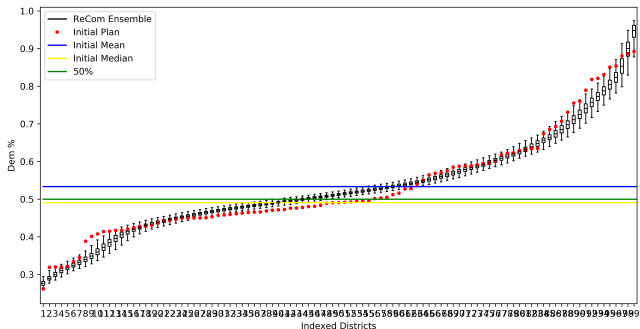
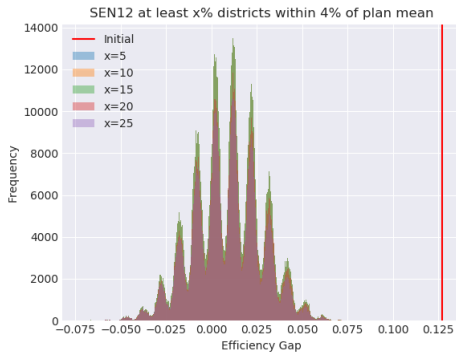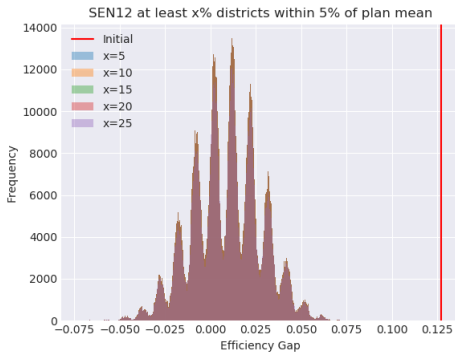# Senate 16 Box Plots
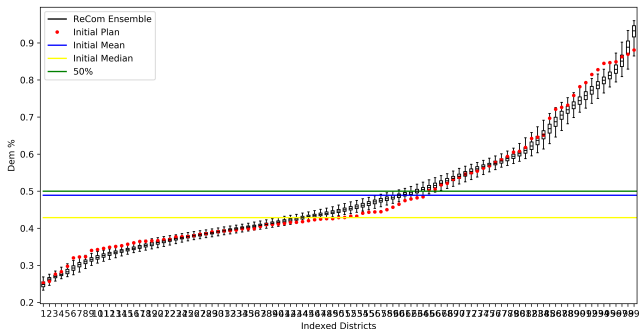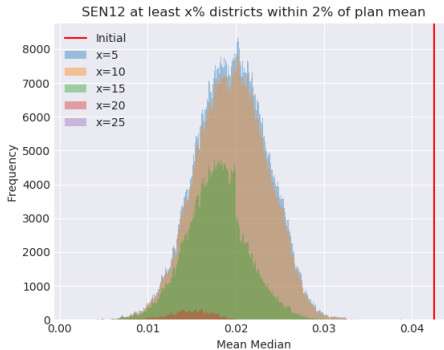
Senate 12 EG Competitiveness

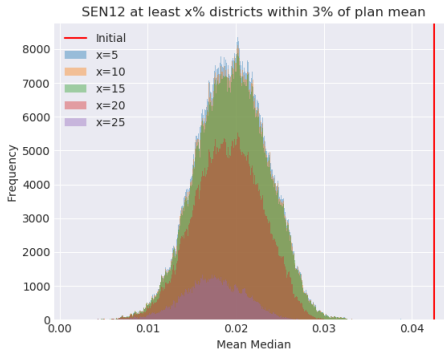# Senate 12 EG Competitiveness
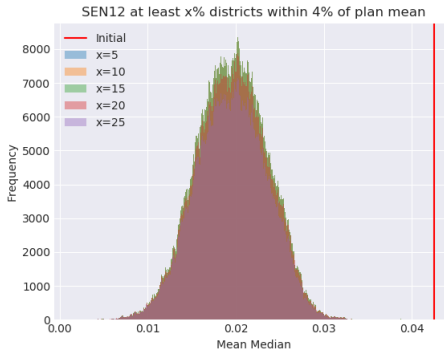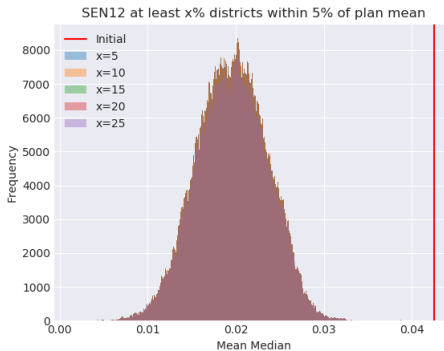
Senate 12 EG Competitiveness

Senate 12 EG Competitiveness

SEN12 at least x% districts within 5% of plan mean

Senate 12 MM Competitiveness

SEN12 at least x% districts within 2% of plan mean

Senate 12 MM Competitiveness

SEN12 at least x% districts within 3% of plan mean

Senate 12 MM Competitiveness

SEN12 at least x% districts within 4% of plan mean

Senate 12 MM Competitiveness

SEN12 at least x% districts within 5% of plan mean

Initial Results

Gerrymandering Detected?

Legal History

- Pressure for algorithmic version of municpality preservation
- Substitute for compactness
- Allows weaker population bounds
- Enforces a multiscale viewpoint on the redistricting problem
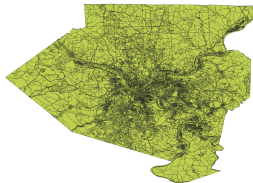
Levels of Resolution

- Blocks
- Precincts
- Wards
- Municipalities
- Counties

Blocks

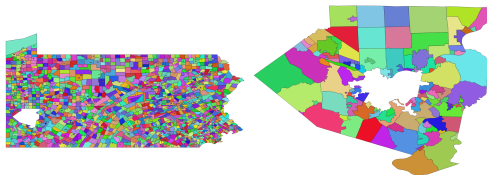(a) Pennsylvania   (b) Allegheny   (c) Philadelphia
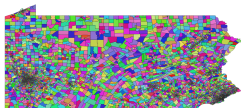
Counties

Municipalities

(a) Pennsylvania

(b) Allegheny

Precincts
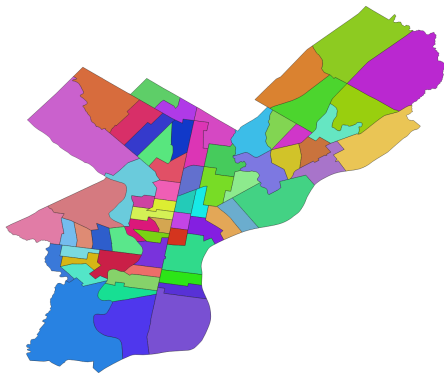

(a) Pennsylvania

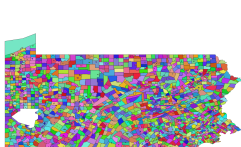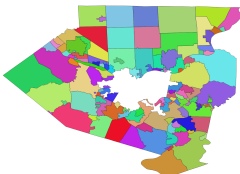(b) Pittsburgh

(c) Philadelphia

Wards

Putting Them Together

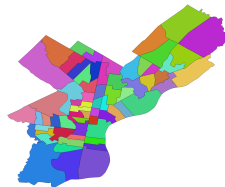(a) Pennsylvania   (b) Allegheny   (c) Pittsburgh
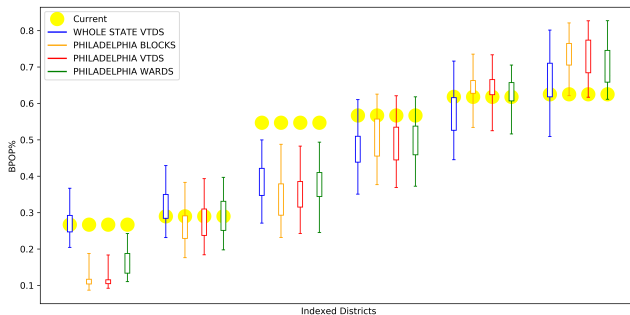
Putting Them Together

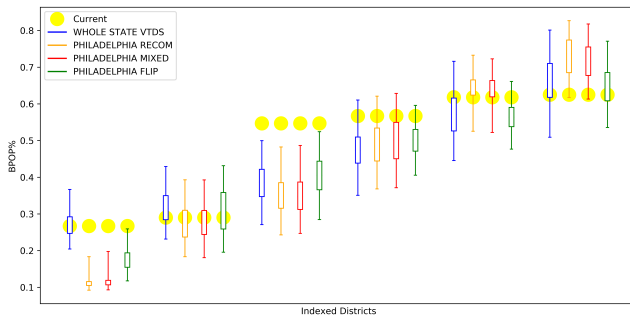
(a) Blocks

(b) Precincts

(c) Wards

Comparison of BPOP Districts

Comparison of BPOP Districts

Initial Results

Better Representation Possible?

MORAL:

Computational Redistricting is NOT a solved problem!

The End

Thanks!