Comparing Methods of Opportunistic Risk-Limiting Audits

Eric Chen, Rohith Raghavan
Mentored by Mayuri Sridhar

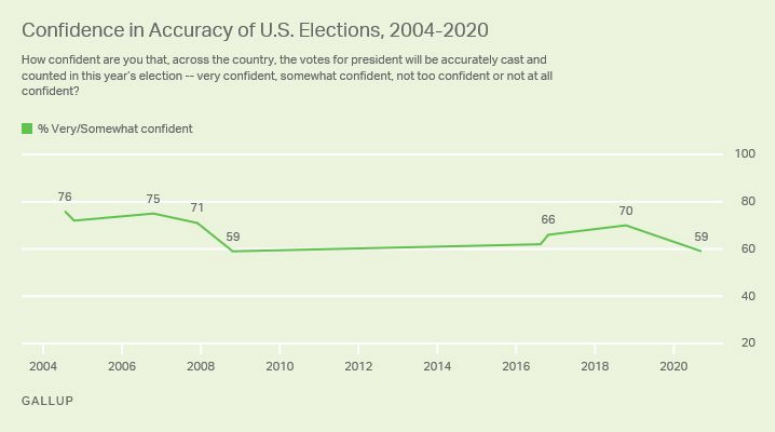
Introduction

What Are Audits?

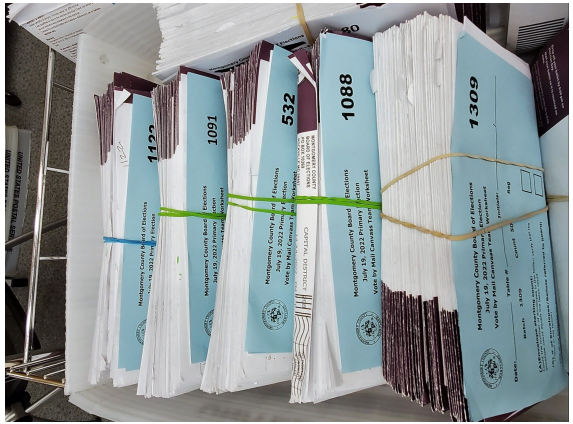
- Reviews of election results from a third party
- Allow for full recount if necessary
- Detect both sabotage and mistakes

Why Run Audits?

- Controversy over election results
- Catches both human and mechanical errors



How Are Audits Run?



- Simplest case: full recount
  - Guarantees accuracy
  - Extremely expensive
- We want an estimate of accuracy but lower ballots required

Introduction to Risk-Limiting Audits

- Instead of full recounts, we only select a sample of ballots
- Compare these ballots to the results to determine accuracy of the election
  - Only determines if outcome is correct, not margin
- Stop once the risk is “low enough”

Risk Limiting Audits

What is a risk-limiting audit?

- Statistical test that uses samples to assign a risk limit
- Audits affirm results when risk limit is met
  - Cannot reject results, only call for a full hand count(escalation)
- Real results of election are results from full hand count, as opposed to reported results

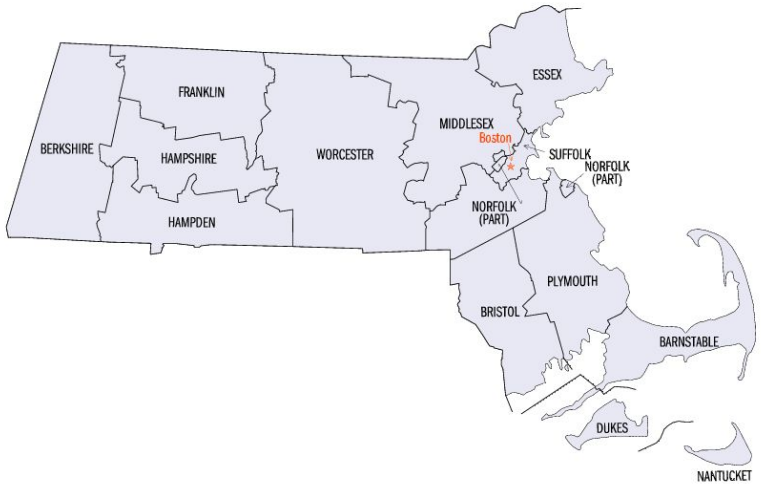
Purpose and Benefits of RLAs

- RLAs attempt to minimize the work done to audit a election, the minimum number of ballots sampled
  - Use a strategy that avoids escalation as much as possible
- RLAs rely only on the proportions of votes, not the number of votes
  - Scales well to larger elections without losing confidence

Stratification

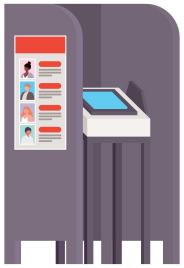
Strata

- Strata are smaller groups which divide the population
  - Different collections of ballots
- Potential strata
  - Voting centers
  - Towns, Counties, States
  - Mail-in and in-person ballots



Why stratify?



- Simple RLAs would be difficult to conduct on a wide scale
  - Need to combine audits from different locations
  - Need to pick a sample at random across whole election
- Different locations may be running audits in different ways
  - Example: replacement / no replacement

Challenges of stratification

- Risk from strata usually cannot be trivially combined
  - Introduces uncertainty, ruining audit
- Logistical challenges
- Using data from different polling software and strategies

Opportunistic Auditing

Opportunistic Auditing

- Elections normally consist of multiple contests at the same time
  - MA governor/senate race on the same ballot as presidential race
- Different ballots contain different races
- Opportunistic auditing = gathering audit information on multiple races at once
- Saves resources

Goal of Our Research

- We aim to create realistic methods of opportunistic auditing
- How can we efficiently audit the national level while also auditing state level contests?
- How do we choose states we want to audit?
- Should we audit the national level first or states first?

BRAVO

BRAVO (Ballot-polling Risk-limiting Audits)

- Ballot-polling and risk-limiting audit
- Works on simple plurality votes
- Follows the steps introduced earlier
  - Draw a ballot
  - Calculate risk
  - Check if risk is low enough to end the audit
  - Repeat until a full recount is necessary

How is risk calculated?

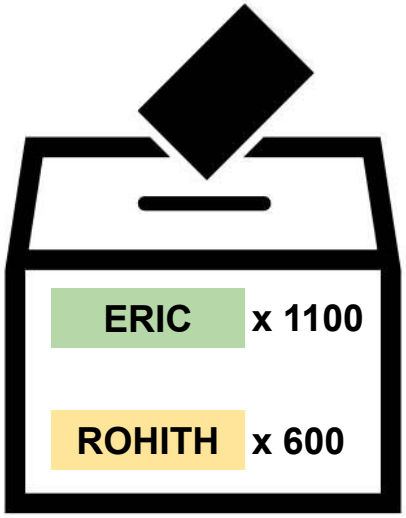
- Set the test statistic, T, as 1. 1/T is the risk
- Let s be the reported proportion of votes the winner received
- Select a valid ballot from the sample
- If the ballot is for the reported winner, multiply T by 2s
- Otherwise, multiply T by 2(1-s)

How is risk calculated?

- If 1/T < risk limit, the audit ends
- Return the ballot to the sample
- Repeat

A demonstration of BRAVO

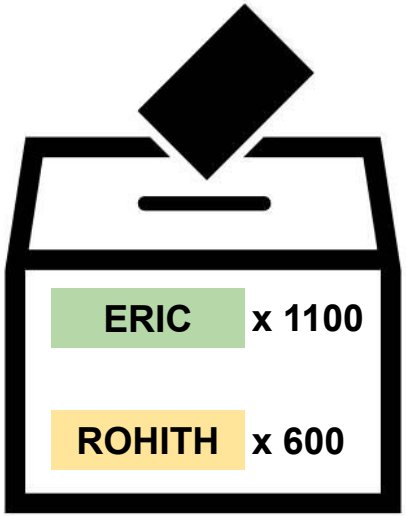
- Let's say we have an election with the following results
  - Votes for Eric: 1100, or 64.7% of the vote
  - Votes for Rohith: 600, or 35.3% of the vote



A demonstration of BRAVO

- We keep track of our current audit data in a table

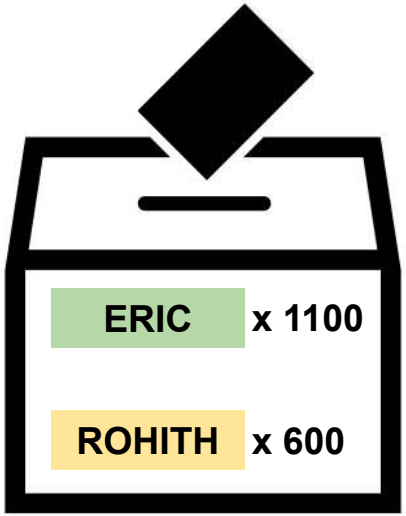
Ballots audited	0
Test statistic (T)	1
Risk (1/T)	1
% of votes for Eric	64.7%
Risk Limit	5%



A demonstration of BRAVO

- We keep track of our current audit data in a table

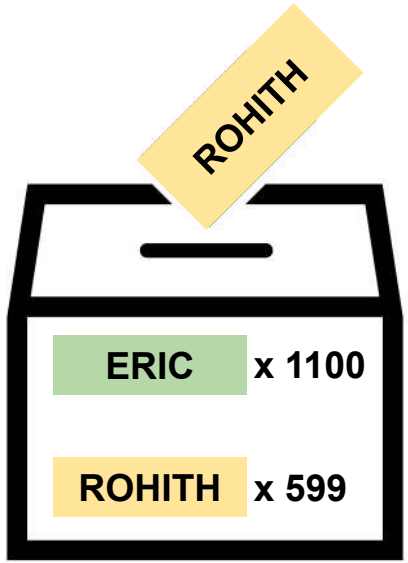
Ballots audited	0
Test statistic (T)	1
Risk (1/T)	1
% of votes for Eric	64.7%
Risk Limit	5%



A demonstration of BRAVO

- We take a random ballot from the box

Ballots audited	0
Test statistic (T)	1
Risk (1/T)	1
% of votes for Eric	64.7%
Risk Limit	0.05



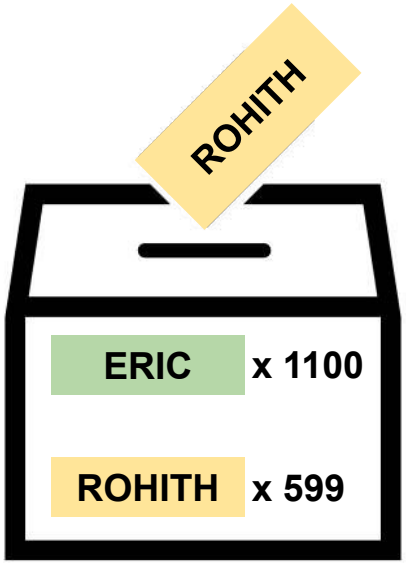
A demonstration of BRAVO

- It is for the projected loser, so multiply T by 2(1-s)

Ballots audited	1
Test statistic (T)	0.705
Risk (1/T)	1.417
% of votes for Eric	64.7%
Risk Limit	0.05

T =2(1-0.647) = 0.705

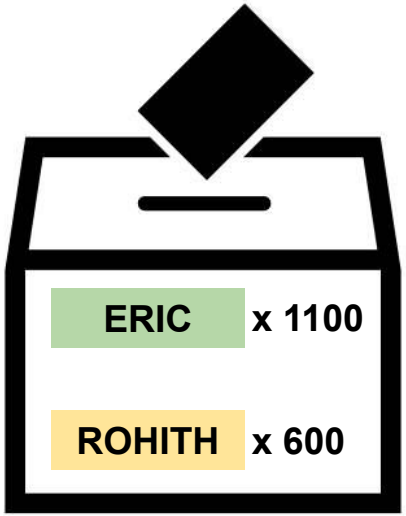
Risk = 1 / T = 1.417



A demonstration of BRAVO

- We replace the ballot

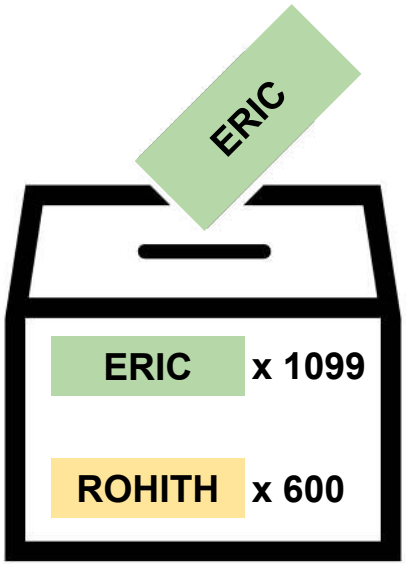
Ballots audited	1
Test statistic (T)	0.705
Risk (1/T)	1.417
% of votes for Eric	64.7%
Risk Limit	0.05



A demonstration of BRAVO

- We take another ballot from the box

Ballots audited	1
Test statistic (T)	0.705
Risk (1/T)	1.417
% of votes for Eric	64.7%
Risk Limit	0.05



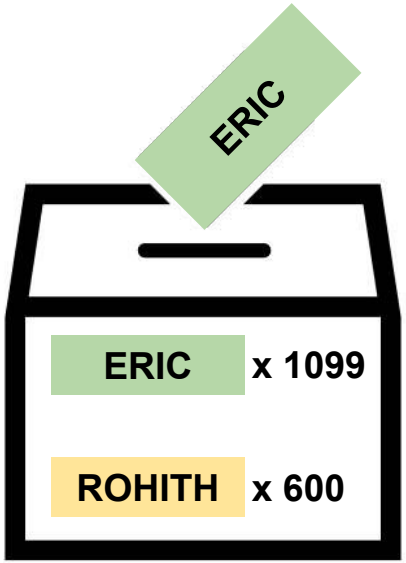
A demonstration of BRAVO

- It is for the projected winner, so multiply T by s / 50%

Ballots audited	2
Test statistic (T)	0.913
Risk (1/T)	1.109
% of votes for Eric	64.7%
Risk Limit	0.05

$T = 0.705 \times 2(0.647) = 0.913$

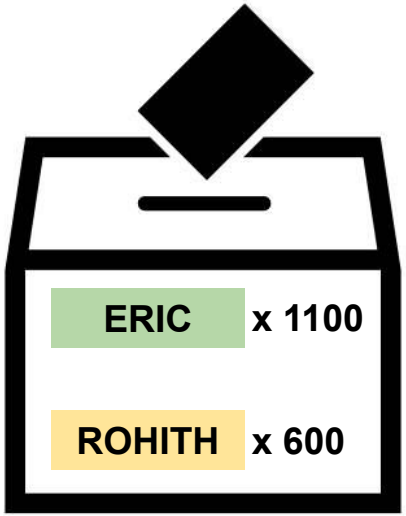
$P = 1 / T = 1.109$



A demonstration of BRAVO

- We replace the ballot

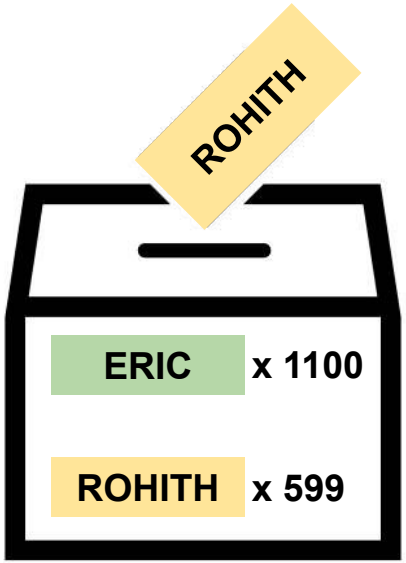
Ballots audited	2
Test statistic (T)	0.913
Risk (1/T)	1.109
% of votes for Eric	64.7%
Risk Limit	0.05



A demonstration of BRAVO

- We take another ballot from the box, and recalculate risk

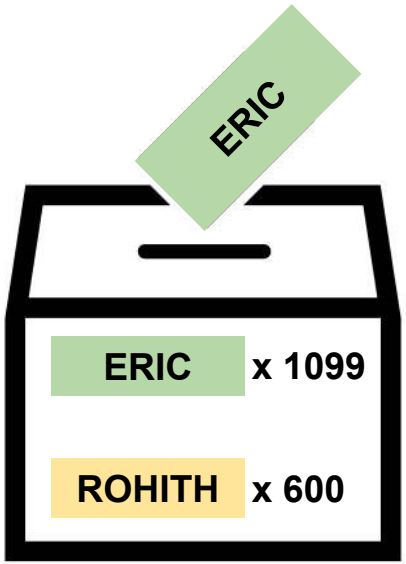
Ballots audited	3
Test statistic (T)	0.645
Risk (1/T)	1.551
% of votes for Eric	64.7%
Risk Limit	0.05



A demonstration of BRAVO

- We continue taking ballots from the box, and recalculate risk

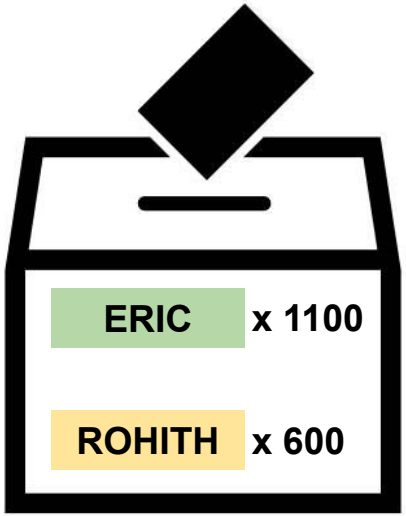
Ballots audited	31
Test statistic (T)	23.188
Risk (1/T)	0.043
% of votes for Eric	64.7%
Risk Limit	0.05



A demonstration of BRAVO

- As the P-value is now below the risk limit, we can stop

Ballots audited	31
Test statistic (T)	23.188
Risk (1/T)	0.043
% of votes for Eric	64.7%
Risk Limit	0.05



A demonstration of BRAVO - Results

- We needed 31 ballots to audit the election
  - Compared to 1700 ballots a full recount would need
- On average, the audit required about 70.9 ballots
- Did not depend on the number of ballots, only proportion
- Audit risk can be calculated easily

Drawbacks of BRAVO

- Reported and actual results must be close
  - Otherwise, the audit is unlikely to conclude
- Ballots must be chosen with replacement
  - Potential for abuse
- Ballots must be chosen randomly from the whole sample
  - Difficult to audit across different locations
- No support for stratifying samples
  - Risks cannot be combined without introducing uncertainty

ALPHA

ALPHA

- More advanced version of BRAVO
- Uses betting martingales
- Dynamically updates η, the alternative hypothesis
  - “Guess” of % of votes that reported winner received
- Allows for sampling without replacement
- More efficient than BRAVO when reported and actual results differ

Stratification in ALPHA

- BRAVO cannot to combine results from different strata
- ALPHA allows for simple multiplication to stratify
  - Multiply each test statistic to get the overall test statistic
  - Does not introduce uncertainty
  - Allows for realistic use of opportunistic auditing

Results

Set Up

- Modified ALPHA and created functionality that allowed opportunistic auditing
- Two states representing two stratum(Massachusetts and New York)
- State level races were governor’s races, global is presidential race

Simulations and Strategies

- Two major strategies tested:
  - Global first, then states
  - States first, then global
- Tested different sizes of margins in stratum and individual races
  - Big margin: 60%-40%, tight margin: 52%-48%
- Tested effect of drift: whether or not governor and presidential margins align
- Equal sized strata, 500 ballots each

Results - Without Drift

- Ballots required in only presidential audit **strongly mirrored/equal** to auditing states before or after opportunistically

Numerical Results - With Drift

Avg ballots used	Presidential Only	President then state	State then president
Big margins in presidential strata and MA gov, tight in NY gov	466.65	672.0	725.25
Big margins in presidential strata, tight in both gov races	513.8	838.15	717.8
Big margin in NY gov and MA pres, tight in MA gov and NY pres	691.55	758.7	771.3

Conclusions

- When there is no drift, almost no additional ballots required
  - Most common situation
- When presidential race has big margins, the strategy used for opportunistic auditing be optimized by using margins in states
- When both states experience opposite forms of drift, both strategies have similar efficacy

Future Work

- Expanding the number of levels in a simulation
  - Local races have low amount of scrutiny, important to audit
- Introducing error/ incorrect reported results
- Greater number of strata to more accurately reflect optimal strategies for real-world auditing
- More research on the costs and benefits of auditing more states or auditing states with closer margins
  - What is the main goal of opportunistic auditing?

Acknowledgements

- Thank you to Slava Gerovitch, Srini Devadas, and the rest of MIT PRIMES for making this project possible
- Thank you to Tanya Khovanova for 3 years of PRIMES STEP and inspiring us to pursue problem solving and research
- Thank you to our mentor, Mayuri Sridhar, for guiding our research, being our sounding board, and supporting us in every way possible
- Thank you to our parents for supporting us through the program