

Comparing Methods of Opportunistic Risk-Limiting Audits

Eric Chen, Rohith Raghavan
Mentored by Mayuri Sridhar

MIT PRIMES October Conference 2024
October 12th, 2024



Introduction

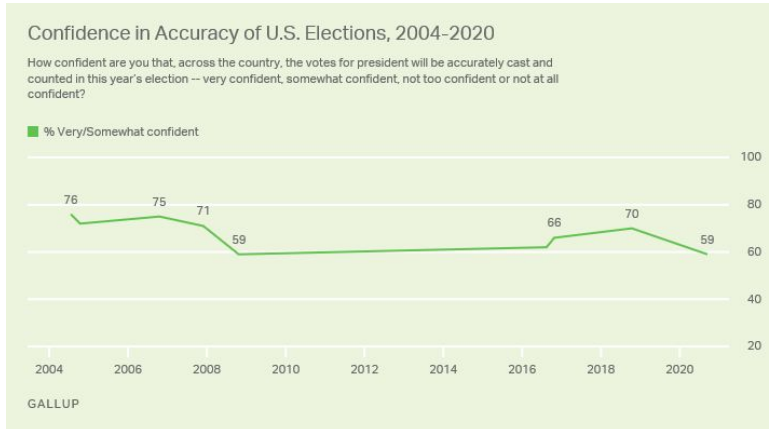
What Are Audits?

- Reviews of election results from a third party
- Detect both sabotage and mistakes



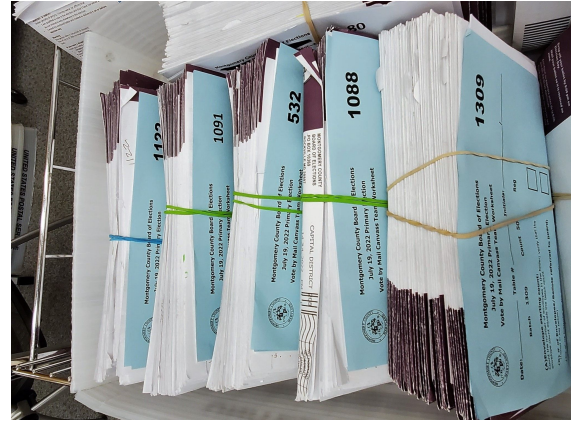
Why Run Audits?

- 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
- Stop once the risk is “low enough” - 5%
- Compare these ballots to the results to determine accuracy of the election

Risk Limiting Audits



What is a risk-limiting audit?

- Hypothesis test based on a sample drawn
- Audits affirm results when risk limit is met
 - Cannot reject results, only call for a full hand count (escalation)



How do risk-limiting audits work?

- We maintain a test statistic
- As we draw ballots, we update this test statistic until we are confident in our results or require escalation.



Benefits of RLAs

- Relies on proportions, not counts, of ballots.
- Avoids escalation unless absolutely necessary.
- **Stratification and Opportunistic Auditing**

Opportunistic Auditing



Strata

- Strata are smaller groups which divide the population
- **Global** races encompass multiple strata, while **local** races are limited to one stratum.
- We examine the US Presidential Election and statewide Governors races.



Opportunistic Auditing

- Elections normally consist of multiple contests at the same time
- **Opportunistic audits** gathers audit information on multiple races at once



Goal of Our Research

- We aim to create realistic methods of opportunistic auditing
- What strategies can we use to select strata to sample in opportunistic audits?
- What should we value more in opportunistic audits?



ALPHA



ALPHA

- More advanced RLA
- Uses betting martingales
 - Allows for dynamic updates based on current sampling results



ALPHA for One-Stratum Audits

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



Stratification in ALPHA

- ALPHA allows for simple multiplication to stratify
 - Multiply each test statistic to get the overall test statistic
 - Allows for realistic use of opportunistic auditing

Results - Two Strata



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

- Simple strategies of round robin
 - 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	672.04	725.25
Big margins in presidential strata, tight in both gov races	513.80	838.15	717.83
Big margin in NY gov and MA pres, tight in MA gov and NY pres	691.55	758.72	771.36

Results - Three Strata



Experimental Setup

- Tested on 3 strata, 40 ballots each
 - Margins of victory varied from 10-20% based on recent NY, NJ, MA results
- Randomly shuffled the ballots in each stratum
- Used ALPHA to audit and each stratum



Strategies

- Round Robin, president only
- Round Robin, president and state completion
- Lowest state governor's T-value
- Average governor's T-value optimized



Results - Data Table

Strategy:	Round Robin(Pres Only)	Round Robin(State Completed)	Lowest Governor T-value	Average-Optimization
Ballots Sampled:	472.08	905.07	1134.67	843.91
State Audit %:	41.3%	100%	56.3%	100%



Results

- Lowest state governor's T value clearly not feasible
 - High workload, unimpressive state results
- When auditing to state completion, average-optimization does slightly better than round robin
- High cost of auditing state audits



Conclusions

- Appears to be some merit to using average-optimization or similar strategies
 - Advanced strategies can be effective
- When high drift exists, significant cost to also auditing states
- Some strategies are not effective



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, Srinivasa 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



Main References

- Mark Lindeman and Philip B. Stark. A gentle introduction to risk-limiting audits. *IEEE Security & Privacy*, 10(5):42–49, 2012.
- Mark Lindeman, Philip B. Stark, and Vincent S. Yates. BRAVO: Ballot-polling risk-limiting audits to verify outcomes. In *2012 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE 12)*, Bellevue, WA, August 2012. USENIX Association.
- Philip B. Stark. Alpha: Audit that learns from previously hand-audited ballots, 2022.