The Athena Class of Ballot Polling Tabulation Audits

Filip Zagórski, Grant McClearn, Sarah Morin, Neal McBurnett, Poorvi L. Vora

https://arxiv.org/abs/2008.02315

Risk Limiting Tabulation Audit

- Designed to check tabulation
- Evidence-based elections require more:
  - Eligibility checks
  - Voter-verified paper trail
  - Compliance audits to ensure that the paper trail accurately represents all the properly cast ballots
- Set desired risk limit for the RLTA up front

Ballot Polling RLTA

Draw a sample of ballots, identified at random from the ballot manifest

If the sample passes the audit

- Stop and declare the RLTA successful

Else

- Draw more ballots or go to a full hand count

Quick Note: Ballot Comparison RLTA

- Far more efficient than any ballot polling RLTA
  - In Colorado, the sample size for a ballot-level comparison audit with a 5.37% margin (2016 Presidential) would be just 90, with a 100% stopping probability if there were (as usual) no discrepancies, vs 3,554 for BRAVO with 90% stopping probability
- Can be used only if you can identify individual ballots to compare with the machine's interpretation of the vote
- When individual ballots cannot be identified
  - Ballot polling RLTAs are a choice
  - Other choices include batch comparison RLTAs

BRAVO: most popular ballot polling RLTA

- Designed for use after each (single) ballot is drawn
- Requires fewest ballots, on average, in this setting
- When ballots are drawn in large numbers at a time:
  - End-of-round BRAVO: decision based on the tally of the sample
  - Selection-ordered BRAVO: decision based on order of the ballots drawn

The ATHENA Class of Audits: MINERVA

- Designed for use when ballots are drawn in large ***rounds.***
- Requires fewer ballots than BRAVO*.*
- We considered first rounds with a 90% probability of stopping if the election tally is as announced, thus limiting the need for going to a next round.
- Example results on statewide Presidential contest, 2016

MINERVA and BRAVO: Distinct Ballots drawn

stopping probability = 90%; risk limit =10% margin > 5%

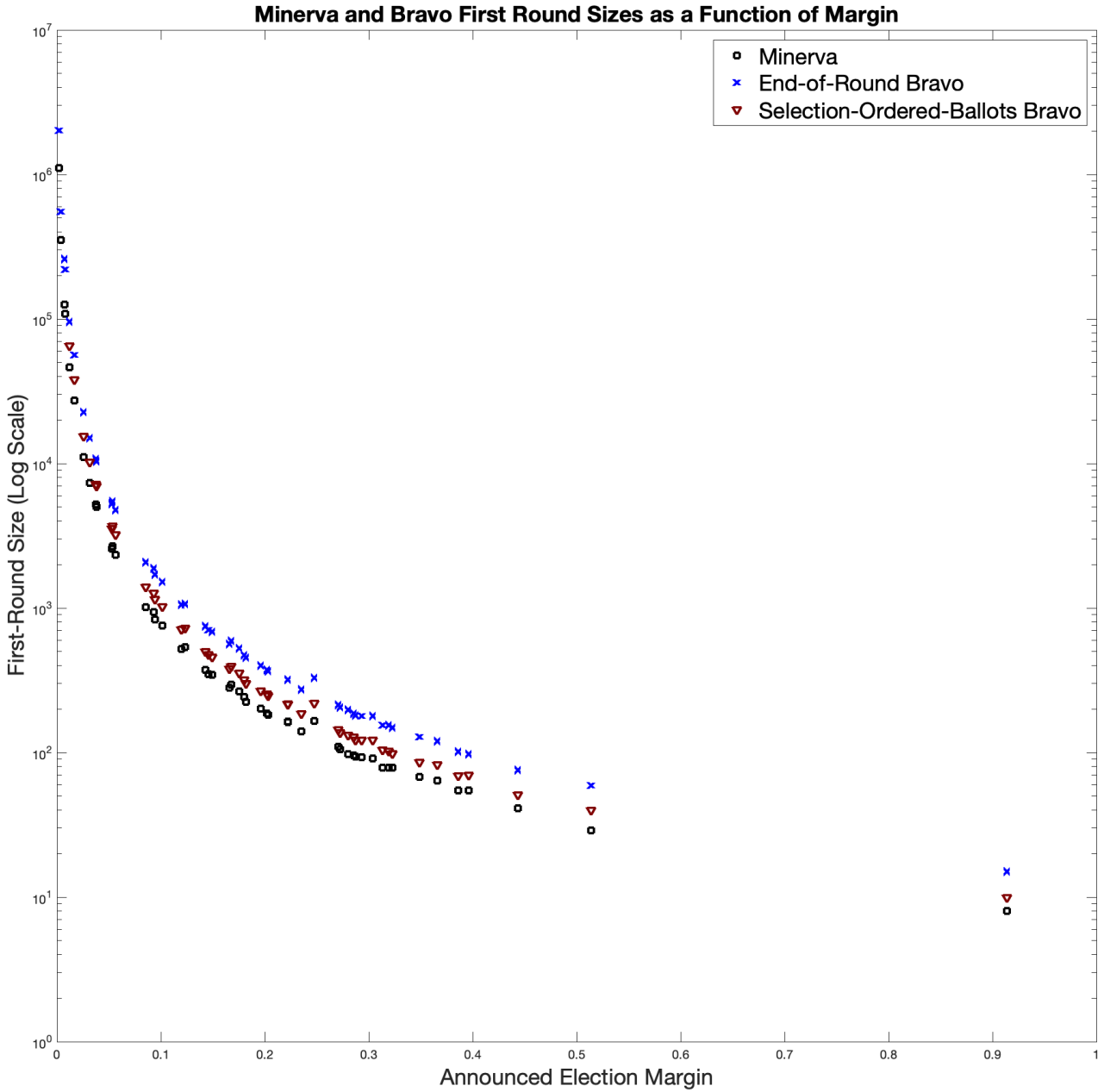
State	Margin in percentage	MINERVA	BRAVO		MINERVA round size as a percentage of BRAVO round size	
			EoR	SB	EoR	SB
California	32.26	79	148	99	53.38	79.80
Alabama	28.75	94	181	122	51.93	77.05
New York	23.54	140	272	186	51.47	75.27
Washington	17.57	265	525	355	50.48	74.65
Rhode Island	16.62	280	562	280	49.82	73.30
New Jersey	14.57	350	703	478	49.79	73.22
Oregon	12.31	535	1,068	724	50.09	73.90
Ohio	8.54	1,018	2,077	1,403	49.01	72.56
Virginia	5.65	2,329	4,788	3,228	48.64	72.13
Colorado	5.37	2,675	5,470	3,685	48.90	72.59
Georgia	5.32	2,567	5,263	3,554	48.77	72.23

MINERVA and BRAVO: Distinct Ballots drawn

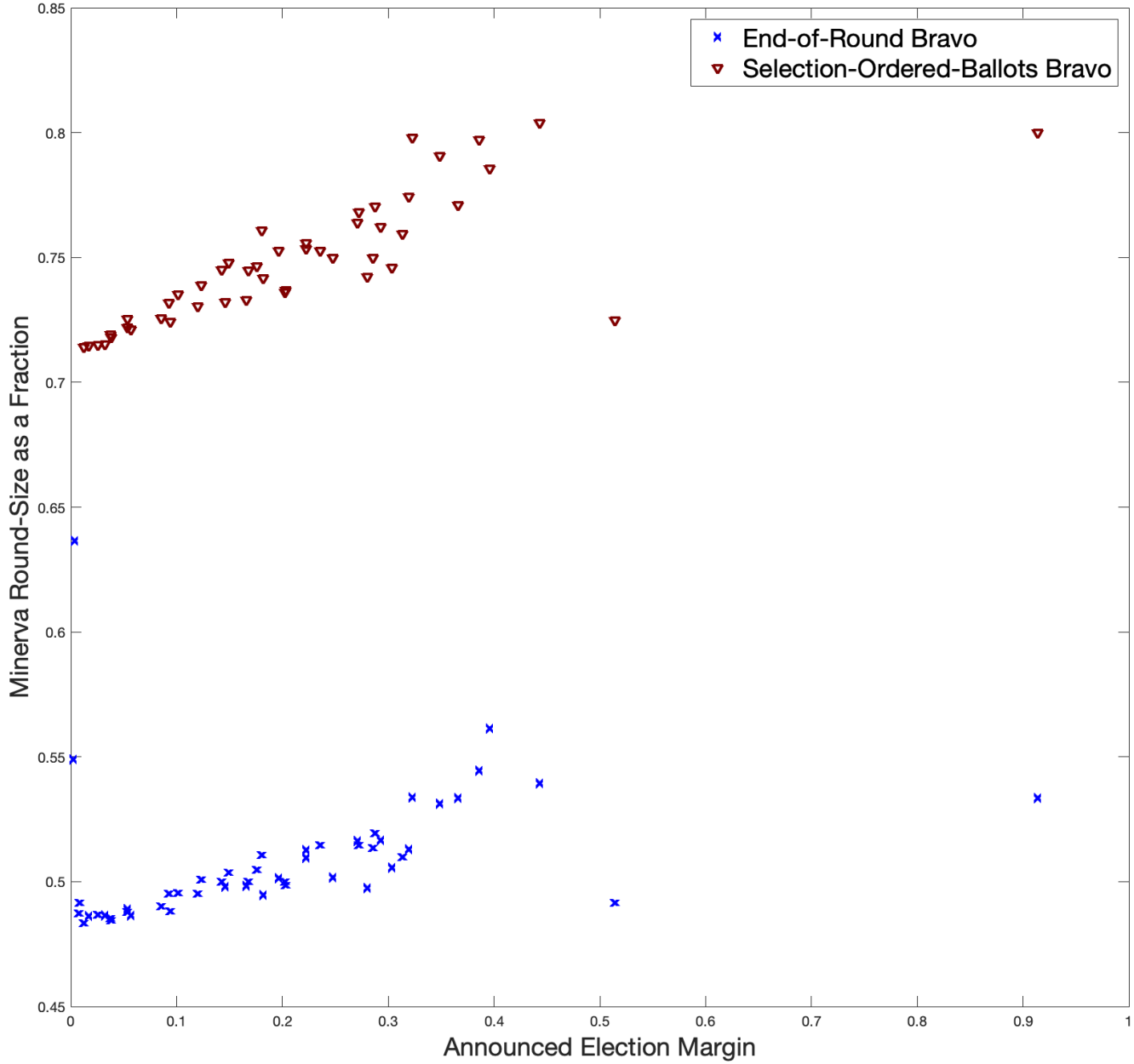
stopping probability = 90%; risk limit =10% margin < 5%

State	Margin in percentage	MINERVA	BRAVO		MINERVA round size as a percentage of BRAVO round size	
			EoR	SB	EoR	SB
North Carolina	3.81	4,998	10,319	6,956	48.43	71.85
Arizona	3.78	5,199	10,710	7,217	48.54	72.04
Nevada	2.59	11,056	22,711	15,357	48.68	71.99
Minnesota	1.66	27,294	56,139	37,939	48.62	71.94
Florida	1.24	46,449	96,115	64,827	48.33	71.65
Wisconsin	0.82	108,592	220,878		49.16	
Pennsylvania	0.75	126,477	259,621		48.72	
Michigan	0.24	1,107,933	2,018,381		54.89	

MINERVA and BRAVO First-Round Sizes (log plot)



MINERVA First-Round Sizes as a Fraction of BRAVO First-Round Sizes



Caveats

- The improvement over BRAVO is expected to:
  - Be smaller for smaller stopping probabilities
  - Vary with risk limit
- These are only first-round results.
  - Future rounds are trickier. E.g.: might need to commit to future round sizes in advance
- When available, ballot comparison uses smaller samples

Where are we?

- ✓Open source code: https://github.com/gwexploratoryaudits/
- ✓Proofs appear reasonable to those who have reviewed them.
- ✓First round simulations.
- Working on simulations to
  - Substantiate proven claims for further rounds
  - Examine audit behavior when election outcome is correct but announced tally is not
- Paper in review.