Leveraging Paper Ballots

Philip B. Stark

Department of Statistics
University of California, Berkeley

Was Archimedes a NY Politician?

> ### [Archimedes] used to say, in the Doric speech of *Syracuse*:
>
> Give me a place to stand and with a lever I will move the whole world.

`http://www.theonion.com/video/florida-to-experiment-with-n`

Pros & Cons of Lever Voting Machines

- +Familiar steam-punk aesthetic
- +Can tabulate votes after nuclear holocaust
- +Not subject to viruses
- +No auditing, no recounts!
- −No auditing, no recounts!
- −No way to assess accuracy or correct errors.
- −High residual vote: voter errors
- −Can be misprogrammed/hacked

Lever machines are great if knowing who really won doesn't matter much.

Vote-Counting Accuracy

- All ways of counting votes make some errors
- Can err capturing voter intent or tabulating
- Lever machines no exception
- If error rate high enough, can alter outcomes
- Need "breadcrumbs" (audit trail) to recover correct outcome
- Lever machines leave no breadcrumbs
- Voter-marked paper is much better: can measure error rate & recover from problems
- Breadcrumbs not enough: have to look!
- NY needs better audits, including voter-intent rules and sound statistics

Automatic Recount Thresholds

- For NYC citywide primaries, no recount if winner gets $> 40.5\%$ or if all margins $> 0.5\%$ and > 10 votes
- No scientific/statistical reason for 40% (or anything similar)
- *Some* reason for 0.5%—but not as good/efficient as good audit
- Intrinsic error rate for voter-marked paper $\approx 0.05\%$–0.5%. Depends also on PCOS v CCOS, ballot design
- Misprogramming & procedure failures can give *much* higher error rates
- Risk-limiting audits deal with all of these: guarantee large chance of correcting wrong outcomes at much lower cost than a recount that wouldn't change the outcome if the voting system supports it

What do we want election audits to do?

- Ensure that the electoral outcome is correct.
- If the outcome is wrong, correct it before it's final/official.

Good audits give strong evidence even w/ small margins

- Full hand count generally unnecessarily expensive and time-consuming.
- Instead, check a random sample by hand.
  Smaller margins require checking bigger samples.
  Even for very small margins, less work than a full hand count.
- Keep checking until there's convincing evidence that the outcome is right—or until all ballots have been examined and the right outcome is known.

Controlling the chance of error

- Sample is drawn at random, so there's a chance a wrong outcome will escape correction—but we can make that chance as small as we want. Statistics says how.
- *Risk* is the largest possible chance that the audit does not correct the outcome, if the outcome is wrong.
- *Risk-limiting audit* ensures that the largest possible chance is still a small chance, like 10%, 5%, 1%.
- Generally, have to check more ballots to make chance smaller.

Random Sampling

"Stirring" is key to reducing work

- Don't have to climb into the bathtub to tell if it's hot: can just stick your toe in—if the water is stirred well.
- Don't have to walk all over town to tell if it's cold outside: the air is mixed well enough that you just have to step outside to get a pretty good idea.
- Don't have to drink a whole pot of soup to tell if it's too salty: a teaspoon is enough—if the pot has been stirred. (Doesn't matter whether the pot holds 1q or 50g.)

How do you stir ballots?

Random sampling is stirring

- Imagine numbering the ballots.
- Write the numbers on ping-pong balls; put in a lotto machine.
- Lotto machine stirs the balls and spits some out.
- The ballots with the numbers on the selected balls are a random sample of ballots.
- Easier to stir balls than ballots. Even easier to generate random numbers.
- Still amounts to putting ballots into a huge cement mixer to stir them, then taking a "teaspoon" of ballots.

Paper rules—if it is right

- Can't correct wrong outcomes without counting the whole audit trail.
- Counting the whole audit trail won't give right answer unless it's adequately accurate and intact.
- Current procedures for protecting, tracking, and accounting for ballots are spotty. Should be top priority!

Risk limit *assumes* outcome is wrong in the hardest-to-find way. Biggest chance the outcome won't be corrected.

Ballot-polling Audits and Comparison Audits

- Ballot polling audit: sample ballots until there is strong evidence that looking at all of them would show the same election outcome.
  Like an exit poll—but of ballots, not voters.
- Comparison audit:
  1. Commit to vote subtotals (or CVRs), e.g., precinct-level results
  2. Check that the subtotals add up exactly to contest results
  3. Check subtotals by hand until there is strong evidence the outcome is right

Tradeoffs

- Ballot polling audit
  - Virtually no set-up costs
  - Requires nothing of voting system
  - Preserves voter anonymity except possibly for sampled ballots
  - Requires more counting than ballot-level comparison audit
  - Does not check tabulation: outcome could be right because errors cancel

- Comparison audit
  - Heavy demands on voting system for reporting and data export
  - Requires LEO to commit to subtotals
  - Requires ability to retrieve ballots that correspond to CVRs or subtotals
  - Checks tabulation (but not for *transitive audits* unless subtotals are cross checked as well)
  - Ballot-level comparison audits require least hand counting

Workload: Ballot-level audit, 2 Candidates
10% Risk Limit

| Margin | Ballots drawn | | | Comparison |
| | Ballot-polling | | | |
	median	90th percentile	Mean	(w/o errors)
40%	22	60	30	12
30%	38	108	53	16
20%	84	244	119	24
10%	332	974	469	48
8%	518	1,520	730	60
4%	2,051	6,053	2,900	120
2%	8,157	24,149	11,556	240
1%	32,547	96,411	46,126	480
0.5%	full hand count probably easier			960
0.1%	fuggedaboutit			4,800

Evidence-based elections

Principle: Trust, but verify

LEOs should give convincing evidence that outcomes are right (or say they can't).
"Trust me" is not convincing.

- Voters create complete, durable, accurate audit trail.
- LEO curates the audit trail adequately.
- Compliance audit to check whether the audit trail is trustworthy enough to determine who won.
  If not, how strong can the evidence be?
- Risk-limiting audit to correct the outcome if it is wrong. Presumes audit trail is OK.
  "Explaining" or "resolving" errors isn't enough.

What can NY do right now to improve EI?

- Don't resurrect lever machines: *leverage the paper trail!*
- Mandate rigorous ballot accounting
- Mandate ballot manifests
- Mandate compliance audits: assess integrity of audit trail
- Ballot-polling RLAs for large contests
- Develop software to support ballot-level comparison RLAs w/ current voting systems ASAP (partial re-scan?)
- Improve audit law: RLA with voter-intent provisions
- Eliminate automatic recount once RLAs are routine
- Plan replacement voting systems that have built-in auditibility