

Implementing Risk-Limiting Post-Election Audits in California

J.L. Hall^{1,2} L.W. Miratrix³ P.B. Stark³ M. Briones⁴ E. Ginnold⁴
F. Oakley⁵ M. Peaden⁶ G. Pellerin⁶ T. Stanionis⁵ T. Webber⁶

¹University of California, Berkeley; School of Information

²Princeton University; Center for Information Technology Policy

³University of California, Berkeley; Department of Statistics

⁴Marin County, California; Registrar of Voters

⁵Yolo County, California; County Clerk/Recorder

⁶Santa Cruz County, California; County Clerk

The Electronic Voting Technology Workshop/
Workshop on Trustworthy Elections 2009



Outline

Risk-Limiting Audits Defined

- What They Are

- What They Are Not

Risk-Limiting Audits in CA

- Marin County, Measure A (Feb. 2008)

- Yolo County, Measure W (Nov. 2008)

- Marin County, Measure B (Nov. 2008)

- Santa Cruz County, County Supervisor (Nov. 2008)

Discussion

- Inadequacy of Election Management Systems (EMS)

- Importance of Auditor/Election Official Communication

Simpler Risk-Limiting Audits?

- Can We Make Risk-Limiting Audits More Simple?

- Our Simpler Risk-Limiting Audit



Post-Election Manual Tally (PEMT) Audits Defined

Post-election audits require:

1. something to check. (*i.e.*, electronic results)
2. something to check against. (*i.e.*, physical audit trail)
3. an method for checking the two. (*i.e.*, hand counts)



Consensus Definition

“Risk-limiting audits have a large, pre-determined minimum chance of leading to a full recount whenever a full recount would show a different outcome.”¹

¹<http://electionaudits.org/principles.html>



Risk-Limiting Audits Defined

To limit risk, an audit must have:²

4. A minimum, pre-specified chance that, if the apparent outcome is wrong, every ballot will be tallied by hand.

Practically, risk-limiting audits have two more aspects:

5. A way to assess the evidence that the apparent outcome is correct, given the errors found by the hand tally.
6. Rules for enlarging the sample if the evidence that the apparent outcome is correct is not sufficiently strong.

²Any of this can be applied to open-audit voting systems.



Current Audits and Audit Policy Do Not Limit Risk

Some problems:

- ▶ Focus typically on initial sample size
 - ▶ Not as important as measuring error and escalation
- ▶ Error should be contextualized at the contest level
 - ▶ Often, escalation applies to machines or geographical regions
- ▶ Often use *ad hoc* error bounds
 - ▶ For example, Within-Precinct Miscount (WPM) is bogus
- ▶ Must get *both* the legal and statistical wording correct
 - ▶ Often mix *detection* and *confirmation* paradigms



But Some States Are Getting Closer...

- ▶ AK, HI, OR, TN, WV use fairly blunt methods to get closer
- ▶ CA, MN and NY have somewhat better schemes...
- ▶ CO is relatively the best:

“risk-limiting audit” means an audit protocol that makes use of statistical methods and is designed to limit to acceptable levels the risk of certifying a preliminary election outcome that constitutes an incorrect outcome.

- ▶ However, what are “statistical methods”?
- ▶ Also, “incorrect outcome” specifies “recount” instead of “full hand (re)count”



Overview

County	Total Ballots	Winner	Loser	Margin	# Ballots Audited	% Ballots Audited
Marin (A)	6,157	4,216	1,661	5.1%	4,336	74%
Yolo	36,418	25,297	8,118	51.4%	2,585	7%
Marin (B)	121,295	61,839	42,047	19.1%	3,347	3%
Santa Cruz	26,655	12,103	9,946	9.6%	7,105	27%



Marin A: The Election, Test and Sample

- ▶ The Election: Kentfield School District Measure A
 - ▶ 9 precincts³, 5,877 ballots cast, 298-vote margin (5.1%)
- ▶ The Test and Sample:
 - ▶ Error measured as overstatement of margin, x .
 - ▶ Weight function, w_p :

$$w_p(x) = \frac{(x - 4)_+}{b_p}$$

- ▶ Stratified random sample of 6 precincts in 2 strata (IP/VBM)

³One had only 6 registered voters, we treated it entirely as error.



Marin A: Risk Calculation and Cost

▶ Risk Calculation:

- ▶ If 1 batch overstated the margin, a random sample of 6/8 batches would have missed it with probability:⁴

$$\frac{\binom{7}{6}}{\binom{8}{6}} = 25\%.$$

▶ Cost:

- ▶ Took $1\frac{3}{4}$ days, total cost: \$1,501, \$0.35 per ballot

⁴ $\binom{x}{y}$ is shorthand for the binomial coefficient $x!/(y!(x-y)!)$.



Yolo: The Election, Test and Sample

- ▶ The Election: Davis Joint Unified School District
 - ▶ 57 precincts, 36,418 ballots, 17,179-vote margin (51.4%)
- ▶ The Test and Sample:
 - ▶ Stratified Random Sample (IP/VBM) with small precincts in one stratum treated entirely as error
 - ▶ Used maximum relative overstatement (MRO) of margins instead of weighted margin overstatement
 - ▶ MRO normalizes the overstatement by the reported margin. . . an overstatement in a contest with a small margin is weighted more



Yolo: Risk Calculation and Cost

- ▶ Risk Calculation:
 - ▶ To limit risk to 25% required sample of 6/103 batches
 - ▶ Found two errors (only one overstatement error), below the threshold to trigger expansion
- ▶ Cost: Not directly relevant
 - ▶ Two authors and one official did the counting!



Marin B: The Election, Test and Sample

- ▶ The Election: Measure B (added two govt. admin. positions)
 - ▶ 189 precincts, 121,295 ballots, 19,792-vote margin (19.1%)
- ▶ The Test and Sample:
 - ▶ Used trinomial bound based on taint, t_p , of each batch
 - ▶ $t_p \equiv e_p/u_p \leq 1$ (e_p is MRO in p)
 - ▶ Compares t_p to a pre-specified threshold, d
 - ▶ Batches have either non-positive t_p ; t_p less than d ; or, t_p greater than d
 - ▶ Bounds risk based on category counts in each bin
 - ▶ Trinomial bound uses weighted sampling with replacement *probability proportional to an error bound* (PPEB)
 - ▶ With stratified random sampling, we would have had to count 44% more ballots



Marin B: Risk Calculation and Cost

▶ Risk Calculation:

- ▶ Chose $d = 0.038$ and $n = 14$ (number of draws) based on previously observed levels of error (see [1])
- ▶ Because sampling is with replacement, we get an expected number of unique precincts:

$$\sum_p \left(1 - \left(1 - \frac{u_p}{U} \right)^n \right) = 13.8$$

- ▶ Audit found no errors⁵
- ▶ Cost: 2 days, \$1,723 or \$0.51 per ballot

⁵However, we apparently audited results that were too preliminary



Santa Cruz County, County Supervisor (Nov. 2008)

Santa Cruz: The Election, Test and Sample

- ▶ The Election: Santa Cruz County Supervisor, 1st District
 - ▶ 76 precincts, 26,655 ballots, 2,139-vote margin (8.0%)
- ▶ The Test and Sample:
 - ▶ PPEB sampling using the trinomial bound



Santa Cruz: Risk Calculation and Cost

- ▶ Risk Calculation:
 - ▶ set $n = 19$ and $d = 0.047$
 - ▶ We did see some error:
 - ▶ largest t_p was 0.036, 1 ballot overstatement in small precinct
 - ▶ largest overstatement was 4 ballots in a large precinct, t_p here was 0.007
 - ▶ No t_p was larger than d , so we could certify at 25% risk
- ▶ Cost: 3 days, cost \$3,248, or \$0.46 per ballot



Inadequacy of Election Management Systems (EMS)

Ugh, EMSs

- ▶ A constant factor was the inadequacy of EMS output

411	Precinct Reporting	2	0	834			3130	3134	3140	3145	3146	3151	315
412	NP - TONY MADRIGAL	2	1	834			442	346	336	568	377	196	37
413	NP - LISA J. MOLYNEUX	2	2	834			184	135	147	225	104	104	15
414	NP - DON LANE	2	3	834			618	395	621	770	465	272	45
415	NP - TIM FITZMAURICE	2	4	834			418	229	317	498	326	166	28
416	NP - J. CRAIG CANADA	2	5	834			109	68	62	107	44	45	6
417	NP - BLAS JACOB (JAY) CA	2	6	834			72	79	50	107	67	34	7
418	NP - RYAN COONERTY	2	7	834			827	482	846	945	565	345	53
419	NP - SIMBA KENYATTA	2	8	834			174	107	148	263	183	97	15
420	NP - KATHERINE BEIERS	2	9	834			576	320	445	654	410	217	41
421	NP - DAVID TERRAZAS	2	10	834			616	381	638	612	349	215	34
422	WRITE-IN	2	11	834			0	0	0	0	0	0	
423	Santa Cruz City Council Vol	3	-1	834									
424	Precinct Reporting	3	0	834	-1	87938	3171	3178	5101	5108	5109	5162	
425	NP - TONY MADRIGAL	3	1	834	11365	87938	304	561	303	349	583	257	
426	NP - LISA J. MOLYNEUX	3	2	834	4224	87938	159	171	121	128	272	55	
427	NP - DON LANE	3	3	834	13944	87938	411	515	385	365	704	244	
428	NP - TIM FITZMAURICE	3	4	834	9171	87938	299	382	267	272	491	194	
429	NP - J. CRAIG CANADA	3	5	834	1945	87938	79	85	61	54	118	18	
430	NP - BLAS JACOB (JAY) CA	3	6	834	2166	87938	74	104	67	70	115	33	
431	NP - RYAN COONERTY	3	7	834	17056	87938	491	628	402	458	895	262	
432	NP - SIMBA KENYATTA	3	8	834	5105	87938	155	280	180	195	289	99	
433	NP - KATHERINE BEIERS	3	9	834	11642	87938	433	429	315	305	626	195	
434	NP - DAVID TERRAZAS	3	10	834	11320	87938	342	376	276	258	570	111	
435	WRITE-IN	3	11	834		87938	0	0	0	0	0	0	

Ugh, EMSs

- ▶ We ended up re-keying batch-level data because of this
 - ▶ No way we can do this for many or big elections
- ▶ Unclear what EMSs are actually capable of
 - ▶ HTML?, XML?, EML?, CSV?, PDF? (yuk!), DB dumps?
- ▶ We had to do some strange DB reporting calisthenics
 - ▶ *E.g.*, Marin EMS could not report results at batch-level
 - ▶ We modified DB reports to remove all but 1 batch, re-ran
- ▶ We'd like to see structured data (EML) with schema (XSD)



Communication is key!

- ▶ Santa Cruz
 - ▶ The totals we used for calculations did not include provisional ballots
 - ▶ However, the audit did include them!
 - ▶ We had to treat all changes in totals due to provisional ballot changes as error
- ▶ Marin Measure B
 - ▶ One week ago, noticed a similar problem in Marin Measure B
 - ▶ Precincts in Marin smaller than 250 registered voters are forced to be VBM
 - ▶ However, the EMS lists these as IP
 - ▶ Used premature results for one precinct marked as IP that was forced-VBM



Can We Make Risk-Limiting Audits More Simple?

Risk-Limiting Audits Don't Have to be as Complex

- ▶ Risk-limiting methods that use statistics based on observed audit discrepancy to decide to escalate are complex
- ▶ Even with an experienced statistician, the logistics are complex and can lead to high uncertainty for election officials



Our Proposal

1. *Basic Audit Level*: A fixed percentage of batches (*e.g.*, 0.5%) from every race is hand counted
2. *Full Recount Trigger*: Any contest with a sufficiently small margin is counted by hand in its entirety
3. *Random Full Hand Counts*:

$$P_r = \frac{f_r}{20} + \frac{1}{1000 \cdot m_r}$$

P_r is the probability of a full hand count, f_r is fraction of voters eligible to vote in the contest and m_r is the margin in the race expressed as a fraction



Conclusions

- ▶ Risk-limiting audits are within reach
- ▶ They're cheap (~ \$0.44 per ballot)
- ▶ They're difficult to administer

- ▶ Future?
 - ▶ Kaplan-Markoff [2] approach appears to be promising
 - ▶ Stratified sampling across Cong. districts is unsolved



Some Further Reading I



Luke W. Miratrix and Philip B. Stark
Election Audits using a Trinomial Bound.

University of California at Berkeley Department of Statistics,
<http://statistics.berkeley.edu/~stark/Preprints/trinomia109.pdf>



Philip B. Stark
Efficient Post-Election Audits of Multiple Contests: 2009 California Tests.

University of California at Berkeley Department of Statistics,
<http://ssrn.com/abstract=1443314>

