NY SD 7

Legislation

Risk-limiting audits

Ballot-level audits

# AUDITS: THE AFTER-MATH OF ELECTION REFORM Conference on Innovative Electoral Reforms and Strategies

Philip B. Stark

Department of Statistics University of California, Berkeley

10-11 December 2010

Legislation

Risk-limiting audit

#### Voters marks vary widely: No VTS can count perfectly

			Presidential General Election		Precinct 100063
and the second of the second of the	and the second se	A DESCRIPTION OF THE PARTY OF T	YOLO COUNTY November 04, 2008		Precinct 700005
	Both Sides		Demonstration &	Proposition 11 NEDISTRICTING INITIATIVE CONSTITUTIONAL AMENDMENT AND STATUTE	Davis Joint Unified School District Measure W Shall the Davis Joint Unified School District preserve existing classroom programs including math and science, English, music,
Presidential General Election	Both Sides		Changes California Constitution to eliminate	Changes authority for establishing state office boundaries from elected representatives to	physical education, librarians, secondary class
VOLO COUNTY November 04, 2008	Precinct 100383		the right of same-sex couples to many. Provides that only marriage between a man and a women is velid or recognized in California.	commission. Establishes multilevels poolies to select commissioners from registered voter pool. Commission comprised of Democratis, Republicans, and representatives of neither	Bits induction in terror and a second
Proposition 1 Proposition 1 Proposition 1 Proposition 1 OCMITH/TONAL AND/ONLY OCMITH/TONAL AND/ONLY Design Cathony Contactive Statistics Proposition 1 Design Cathony Contactive Statis	Davis Joint United School Descet Message W Dritt the Davis Joint Officed School Descet Immore assign descroom graphic water including match and searces. Scriptics, source refract description, thereins	*nnnn:********************************	revenue loss, mainly sales taxes, totaling in the several tens of millions of dollars, to state and local governments. In the long run, likely table facal impact on state and local povernments.	party. Placel impact: Potential increase in state redistricting costs ence every ten years due to the entities performing redistricting. Any increase in costs probably would not be	parcel for all other parcels?
and a eveny is unit or recogniser in an advecting of the	Non-reduction, subsidiar and co-cumulate programm including durate, defects, and providen ty being autolocide to key a sponsition for a provide of these preservent to calculate from annotation of 2520 Dig per feeding unit for multi-baseling presents and 3120-05 per primer for all does previous?	+	Yes No	Yes	
Monte in an angelie de la la en en la sur la	Ves No g		Proposition 9 CRIMINAL JUSTICE SYSTEM, VICTIMS' RIGHTS, PARIOLE INITIATIVE CONSTITUTIONAL AMENEMENT AND STATUTE.	Proposition 12 VETERANS' BOND ACT OF 2008. This act provides for a bond issue of nine hundred million dollars (\$900.002.000) to	
Precedence			Requires notification to victim and opportunity for input during phases of criminal justice process, including ball, pices, semimuling and parole. Establishes uclim safety in consideration for ball or parole.	provide farm and home aid for California veteram. Fiscal impact: Costs of about \$1.8 billion to pily off both the principal (\$900 million) and insecond r05% willion) on the bonds; costs pair	
In Paul aling present information (2002LII) process, horizoity of press, and and an exerce at the California status, Excluderes status aling press, and aling press, finishina press, finishina and california aline and press, finishina press, finishina press, finishina press, finishina press, finishi press, finishina press, finishina press, finishina press, and press, finishina press, finishina press, finishi press, finishina press, finishina press, and press, finishina press, finishina press, finishi press, finishina press, finishina press, finishi press, finishina press, finishi press, finishi press, finishi pr			Fiscal Impact: Potential loss of state saving on prison operations and increased county (a costs amounting to hundreds of millions of dollars annually. Potential net savings in the logs tens of millions of dollars annually on	by participating veterans. Average payment sor principal and interest of about \$50 million per year for 30 years. Bonds Yes	
Contraction of the strength in the strenght in the strength in the strength in the strength in the streng			Yes No	Bonds No Los Rios Community College Distric Measure M Tanat son Ros Community College Distric	
Proposition 10 ALTERNATIVE FUEL VEHICLES AND INCOLONAL COLORY BOLISS			Proposition 10 ALTERNATIVE FUEL VEHICLES AND RENEWABLE EMERGY BONDS. INITIATIVE STATUTE.	be authorized to issue \$475 million in bonds a the lowest available interest sates to improve student academic performance by building classrooms, facilities and labs throughout the result including for facehing prent	×
And work is the noncontraction of the strength			Authorizes 55 billion in bonds paid from th Qeneral Punit, to help consumers and offer purchase certain whiches, and so and research in renewable energy and alterna faal vehicles. Prevail trepart: State cett of about 513 hel	rs programs, architecture, engineering and construction management, computer actionor acity childhood development; and fire and police public calefy programs at the American Disco Committee River, El Dorado, Folson.	a. n
Mar of allow and official and an an and an		64	crear 30 years to replay bonds. Increased and local investigation of potentially foldering re- terns of millions of dollars through 2010. Protectal state administrative dollar up to \$10 million annually.	and Sacramento City College campuses?"	
Via	***	1063031	Bonds Yes	City of Davis Measure N Stast the Proposed Creater of the City of Davis De Adlights?	
	-				
Non Barris	25			19	
				1	oth Sides



Why Audit?	NY SD 7	Legislation	Risk-lim	iting audits	Ballot-level audits
	Ary SD 7 Special Election November 2009 City of Davis November 03, 2009 Instruction Text: Prese use a black or blas fits parts to mark your choice To vark for your choice in each context, completely fit you'ded to the list of your choice. Pressure 2009 P	es on the ballot. In the box eral Plan to nech property from Lipo and	Precir	iting audits Special Election November City of Davis November 03, 2009 Instruction Text: Please use a black or blue ink pen to n To vole for your choice in each contex provided to the left of your choice ange the land use designations for the apricultare to residential uses, as set for applications Ranch Project Fee Withonse Ranch Project be approved No	2009 ark your choices on the ballot, completely fill in the box the Davis General Plan to e Widhorse Ranch property from ch in the Resolution and tures for development of the
nni			1000		



NY SD 7

Legislation

Risk-limiting audits

Ballot-level audits

### Why audit?

#### • Any way of counting votes makes mistakes.

- If there are enough mistakes, apparent winner could be wrong.
- If there's a complete, accurate audit trail, can ensure big chance of correcting wrong outcomes—but you have to count by hand.
- Other reasons too: process improvement, deterrence, etc.
- Compliance audits vs. materiality audits.

▲□▶▲□▶▲□▶▲□▶ □ ● ● ●

Y SD 7

Legislation

Risk-limiting audits

Ballot-level audits

- Any way of counting votes makes mistakes.
- If there are enough mistakes, apparent winner could be wrong.
- If there's a complete, accurate audit trail, can ensure big chance of correcting wrong outcomes—but you have to count by hand.
- Other reasons too: process improvement, deterrence, etc.
- Compliance audits vs. materiality audits.



egislation

Risk-limiting audits

Ballot-level audits

- Any way of counting votes makes mistakes.
- If there are enough mistakes, apparent winner could be wrong.
- If there's a complete, accurate audit trail, can ensure big chance of correcting wrong outcomes—but you have to count by hand.
- Other reasons too: process improvement, deterrence, etc.
- Compliance audits vs. materiality audits.



Legislation

Risk-limiting audits

Ballot-level audits

- Any way of counting votes makes mistakes.
- If there are enough mistakes, apparent winner could be wrong.
- If there's a complete, accurate audit trail, can ensure big chance of correcting wrong outcomes—but you have to count by hand.
- Other reasons too: process improvement, deterrence, etc.
- Compliance audits vs. materiality audits.



Legislation

Risk-limiting audits

- Any way of counting votes makes mistakes.
- If there are enough mistakes, apparent winner could be wrong.
- If there's a complete, accurate audit trail, can ensure big chance of correcting wrong outcomes—but you have to count by hand.
- Other reasons too: process improvement, deterrence, etc.
- Compliance audits vs. materiality audits.

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ● □ ● ● ●

- Essential that voters create complete, durable, accurate audit trail.
- Essential that voting systems enable auditors to access reported results (total ballots, counts for each candidate, registered voters) in auditable batches. (Smaller batches are better.)
- Essential to select batches at random, *after* the results are posted. (Can supplement with "targeted" samples.)
- Need a plan for dealing with discrepancies, possibly leading to full count. "Reconciling" or "explaining" isn't enough.

NY SD 7

Legislation

**Risk-limiting audits** 

Ballot-level audits

- Essential that voters create complete, durable, accurate audit trail.
- Essential that voting systems enable auditors to access reported results (total ballots, counts for each candidate, registered voters) in auditable batches. (Smaller batches are better.)
- Essential to select batches at random, *after* the results are posted. (Can supplement with "targeted" samples.)
- Need a plan for dealing with discrepancies, possibly leading to full count. "Reconciling" or "explaining" isn't enough.

NY SD 7

Legislation

Risk-limiting audits

Ballot-level audits

- Essential that voters create complete, durable, accurate audit trail.
- Essential that voting systems enable auditors to access reported results (total ballots, counts for each candidate, registered voters) in auditable batches. (Smaller batches are better.)
- Essential to select batches at random, *after* the results are posted. (Can supplement with "targeted" samples.)
- Need a plan for dealing with discrepancies, possibly leading to full count. "Reconciling" or "explaining" isn't enough.

NY SD 7

Legislation

Risk-limiting audits

Ballot-level audits

- Essential that voters create complete, durable, accurate audit trail.
- Essential that voting systems enable auditors to access reported results (total ballots, counts for each candidate, registered voters) in auditable batches. (Smaller batches are better.)
- Essential to select batches at random, *after* the results are posted. (Can supplement with "targeted" samples.)
- Need a plan for dealing with discrepancies, possibly leading to full count. "Reconciling" or "explaining" isn't enough.

NY SD 7

Legislation

Risk-limiting audit

Ballot-level audits

### New York's Recent Reforms

Moved to precinct-count optically scanned paper ballots.

Introduced audit laws, starting with 3% of machines (scanners).

Irreconcilable differences between hand count and machine count can lead to counting more ballots by hand: 5%, 12%, or all.



- Balance of power in NY Senate: Either 31 seats for each party, or 32 for Republicans.
- Reported margin of 451 votes (0.5%) for Republican candidate Martins.

• Disagreement about purpose and requirements of auditing.



- Balance of power in NY Senate: Either 31 seats for each party, or 32 for Republicans.
- Reported margin of 451 votes (0.5%) for Republican candidate Martins.

• Disagreement about purpose and requirements of auditing.



- Balance of power in NY Senate: Either 31 seats for each party, or 32 for Republicans.
- Reported margin of 451 votes (0.5%) for Republican candidate Martins.

• Disagreement about purpose and requirements of auditing.

- 7 of 249 of machines audited at random (3%).
- 3 of 7 (i.e., 43%) showed errors. Net error favored the apparent winner.
- Republicans: the errors were "reconciled": Machines are fine. Democrats disagree.
- Judge Warshawsky: "In my opinion, reconcilable would be 'Is there a clear reason why the deficiency occurred?' "
- Is "the machine was mis-programmed" a clear reason?
- In my opinion, reconcilable would be "We counted again by hand and found that the error was in the hand count, not the machine count."
- "Clear reason" is irrelevant for whether the apparent outcome of the contest is correct. Size of the difference matters.

- 7 of 249 of machines audited at random (3%).
- 3 of 7 (i.e., 43%) showed errors. Net error favored the apparent winner.
- Republicans: the errors were "reconciled": Machines are fine. Democrats disagree.
- Judge Warshawsky: "In my opinion, reconcilable would be 'Is there a clear reason why the deficiency occurred?' "
- Is "the machine was mis-programmed" a clear reason?
- In my opinion, reconcilable would be "We counted again by hand and found that the error was in the hand count, not the machine count."
- "Clear reason" is irrelevant for whether the apparent outcome of the contest is correct. Size of the difference matters.

NY SD 7

Legislation

Risk-limiting audit

Ballot-level audits

- 7 of 249 of machines audited at random (3%).
- 3 of 7 (i.e., 43%) showed errors. Net error favored the apparent winner.
- Republicans: the errors were "reconciled": Machines are fine. Democrats disagree.
- Judge Warshawsky: "In my opinion, reconcilable would be 'Is there a clear reason why the deficiency occurred?' "
- Is "the machine was mis-programmed" a clear reason?
- In my opinion, reconcilable would be "We counted again by hand and found that the error was in the hand count, not the machine count."
- "Clear reason" is irrelevant for whether the apparent outcome of the contest is correct. Size of the difference matters.

NY SD 7

Legislation

Risk-limiting audit

Ballot-level audits

- 7 of 249 of machines audited at random (3%).
- 3 of 7 (i.e., 43%) showed errors. Net error favored the apparent winner.
- Republicans: the errors were "reconciled": Machines are fine. Democrats disagree.
- Judge Warshawsky: "In my opinion, reconcilable would be 'Is there a clear reason why the deficiency occurred?' "
- Is "the machine was mis-programmed" a clear reason?
- In my opinion, reconcilable would be "We counted again by hand and found that the error was in the hand count, not the machine count."
- "Clear reason" is irrelevant for whether the apparent outcome of the contest is correct. Size of the difference matters.

NY SD 7

Legislation

Risk-limiting audit

Ballot-level audits

- 7 of 249 of machines audited at random (3%).
- 3 of 7 (i.e., 43%) showed errors. Net error favored the apparent winner.
- Republicans: the errors were "reconciled": Machines are fine. Democrats disagree.
- Judge Warshawsky: "In my opinion, reconcilable would be 'Is there a clear reason why the deficiency occurred?' "
- Is "the machine was mis-programmed" a clear reason?
- In my opinion, reconcilable would be "We counted again by hand and found that the error was in the hand count, not the machine count."
- "Clear reason" is irrelevant for whether the apparent outcome of the contest is correct. Size of the difference matters.

NY SD 7

Legislation

Risk-limiting audit

Ballot-level audits

- 7 of 249 of machines audited at random (3%).
- 3 of 7 (i.e., 43%) showed errors. Net error favored the apparent winner.
- Republicans: the errors were "reconciled": Machines are fine. Democrats disagree.
- Judge Warshawsky: "In my opinion, reconcilable would be 'Is there a clear reason why the deficiency occurred?' "
- Is "the machine was mis-programmed" a clear reason?
- In my opinion, reconcilable would be "We counted again by hand and found that the error was in the hand count, not the machine count."
- "Clear reason" is irrelevant for whether the apparent outcome of the contest is correct. Size of the difference matters.

NY SD 7

Legislation

Risk-limiting audit

Ballot-level audits

- 7 of 249 of machines audited at random (3%).
- 3 of 7 (i.e., 43%) showed errors. Net error favored the apparent winner.
- Republicans: the errors were "reconciled": Machines are fine. Democrats disagree.
- Judge Warshawsky: "In my opinion, reconcilable would be 'Is there a clear reason why the deficiency occurred?' "
- Is "the machine was mis-programmed" a clear reason?
- In my opinion, reconcilable would be "We counted again by hand and found that the error was in the hand count, not the machine count."
- "Clear reason" is irrelevant for whether the apparent outcome of the contest is correct. Size of the difference matters.

ヘロト 人間 とくほとくほとう

æ

### Example NY SD7 Audit report

3% AUDIT: "SUPERVIS	OR TEAM" RESULTS
T/AD/ED _417023_ POLL SITE _Mineola Historical Society	MACHINE NUMBER AUDIT # DRAWN _5
REPUBLICAN SUPERVISOR PASSx_ FAIL Basic for Decision	DEMOCRATIC SUPERVISOR PASS FAILx Basis for Decision:
2 <sup>nd</sup> ALDDY TEAM RECONCLED NUMBER OF BALLOTS     3 different billion in question help to record the discreption.     Addef of the 7 <sup>th</sup> SD race influent one more vote for Marins on line 8, while methods once and a new vote. Al where the question by such that the 1 ballow new Over wee ballow that was subtributed to wash sub- down and the sub-ballow that was subtributed to ballow new A particular by such that discreptions.     Addef of the machine manually also would recorded the discreption.     SIGNATURE     Communities Mar 752 Nor	The DS200 counted and shukked 290 ballon and the same number of ballon was conside in the hand-count. For shorth read, bower it is hand-count results is an intervent it need, bound from the votes shukked by the machine. In the rate of counters, for example, the hand count resulted in one additional works of Luce B and only the result of an ore additional works of Luce B and one in we work on line B; the hand-count is old offered from the machine that necoded by the machine that necoded by the machine in definition of the ballon of the hand count overvoort dut had been recorded by the machine there is no way to 'recordie' these thalation discreptions. Becaused be audit revealed that the DS200 did NOT security oblastics the ballon, the machine SIGNATURE <u>SUMPACE</u> NO LAC
L	

### HO 13-24 machine 952:

It was evident that the discrepancy between the audit (manual count) of the votes reflected on the ballots in the ballot box, and the votes reflected on the scanner result tape, are attributable to the two additional ballots found in the ballot box. The disparity between the manual count, and the scanner result tape, are precisely equal to the votes reflected when counting all off [sic] the ballots in the ballot box – including the two additional ballots. This is not "scanner error," but is instead attributable to any one of a number of alternative possibilities .... The presence of more ballots in the ballot box does not demonstrate that the scanner has "failed," merely that the machine operated as it was designed to do – but with the result that some number (in this case, two) ballots were not scanned. ... PASS



#### 2-4059 Machine 104

1 additional ballot found and explains all discrepancies. ... PASS

#### H18-12 Machine 259

One additional ballot was counted by the scanner than was found in the bin ... PASS

### **Proffered testimony**

The audit results not surprising even if a full hand count would show Mr. Johnson to be the winner.

Substantial possibility that the machine with the largest error was not one of the machines that was audited. 97% chance that auditing 7 of 249 machines won't check the machine with the largest error.

Average of less than two errors per machine could account for the apparent margin of about 450 votes.

Average of one error per 200 ballots could account for the apparent margin.

NY SD 7

Legislation

Risk-limiting audits

Ballot-level audits

#### Proffer, contd.

Not a surprising level of error in precinct-count optically scanned ballots. Consistent with the errors the audit did find, within the statistical variability expected from "the luck of the draw."

Large potential for error: the 242 unaudited machines could hold enough error to account for the apparent margin 186 times over. Sixty-six of the 242 unaudited machines could *individually* hold enough error to account for the apparent margin.

NY SD 7

Legislation

Risk-limiting audits

Ballot-level audits

### Proffer, contd.

Substantial chance that a 3% or 8% audit would find little or no error even if Sen. Johnson is the true winner.

If 30 of the 249 machines have errors of 15 votes or more—enough to account for the apparent margin—chance the 3% audit would have found any of those machines is under 60%.

If 20 of the 249 machines have errors of 23 votes or more–enough to account for the apparent margin–chance the 3% audit would have found any of those machines is under 45%.

If 20 of the 242 unaudited machines have errors of 23 votes or more (enough to account for the apparent margin) and an additional 5% of the machines are audited, chance the additional audit would find any of those 20 is under 69%.

### Proffered testimony, contd.

Margin is so small compared to the possible errors that very large percentage of machines must be audited to give strong evidence that Mr. Martins is indeed the winner.

3% is not sufficient.

8% is not sufficient.

To have 90% statistical confidence that Mr. Martins won requires auditing a minimum of 90% of the machines selected randomly: an additional 218 machines.

This is true if the audit finds that those 218 machines have counted perfectly. If the audit of those 218 machines found many errors, still more machines would have to be audited.

< □ > < 同 > < Ξ > < Ξ > < Ξ > < Ξ < </p>

### What should an election audit law do?

Enunciate principles, not methods.

Methods are best left to regulation: Easier to improve, fix, etc.

Mutual distrust among election integrity advocates, elections officials, and legislators: unfortunate, makes good legislation harder.

California AB 2023 is an example of a good, risk-limiting law: http://www.leginfo.ca.gov/pub/09-10/bill/asm/ab\_ 2001-2050/ab\_2023\_bill\_20100325\_amended\_asm\_v98.html

< □ > < 同 > < Ξ > < Ξ > < Ξ > < Ξ < </p>

### A Better Way: Risk-Limiting Audits

Risk-limiting audit: audit that has a guaranteed minimum chance of correcting a wrong outcome (by counting the whole audit trail). Endorsed by CC, VV, LWV, CEIMN, ASA, ...

**Risk:** maximum chance that the audit fails to correct an apparent outcome that is incorrect, no matter what caused the outcome to be incorrect.

Simultaneous risk-limiting audit: guaranteed minimum chance of correcting *all* the contests that have incorrect apparent outcomes.

Simultaneous risk: the maximum chance that the audit won't correct one or more of the apparent outcomes that are incorrect

## **Risk-limiting Audits in California**

Marin County (February 2008; November 2008, 2009)

Yolo County (November 2008, 2009)

Santa Cruz County (November 2008)

Measures requiring super-majority, simple measures, multi-candidate contests, vote-for-*n* contests.

Contest sizes ranged from about 200 ballots to 121,000 ballots.

Counting burden ranged from 32 ballots to 7,000 ballots.

Cost per audited ballot ranged from nil to about \$0.55.

< □ > < 同 > < Ξ > < Ξ > < Ξ > < Ξ < </p>

### Super-simple simultaneous audits

Truly simple: audit rules that allow elections officials to confirm that the outcomes of most contests are right, with one (small) sample.

Risk-limiting: large chance of correcting any outcomes that are wrong—i.e., that disagree with the outcome full hand count of the audit trail would show. (Correct them by conducting a full hand count.)

Exploit statistical efficiency of *ballot-level auditing*, which compares CVR with human interpretation of individual ballots.

Spend some efficiency to buy logistic and computational simplicity.

Ballot-level audits

▲□▶ ▲□▶ ▲ □▶ ▲ □▶ ▲ □ ● ● ● ●

#### Sample size for ballot-level audits

"diluted"		risk limit	
margin	10%	5%	1%
5%	139	180	277
2%	346	450	691
1%	691	899	1382
0.5%	1382	1798	2764

With these sample sizes, can stop the audit if rate of overstatements is less than 20% of the reported margin.

NY SD 7

Legislation

Risk-limiting audits

Ballot-level audits

#### Secret sauce

To implement ballot-level audits on a wide scale may require changes to vote tabulation systems: have to associate individual cast vote records (CVRs) with individual physical ballots.

Auditing using an unofficial vote tabulation system that does produce CVRs—such as those of Clear Ballot Group, the Humboldt Transparency Project, or TrueBallot—and confirming transitively that the apparent outcome is correct, might be the best interim option. (See Calendrino et al. 2007)

### **IRV and Ballot-Level Auditing**

Even if precinct totals are perfect, cannot reconstruct RCV winner.

Ballot-level audits are (nearly) essential.

(Example by Rivest) Two scenarios, 5 ballots, 3 candidates (A, B, C). Same totals for each candidate for each rank, but different winners.

	Rank order		
Ballot	Scenario 1	Scenario 2	
1	ABC	ACB	
2	ACB	ACB	
3	BAC	BAC	
4	BCA	BAC	
5	САВ	СВА	
winner	A	В	

### What do we need for efficient audits?

- Laws that allow/require risk-limiting audits, but mostly ...
- Data plumbing:

Structured, small batch data export from VTSs. A way to associate individual CVRs with physical ballots.

• Reducing counting effort is mostly about reducing batch sizes.

### What do we need for efficient audits?

- Laws that allow/require risk-limiting audits, but mostly ...
- Data plumbing:

Structured, small batch data export from VTSs. A way to associate individual CVRs with physical ballots.

Reducing counting effort is mostly about reducing batch sizes.

### What do we need for efficient audits?

- Laws that allow/require risk-limiting audits, but mostly ...
- Data plumbing:

Structured, small batch data export from VTSs. A way to associate individual CVRs with physical ballots.

• Reducing counting effort is mostly about reducing batch sizes.