Risk-Limiting Post-Election Audits: Statistics, Policy, and Politics

Philip B. Stark

Department of Statistics, UC Berkeley

1 November 2010
Rice University

Outline

Abstract

Risk-limiting post election audits have a guaranteed minimum chance of correcting any electoral outcomes that are incorrect ("incorrect" means that a full hand count of the audit trail would find a different set of winners). This approach to ensuring the accuracy of election results has been endorsed by the American Statistical Association, Common Cause, The League of Women Voters, Verified Voting, and other groups concerned with election integrity. There have been only six such audits so far, in California. California AB 2023, which requires an official pilot of risk-limiting audits, was signed into law this July after unanimous, bipartisan votes in both legislative bodies. I will discuss the theory behind risk-limiting audits (couching auditing as a sequential nonparametric statistical hypothesis test), factors that affect efficiency and cost, lessons learned in the field conducting audits and working with elections officials, and implementation plans for AB 2023. I also might ramble a bit about the process of getting AB 2023 endorsed and passed and my concerns for its success.

[Election Leak] [CNN: DC hack] [Voting Machine Wins] [Homer Votes—sort of]

Vote-flipping in North Carolina

NC GOP leader: Touchscreen voting machines have programming flaw, by Michael Biesecker

The chairman of the N.C. Republican Party alleged Thursday that a programming flaw with touchscreen voting machines used for early voting in 36 counties is causing votes intended for GOP candidates to be counted for Democrats.

Tom Fetzer, the Republican chairman, said that if the State Board of Elections does not enact a list of demands intended to remedy the problem by the end of today, the party's lawyers will be in federal court Friday morning seeking a statewide injunction.

. . .

Johnnie McLean, deputy director of the state elections board, said Thursday that her office has received no widespread reports of problems.

"In every election we will have scattered reports of machines where the screens need to be recalibrated," McLean said. "That sort of comes with the territory with touch-screen technology."

NEWS OBSERVER, 28 OCTOBER 2010, `http://www.newsobserver.com/2010/10/28/766257/nc-republican-party-chair-touchscreen.html#ixzz13gTJCCvp`

Humboldt County CA, 2008

Serious Error in Diebold Voting Software Caused Lost Ballots in California County, by Kim Zetter

Election officials in a small county in California discovered by chance last week that the tabulation software they used to tally votes in this year's general election dropped 197 paper ballots from the totals at one precinct. The system's audit log also appears to have deleted any sign that the ballots had ever been recorded.

Premier has acknowledged . . . its software caused the system to delete votes. The company has apparently known about the problem since 2004 . . .

[RoV] Crnich would never have discovered the problem through her standard canvassing procedures . . . nor would she have discovered it while conducting a mandatory manual audit that California counties are required to do.

Crnich discovered the missing ballots only because she happened to implement a new and innovative auditing system this year that was spearheaded by members of the public who helped her develop it.

WIRED NEWS, 8 DECEMBER 2008, `http://blog.wired.com/27bstroke6/2008/12/unique-election.html`

Polk County NC, 2008

Owens victory in Polk is in doubt, by Times-News staff

Ted Owens went to sleep Tuesday night thinking he had earned another term . . . A recount Wednesday showed he may not have. . . .

Computer software initially displayed figures that were different than those shown by the voting machines . . .

The software installed in the stand-alone computer that ballot results are fed into was the problem . . . [Elections Director Dale Edwards] said there was no explanation as to why the computer counted the wrong numbers, and no one is at fault.

BLUERIDGENOW.COM TIMES-NEWS, 6 NOVEMBER 2008, http://www.blueridgenow.com/article/20081106/NEWS/811050255

Santa Clara County, CA, 2008

Few problems reported in area despite record turnout, by Karen de Sá and Lisa Fernandez

Record-high voting in the Bay Area on Tuesday mostly defied predictions of unwieldy waits and overwhelmed polls. But in Santa Clara County, concerns about touch-screen voting machines will likely increase following significant malfunctions.

Fifty-seven of the county's Sequoia Voting Systems machines failed on Election Day, resulting in hourslong delays before replacements arrived.

MERCURY NEWS, 4 NOVEMBER 2008, `http://www.mercurynews.com/elections/ci_10901166?nclick_check=1`

Leon County, FL, 2008

Ballots not being recorded at two Leon County polling places, by Angeline J. Taylor

Leon County Supervisor of Elections Ion Sancho has reported that ballots . . . are not being read properly. The problem, he said, rests with a new machine that has been purchased for polling sites throughout the state. . . .

"Certain ballots are being rejected across the state," he said. . . . If the machine reads the ballot card as too long, the . . . machine will simply not read the card.

TALLAHASSEE DEMOCRAT, 20 OCTOBER 2008, `http://www.tallahassee.com/article/20081020/BREAKINGNEWS/81020024`

Palm Beach County, FL, 2008

Florida Primary Recount Surfaces Grave Voting Problems One Month Before Presidential Election, by Kim Zetter

At issue is an August 26 primary election in which officials discovered, during a recount of a close judicial race, that more than 3,400 ballots had mysteriously disappeared after they were initially counted on election day. The recount a week later, minus the missing ballots, flipped the results of the race to a different winner.

. . . officials found an additional 227 ballots that were never counted on election day . . . in boxes in the county's tabulation center.

Palm Beach County was using new optical-scan machines that it recently purchased from Sequoia Voting Systems for $5.5 million.

Palm Beach County, FL, 2008, cont'd

[In a re-scan of ballots the machines had rejected] [o]fficials expected the machines would reject the same ballots again. But that didn't happen. During a first test of 160 ballots, the machines accepted three of them. In a second test of 102 ballots, the machines accepted 13 of them . . . When the same ballots were run through the machines again, 90 of the ballots were accepted.

[T]he county then re-scanned two batches of 51 ballots each that had initially been rejected for having no vote cast in the judicial race, but that were found in a manual examination to contain legitimate votes for one candidate or the other. The first batch of 51 ballots were found to have legitimate votes for Abramson. The second batch of 51 ballots were found to have legitimate votes for Wennet.

In the first batch of 51 ballots . . . 11 of the ballots that had previously been rejected as undervotes were now accepted . . . the remaining 40 ballots were rejected as having no votes. In the second batch of 51 ballots . . . the same machine accepted 2 ballots and rejected 49.

Palm Beach County, FL, 2008, cont'd

The same two batches of ballots were then run through the second . . . machine. [I]n the first batch . . . the machine accepted 41 . . . and rejected 10 others. In the second batch . . . the machine accepted 49 of the ballots and rejected 2—the exact opposite of the results from the first machine.

WIRED NEWS, 7 OCTOBER 2008, `http://blog.wired.com/27bstroke6/2008/10/florida-countys.html`

Washington, DC, 2008

Report Blames Speed In Primary Vote Error; Exact Cause of Defect Not Pinpointed, by Nikita Stewart

Speed might have contributed to the Sept. 9 primary debacle involving thousands of phantom votes, according to a D.C. Board of Elections and Ethics report issued yesterday. . . . [T]he report does not offer a definitive explanation. . .

The infamous Precinct 141 cartridge "had inexplicably added randomly generated numbers to the totals that had been reported," according to the report written by the elections board's internal investigative team.

. . . 4,759 votes were reflected instead of the actual 326 cast there.

WASHINGTON POST, 2 OCTOBER 2008; PAGE B02

see also hearings at
`http://www.octt.dc.gov/services/on_demand_video/channel13/October2008/10_03_08_PUBSVRC_2.asx`

New Jersey 2008

County finds vote errors: Discrepancies discovered in 5% of machines, by Robert Stern

Five percent of the 600 electronic voting machines used in Mercer County during the Feb. 5 presidential primary recorded inaccurate voter turnout totals, county officials said yesterday . . .

23 FEBRUARY 2008, NEW JERSEY TIMES

Ohio 2004

Machine Error Gives Bush Thousands of Extra Ohio Votes, by John McCarthy

COLUMBUS, Ohio – An error with an electronic voting system gave President Bush 3,893 extra votes in suburban Columbus, elections officials said. Franklin County's unofficial results had Bush receiving 4,258 votes to Democrat John Kerry's 260 votes in a precinct in Gahanna. Records show only 638 voters cast ballots in that precinct. Bush's total should have been recorded as 365.

5 NOVEMBER 2004, ASSOCIATED PRESS

Florida 2004

Broward Machines Count Backward, by Eliot Kleinberg

[E]arly Thursday, as Broward County elections officials wrapped up after a long day of canvassing votes, something unusual caught their eye. Tallies should go up as more votes are counted. Thats simple math. But in some races, the numbers had gone . . . down.

Officials found the software used in Broward can handle only 32,000 votes per precinct. After that, the system starts counting backward.

. . . The problem cropped up in the 2002 election. . . . Broward elections officials said they had thought the problem was fixed.

5 NOVEMBER 2004, THE PALM BEACH POST

What's the issue?

- 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.

What's the issue?

- 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.

What's the issue?

- 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.

What's important here?

Crucial question:

When to *stop* auditing, not how many ballots to audit initially.

Solution:

If there's compelling evidence that outcome is right, stop; else, audit more.

Current audit laws have the wrong focus: essentially useless for correcting wrong outcomes. (California just passed AB 2023, which calls for a pilot of a statistically sound approach: risk-limiting audits.)

Efficiency is primarily about batch sizes: Need data plumbing.

California Elections Code §15360

[T]he official conducting the election shall conduct a public manual tally of the ballots tabulated by those devices, including absent voters' ballots, cast in 1 percent of the precincts chosen at random by the elections official . . .

The elections official shall use either a random number generator or other method specified in regulations . . .

The official conducting the election shall include a report on the results of the 1 percent manual tally in the certification of the official canvass of the vote. This report shall identify any discrepancies between the machine count and the manual tally and a description of how each of these discrepancies was resolved . . .

NJ S507 [1R] (Gill)

[officials] shall conduct random hand counts of the voter-verified paper records in at least two percent of the election districts where elections are held for federal or State office . . .

Any procedure designed, adopted, and implemented by the audit team shall be implemented to ensure with at least 99% statistical power that for each federal, gubernatorial or other Statewide election held in the State, a 100% manual recount of the voter-verifiable paper records would not alter the electoral outcome reported by the audit . . .

[procedures] shall be based upon scientifically reasonable assumptions . . . including but not limited to: the possibility that within any election district up to 20% of the total votes cast may have been counted for a candidate or ballot position other than the one intended by the voters[.]

Say what?

Others

Oregon and New Mexico have audit laws that allow the sample (of races and/or ballots) to be selected before the election.

Rep. Rush Holt has proposed federal legislation that has tiered sampling fractions, depending on the margin—but no requirement for followup if errors are found.

Can't correct wrong outcomes without counting the whole audit trail.

What should an election audit law do?

Legislation should 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 is an unfortunate but important consideration.

California AB 2023 (Saldaña, sponsored by SoS Bowen)

First proposed audit bill that limits risk!

(b)(3) “Risk-limiting audit” means a manual tally employing a statistical method that ensures a large, predetermined minimum chance of requiring a full manual tally whenever a full manual tally would show an electoral outcome that differs from the outcome reported by the vote tabulating device for the audited contest. A risk-limiting audit shall begin with a hand tally of the votes in one or more audit units and shall continue to hand tally votes in additional audit units until there is strong statistical evidence that the electoral outcome is correct. In the event that counting additional audit units does not provide strong statistical evidence that the electoral outcome is correct, the audit shall continue until there has been a full manual tally to determine the correct electoral outcome of the audited contest.

http://www.leginfo.ca.gov/pub/09-10/bill/asm/ab_2001-2050/ab_2023_bill_20100325_amended_asm_v98.html

Role of statistics

Limiting the risk is easy

No statistics needed: just count all the ballots by hand.

Statistics lets you do less counting when the outcome is right, but still ensure a big chance of a full hand count when outcome is wrong.

Definitions

Outcome: set of winners, not exact vote totals.

Machine-count outcome, apparent outcome: outcome that will become officially final unless an audit or other action intervenes.

Apparent winner: won according to apparent outcome

Hand-count outcome, true outcome, correct outcome: outcome a full manual tally of the audit trail would show

True winner: would win according to full hand tally, if there were a full hand tally

more definitions . . .

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

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

Crucial ingredients for risk-limiting 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. “Explaining” or “resolving” isn’t enough.
- Only one current audit law limits risk: AB 2023.
- Compliance audits vs. materiality audits.

Crucial ingredients for risk-limiting 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. "Explaining" or "resolving" isn't enough.
- Only one current audit law limits risk: AB 2023.
- Compliance audits vs. materiality audits.

Crucial ingredients for risk-limiting 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. "Explaining" or "resolving" isn't enough.
- Only one current audit law limits risk: AB 2023.
- Compliance audits vs. materiality audits.

Crucial ingredients for risk-limiting 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. "Explaining" or "resolving" isn't enough.
- Only one current audit law limits risk: AB 2023.
- Compliance audits vs. materiality audits.

Crucial ingredients for risk-limiting 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. “Explaining” or “resolving” isn’t enough.
- Only one current audit law limits risk: AB 2023.
- Compliance audits vs. materiality audits.

Crucial ingredients for risk-limiting 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. "Explaining" or "resolving" isn't enough.
- Only one current audit law limits risk: AB 2023.
- Compliance audits vs. materiality audits.

Assessing Evidence

- How strong is the evidence that the outcome is correct, given how the sample was drawn, the margin, the errors found, etc.?
- What is the biggest chance that—if the outcome is wrong—the audit would have found as little error as it did? (The definition of "little" differs across sampling methods, etc.)
- *P*-value of the hypothesis that the apparent outcome of one or more contests is wrong.

Assessing Evidence

- How strong is the evidence that the outcome is correct, given how the sample was drawn, the margin, the errors found, etc.?
- What is the biggest chance that—if the outcome is wrong—the audit would have found as little error as it did? (The definition of “little” differs across sampling methods, etc.)
- *P*-value of the hypothesis that the apparent outcome of one or more contests is wrong.

Assessing Evidence

- How strong is the evidence that the outcome is correct, given how the sample was drawn, the margin, the errors found, etc.?
- What is the biggest chance that—if the outcome is wrong—the audit would have found as little error as it did? (The definition of "little" differs across sampling methods, etc.)
- *P*-value of the hypothesis that the apparent outcome of one or more contests is wrong.

MACRO

Sufficient condition for all outcomes to be right:

For every winner and loser, the overstatement errors minus the understatement errors amount to less than 100% of the margin between that pair of candidates.

MACRO (maximum across-race relative overstatement) summarizes overstatement errors within and across contests.

If the MACRO summed over all ballots is less than 100%, all outcomes of all contests are correct.

Bounding the overstatement error in each batch

Constraints on the error are crucial.

If number of valid ballots cast in batch p for contest c is at most b_{cp} then

$$e_{pw\ell} \leq (v_{wp} - v_{\ell p} + b_{cp})/V_{w\ell}.$$

Hence,

$$e_p \leq \max_{c \in \{1,\ldots,C\}} \max_{w \in \mathcal{W}_c, \ell \in \mathcal{L}_c} \frac{v_{wp} - v_{\ell p} + b_{cp}}{V_{w\ell}} \equiv u_p.$$

u_p is a limit on e_p, the maximum relative overstatement of *any* margin that can be concealed in batch p, the MACRO in batch p.

$U \equiv \sum_p u_p$, bound on total error.

Sampling Designs

Generally, cluster samples of some kind.

Simple, Stratified (by county, voting method, other), PPEB/PPS, NEGEXP

Stratified PPEB?

Sampling scheme affects choice of test statistic—analytic tractability

Weighted max, binning for simple & stratified sampling, NEGEXP, PPEB.

More efficient choices possible for PPEB: Kaplan-Markov

Taint

taint of batch p

$$\tau_p = \frac{e_p}{u_p} \le 1.$$

Draw batches with replacement s.t. in each draw

$$\mathbf{P}\{\text{draw batch } p\} = u_p/U.$$

PPS, used in financial auditing.

Taint of jth draw is T_j. $\{T_j\}$ are iid, $\mathbf{E}T_j = E/U$.

Can stop the audit if can reject the hypothesis $\mathbf{E}T_j \ge 1/U$.

Reduces auditing to testing hypothesis about the mean of a bounded random variable.

Sequential risk-limiting audit using Kaplan-Markov bound

0. Calculate error bounds $\{u_p\}$, U. Set $n = 1$. Pick $\alpha \in (0, 1)$ and $m > 0$.
1. Draw a batch using PPEB. Audit it if it has not already been audited.
2. Find $T_n \equiv t_p \equiv e_p/u_p$, taint of the batch p drawn at stage n.
3. Compute

$$P_n \equiv \prod_{j=1}^{n} \frac{1 - 1/U}{1 - T_j}. \quad \text{See this month's WIRED, p.56} \tag{1}$$

- 4. If $P_n < \alpha$, stop; report apparent outcomes. If $n = m$, audit remaining batches. If all batches have been audited, stop; report known outcomes. Else, $n \leftarrow n + 1$ and go to 1.

Sequential risk-limiting audit using Kaplan-Markov bound

0. Calculate error bounds $\{u_p\}$, U. Set $n = 1$. Pick $\alpha \in (0, 1)$ and $m > 0$.
1. Draw a batch using PPEB. Audit it if it has not already been audited.
2. Find $T_n \equiv t_p \equiv e_p/u_p$, taint of the batch p drawn at stage n.
3. Compute

$$P_n \equiv \prod_{j=1}^{n} \frac{1 - 1/U}{1 - T_j}. \quad \text{See this month's WIRED, p.56} \tag{1}$$

- 4. If $P_n < \alpha$, stop; report apparent outcomes. If $n = m$, audit remaining batches. If all batches have been audited, stop; report known outcomes. Else, $n \leftarrow n + 1$ and go to 1.

Sequential risk-limiting audit using Kaplan-Markov bound

0. Calculate error bounds $\{u_p\}$, U. Set $n = 1$. Pick $\alpha \in (0, 1)$ and $m > 0$.
1. Draw a batch using PPEB. Audit it if it has not already been audited.
2. Find $T_n \equiv t_p \equiv e_p / u_p$, taint of the batch p drawn at stage n.
3. Compute

$$P_n \equiv \prod_{j=1}^{n} \frac{1 - 1/U}{1 - T_j}. \quad (1)$$

See this month's WIRED, p.56

- 4. If $P_n < \alpha$, stop; report apparent outcomes. If $n = m$, audit remaining batches. If all batches have been audited, stop; report known outcomes. Else, $n \leftarrow n + 1$ and go to 1.

Sequential risk-limiting audit using Kaplan-Markov bound

0. Calculate error bounds $\{u_p\}$, U. Set $n = 1$. Pick $\alpha \in (0, 1)$ and $m > 0$.
1. Draw a batch using PPEB. Audit it if it has not already been audited.
2. Find $T_n \equiv t_p \equiv e_p/u_p$, taint of the batch p drawn at stage n.
3. Compute

$$P_n \equiv \prod_{j=1}^{n} \frac{1 - 1/U}{1 - T_j}. \tag{1}$$

See this month's WIRED, p.56

- 4. If $P_n < \alpha$, stop; report apparent outcomes. If $n = m$, audit remaining batches. If all batches have been audited, stop; report known outcomes. Else, $n \leftarrow n + 1$ and go to 1.

Sequential risk-limiting audit using Kaplan-Markov bound

0. Calculate error bounds $\{u_p\}$, U. Set $n = 1$. Pick $\alpha \in (0, 1)$ and $m > 0$.
1. Draw a batch using PPEB. Audit it if it has not already been audited.
2. Find $T_n \equiv t_p \equiv e_p/u_p$, taint of the batch p drawn at stage n.
3. Compute

$$P_n \equiv \prod_{j=1}^{n} \frac{1 - 1/U}{1 - T_j}. \tag{1}$$

See this month's WIRED, p.56

- 4. If $P_n < \alpha$, stop; report apparent outcomes. If $n = m$, audit remaining batches. If all batches have been audited, stop; report known outcomes. Else, $n \leftarrow n + 1$ and go to 1.

This sequential procedure is risk-limiting

If any outcome is wrong,

$$\mathbf{P}\{\text{stop without auditing every batch}\} < \alpha.$$

Chance $\ge 1 - \alpha$ of correcting wrong outcomes by full hand count.

Remarkably efficient if batches are not too big.

Pilot 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.

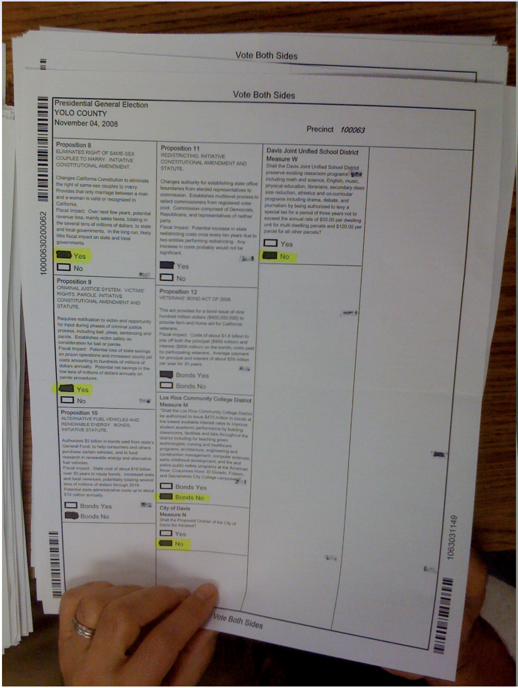
Counting burden ranged from 32 ballots to 7,000 ballots.

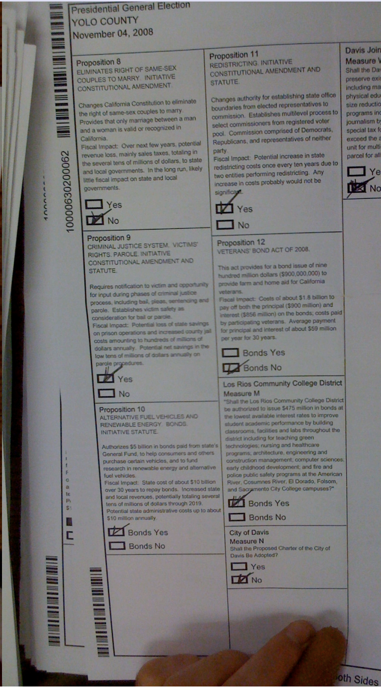
Cost per audited ballot ranged from nil to about \$0.55.

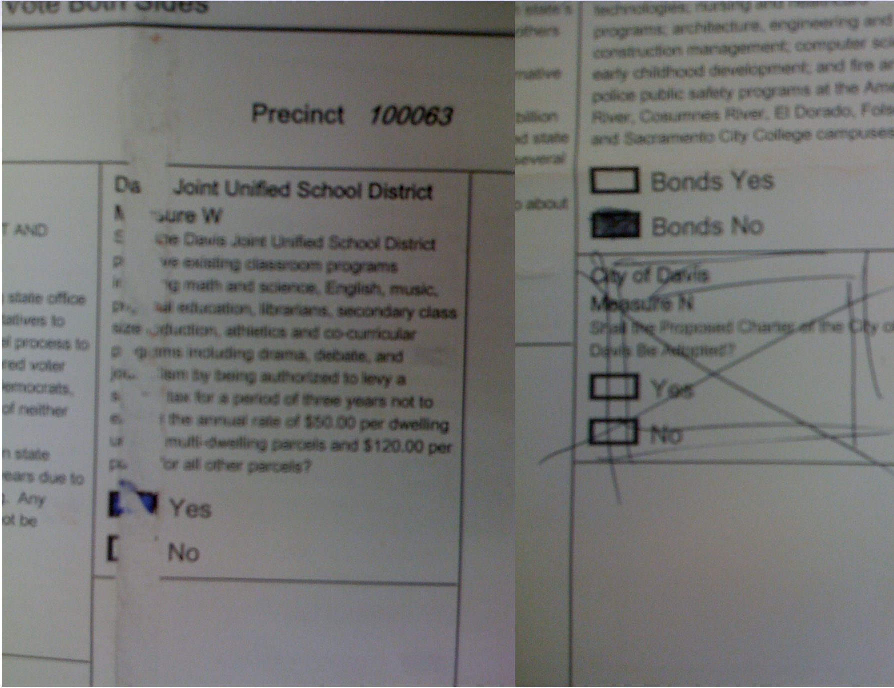
2008 Yolo County, CA Measure W Audit


Hefty
Voted Ballots

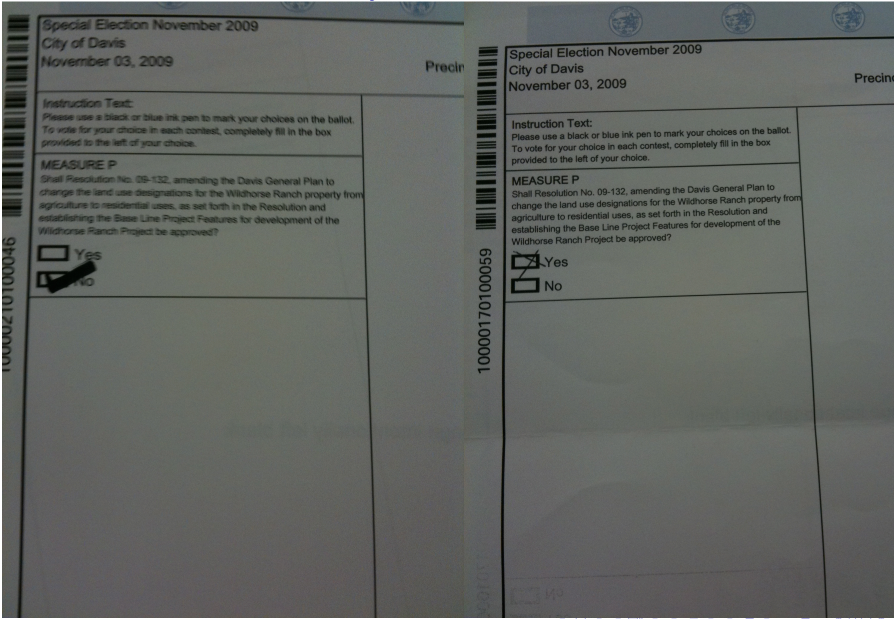

TALLY SHEET
NUMBER OF VOTES CAST FOR EACH CANDIDATE
450
153
33
483
153
SIGN CERTIFICATE ON FRONT COVER
TALLY SHEET
NUMBER OF VOTES CAST FOR EACH CANDIDATE
Measure W Yes
450
Measure W No
153
Measure W Yes
33
Measure W No
SIGN CERTIFICATE ON FRONT COVER

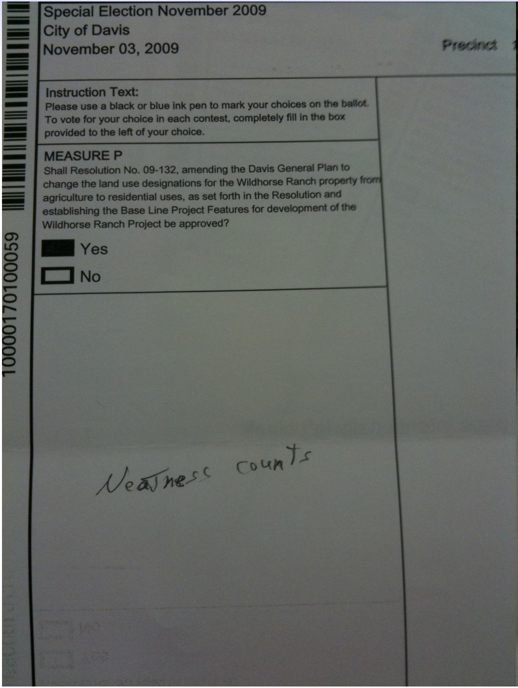

Vote Both Sides
Presidential General Election
YOLO COUNTY
November 04, 2008
Precinct 100063
Proposition 8
Proposition 9
Proposition 10
Proposition 11
Proposition 12
Davis Joint Unified School District
Measure W
Los Rios Community College District
Measure M
City of Davis
Measure N
Yes
No
Bonds Yes
Bonds No
Vote Both Sides

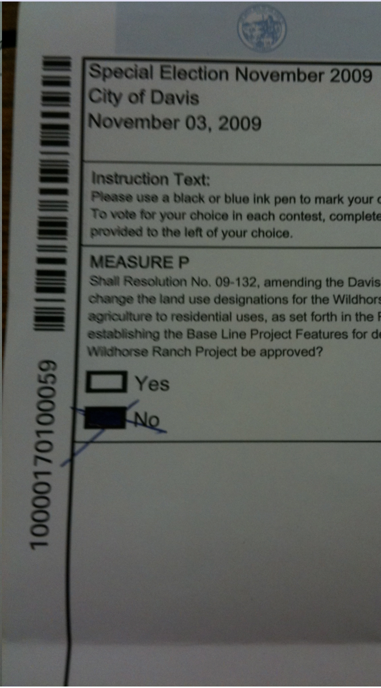

Presidential General Election
YOLO COUNTY
November 04, 2008
Proposition 8
ELIMINATES RIGHT OF SAME-SEX COUPLES TO MARRY. INITIATIVE CONSTITUTIONAL AMENDMENT.
Proposition 9
CRIMINAL JUSTICE SYSTEM. VICTIMS' RIGHTS. PAROLE. INITIATIVE CONSTITUTIONAL AMENDMENT AND STATUTE.
Proposition 10
ALTERNATIVE FUEL VEHICLES AND RENEWABLE ENERGY. BONDS. INITIATIVE STATUTE.
Proposition 11
REDISTRICTING. INITIATIVE CONSTITUTIONAL AMENDMENT AND STATUTE.
Proposition 12
VETERANS' BOND ACT OF 2008.
Los Rios Community College District
Measure M
City of Davis
Measure N
Shall the Proposed Charter of the City of Davis Be Adopted?
Yes
No
Bonds Yes
Bonds No


Precinct 100063
Joint Unified School District
Yes
No
Bonds Yes
Bonds No
City of Davis
Measure N
Yes
No

2009 Yolo County, CA Measure P Audit




Special Election November 2009
City of Davis
November 03, 2009
Precinct
Instruction Text:
Please use a black or blue ink pen to mark your choices on the ballot.
To vote for your choice in each contest, completely fill in the box
provided to the left of your choice.
MEASURE P
Shall Resolution No. 09-132, amending the Davis General Plan to
change the land use designations for the Wildhorse Ranch property from
agriculture to residential uses, as set forth in the Resolution and
establishing the Base Line Project Features for development of the
Wildhorse Ranch Project be approved?
Yes
No
Neatness counts
10000170100059


Special Election November 2009
City of Davis
November 03, 2009
Instruction Text:
Please use a black or blue ink pen to mark your
To vote for your choice in each contest,
provided to the left of your choice.
MEASURE P
Shall Resolution No. 09-132, amending the Davis
change the land use designations for the
agriculture to residential uses, as set forth in the
establishing the Base Line Project Features for
Wildhorse Ranch Project be approved?
Yes
No
10000170100059

Yolo County Measure P, November 2009

Reg. voters	ballots	precincts	batches	yes	no
38,247	12,675	31	62	3,201	9,465

(VBM) and in-person (IP) ballots were tabulated separately (62 batches).

$U = 3.0235$.

For $\alpha = 10\%$, initial sample size 6 batches; gave 4 distinct batches, 1,437 ballots.

Ballot-level auditing would save *lots* of work

Can determine the initial sample size for a Kaplan-Markov ballot-level audit even though the cast vote records (CVRs) were not available.

For $\alpha = 10\%$ would need to look at CVRs for $n = 6$ ballots.

For $\alpha = 1\%$, $n = 12$ ballots.

C.f., 1,437 ballots for actual batch sizes.

Super-simple simultaneous audits

Goal

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.

Have to match CVRs to physical ballots. Requires new voting systems or *transitive auditing* using parallel systems (e.g., Clear Ballot Group, Humboldt ETP, TrueBallot) *a la* Calendrino et al. (2007)

Advantages of super-simple method

- Audit entire collection of contests with one simple random sample of ballots.
- Super simple: initial sample size is a constant—the sample size multiplier ρ—divided by the "diluted margin." ρ set once and for all: doesn't depend on any particulars of the contests, margins, etc.
- Audit expands if too many ballots with errors that overstate a margin by one vote, or any ballots that overstate a margin by two votes. Determining when to stop is simple.
- Chance of correcting all wrong outcomes is guaranteed to be at least as high as claimed.

Advantages of super-simple method

- Audit entire collection of contests with one simple random sample of ballots.
- Super simple: initial sample size is a constant—the sample size multiplier ρ—divided by the “diluted margin.” ρ set once and for all: doesn’t depend on any particulars of the contests, margins, etc.
- Audit expands if too many ballots with errors that overstate a margin by one vote, or any ballots that overstate a margin by two votes. Determining when to stop is simple.
- Chance of correcting all wrong outcomes is guaranteed to be at least as high as claimed.

Advantages of super-simple method

- Audit entire collection of contests with one simple random sample of ballots.
- Super simple: initial sample size is a constant—the sample size multiplier ρ—divided by the "diluted margin." ρ set once and for all: doesn't depend on any particulars of the contests, margins, etc.
- Audit expands if too many ballots with errors that overstate a margin by one vote, or any ballots that overstate a margin by two votes. Determining when to stop is simple.
- Chance of correcting all wrong outcomes is guaranteed to be at least as high as claimed.

Advantages of super-simple method

- Audit entire collection of contests with one simple random sample of ballots.
- Super simple: initial sample size is a constant—the sample size multiplier ρ—divided by the "diluted margin." ρ set once and for all: doesn't depend on any particulars of the contests, margins, etc.
- Audit expands if too many ballots with errors that overstate a margin by one vote, or any ballots that overstate a margin by two votes. Determining when to stop is simple.
- Chance of correcting all wrong outcomes is guaranteed to be at least as high as claimed.

Ballot-level auditing

	contest				
	1	2	3	4	5
CVR	undervote	winner	loser	not on ballot	not on ballot
Hand	loser	loser	winner	not on ballot	not on ballot
overstatement	1	2	-2	0	0

Hypothetical CVR and hand interpretation of a ballot that contains three of five contests under audit. “Winner” and “loser” denote an apparent winner and an apparent loser, respectively. The maximum overstatement is two votes.

	contest				
	1	2	3	4	5
CVR	undervote	winner	loser	winner	not on ballot
Hand	loser	overvote	loser	winner	not on ballot
overstatement	1	1	0	0	0

Hypothetical CVR and hand interpretation of a ballot that contains four of five contests under audit. "Winner" and "loser" denote an apparent winner and an apparent loser, respectively. In contest 3, the CVR and hand count found votes for one and the same apparent loser, and in contest 4, the CVR and hand count found votes for one and the same apparent winner. There are two overstatement errors, but the maximum overstatement is one vote.

New procedure requires setting 3 numbers:

- simultaneous risk limit α. Might be set by legislation.
- error inflation factor $\gamma \geq 100\%$. Controls tradeoff between initial sample size and additional counting when the sample finds many overstatements. γ ffects operating characteristics but not risk.
- error tolerance $\lambda < 100\%$. Tolerable rate of 1-vote maximum overstatements in initial sample as fraction of μ. If rate of ballots in the sample with 1-vote maximum overstatements is no more than $\lambda\mu$ and there are no 2-vote overstatement, audit stops. λ affects operating characteristics but not risk.

New procedure requires setting 3 numbers:

- simultaneous risk limit α. Might be set by legislation.
- error inflation factor $\gamma \geq 100\%$. Controls tradeoff between initial sample size and additional counting when the sample finds many overstatements. γ ffects operating characteristics but not risk.
- error tolerance $\lambda < 100\%$. Tolerable rate of 1-vote maximum overstatements in initial sample as fraction of μ. If rate of ballots in the sample with 1-vote maximum overstatements is no more than $\lambda\mu$ and there are no 2-vote overstatement, audit stops. λ affects operating characteristics but not risk.

New procedure requires setting 3 numbers:

- simultaneous risk limit α. Might be set by legislation.
- error inflation factor $\gamma \geq 100\%$. Controls tradeoff between initial sample size and additional counting when the sample finds many overstatements. γ ffects operating characteristics but not risk.
- error tolerance $\lambda < 100\%$. Tolerable rate of 1-vote maximum overstatements in initial sample as fraction of μ. If rate of ballots in the sample with 1-vote maximum overstatements is no more than $\lambda\mu$ and there are no 2-vote overstatement, audit stops. λ affects operating characteristics but not risk.

Super-simple simultaneous procedure

1. Pick risk limit $\alpha \in (0, 1)$, $\gamma \geq 100\%$, $\lambda < 100\%$
2. Calculate the sample-size multiplier ρ

$$\rho = \frac{-\log \alpha}{\frac{1}{2\gamma} + \lambda \log(1 - \frac{1}{2\gamma})}.$$

   For $\alpha = 10\%$, $\gamma = 110\%$ and $\lambda = 50\%$, $\rho = 15.2$. ρ doesn't depend on the audit data or particulars of the contests.
3. Calculate the diluted margin μ.
4. Audit simple random sample of at least $n = \lceil \rho/\mu \rceil$ ballots. If fewer than $n\lambda\mu$ of those have one-vote maximum overstatements and none has a two-vote overstatement, stop. Otherwise, Kaplan-Markov *P*-value determines when to stop.

Super-simple simultaneous procedure

1. Pick risk limit $\alpha \in (0, 1)$, $\gamma \geq 100\%$, $\lambda < 100\%$
2. Calculate the sample-size multiplier ρ

$$\rho = \frac{-\log \alpha}{\frac{1}{2\gamma} + \lambda \log(1 - \frac{1}{2\gamma})}.$$

   For $\alpha = 10\%$, $\gamma = 110\%$ and $\lambda = 50\%$, $\rho = 15.2$. ρ doesn't depend on the audit data or particulars of the contests.
3. Calculate the diluted margin μ.
4. Audit simple random sample of at least $n = \lceil \rho/\mu \rceil$ ballots. If fewer than $n\lambda\mu$ of those have one-vote maximum overstatements and none has a two-vote overstatement, stop. Otherwise, Kaplan-Markov P-value determines when to stop.

Super-simple simultaneous procedure

1. Pick risk limit $\alpha \in (0, 1)$, $\gamma \geq 100\%$, $\lambda < 100\%$
2. Calculate the sample-size multiplier ρ

$$\rho = \frac{-\log \alpha}{\frac{1}{2\gamma} + \lambda \log(1 - \frac{1}{2\gamma})}.$$

   For $\alpha = 10\%$, $\gamma = 110\%$ and $\lambda = 50\%$, $\rho = 15.2$. ρ doesn't depend on the audit data or particulars of the contests.
3. Calculate the diluted margin μ.
4. Audit simple random sample of at least $n = \lceil \rho/\mu \rceil$ ballots. If fewer than $n\lambda\mu$ of those have one-vote maximum overstatements and none has a two-vote overstatement, stop. Otherwise, Kaplan-Markov P-value determines when to stop.

Super-simple simultaneous procedure

1. Pick risk limit $\alpha \in (0, 1)$, $\gamma \geq 100\%$, $\lambda < 100\%$
2. Calculate the sample-size multiplier ρ

$$\rho = \frac{-\log \alpha}{\frac{1}{2\gamma} + \lambda \log(1 - \frac{1}{2\gamma})}.$$

   For $\alpha = 10\%$, $\gamma = 110\%$ and $\lambda = 50\%$, $\rho = 15.2$. ρ doesn't depend on the audit data or particulars of the contests.
3. Calculate the diluted margin μ.
4. Audit simple random sample of at least $n = \lceil \rho/\mu \rceil$ ballots. If fewer than $n\lambda\mu$ of those have one-vote maximum overstatements and none has a two-vote overstatement, stop. Otherwise, Kaplan-Markov P-value determines when to stop.

Examples

diluted margin μ	$\lambda = 50\%$ risk limit α 10%	5%	1%	$\lambda = 20\%$ risk limit α 10%	5%	1%
5%	305	396	609	139	180	277
2%	761	989	1521	346	450	691
1%	1521	1978	3041	691	899	1382
0.5%	3041	3956	6081	1382	1798	2764
multiplier ρ	15.20	19.78	30.40	6.91	8.99	13.82

Initial sample sizes *n* and multipliers ρ for $\gamma = 110\%$. Column 1: diluted margin of victory μ. Columns 2–4: *n* for various risk limits if the audit is to stop when the percentage of ballots in the sample that overstate a margin by 1 vote is not more than 50% of the diluted margin. Columns 5–7: *n* for various simultaneous risk limits if audit is to stop when the percentage of ballots in the sample that overstate a margin by 1 vote is not more than 20% of the diluted margin. Last row: sample sizes *n* are equal to these "multipliers" divided by diluted margin μ.

Conclusions

Very simple formula for initial sample for risk-limiting audit. Allows audit to stop if, in the initial sample, rate of 1-vote maximum overstatements is at most a pre-specified fraction of the margin and there are no 2-vote overstatements.

Method requires choosing 3 numbers.

Simple but somewhat inefficient: More ballots have to be counted by hand than if sharper bounds were used, but those methods require far more complex math.

Ballot-level audits are so efficient that total cost still low.

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)

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.