Evidence-Based Elections

Philip B. Stark

Department of Statistics
University of California, Berkeley

Data, Society, and Inference Seminar
Stanford University and University of California, Berkeley
8 October 2012

Credit

Many collaborators, including:

Josh Benaloh, Joe Hall, Mike Higgins, Doug Jones, Eric Lazarus, Mark Lindeman, Luke Miratrix, Olivier Pereira, Ron Rivest, David Wagner, Dan Wallach, Kai Wang, Vince Yates.

Lots of help from elections officials, especially:

Jennie Bretschneider, Elaine Ginnold, Neal Kelley, Freddie Oakley, Tom Stanionis.

No right is more precious in a free country than that of having a voice in the election of those who make the laws under which, as good citizens, we must live. Other rights, even the most basic, are illusory if the right to vote is undermined. – US Supreme Court, 1964 (Wesberry et al. v Sanders)

It doesn't matter who votes. What matters is who counts the votes. – Josef Stalin

The purpose of elections is to convince the losers that they lost. – Dan Wallach

The difference between theory and practice is smaller in theory than it is in practice. – Various

Palm Beach

Software maker takes blame in Wellington vote count mess, by George Bennett

The supplier of Palm Beach County's voting and tabulating equipment says a software "shortcoming" led to votes being assigned to the wrong candidates and the elections office declaring the wrong winners in two recent Wellington council races. . . . Unbeknownst to elections officials, the vote totals for the mayor's race ended up being reported and later certified as the results of the Seat 1 race. The Seat 1 vote totals were certified as the Seat 4 results and the Seat 4 vote totals were certified as the mayoral results.

The problem wasn't discovered until six days after the election, during a routine audit. . . . The fact that the audit is conducted after winners are certified is a requirement of state law.

THE PALM BEACH POST, 23 MARCH 2012, `http://www.sun-sentinel.com/news/palm-beach/pb-bucher-election-machines-20120323,0,7453964.story`

NY

Board of Elections does nothing as hundreds of Bronx votes go missing

More than six months ago, voting experts at New York University Law School's Brennan Center detected an alarming pattern at one polling place in the South Bronx: The tallies from the electronic scanning machines at Public School 65 included high proportions of invalidated votes. . . . The board did nothing. Actually, the board did worse than nothing. It refused to check – even when asked to do so by state election officials. . . . [W]e discovered that voters had done their part correctly, while one of the three scanners at PS 65 misread and miscounted votes. Here are the disgraceful findings:

In the September primary, the scanner processed 103 ballots and made errors on 69 of them, a failure rate approaching 70%.

In the November general election, the scanner handled 289 ballots and misread votes on 156 of them, a 54% failure rate.

NEW YORK DAILY NEWS, 27 FEBRUARY 2012,

http://www.nydailynews.com/opinion/voters-damned-article-1.1028275#ixzz1nb60Oaz2

Saguache County, Colorado crimes

Grand jury has its hands full with Saguache election case, by Troy Hooper

A disputed election in south-central Colorado is now in the hands of a grand jury that is reviewing allegations that the clerk and other officials committed crimes when they tallied the votes.

The officials under investigation stood to benefit from the election's outcome — most notably Saguache County Clerk Melinda Myers — who, along with County Commissioner Linda Joseph, at first lost but then won their races after Myers declared the races had to be retabulated due to a technical glitch.

[Myers won't let the Colorado Secretary of State inspect the ballots.] "There are processes that we are avowed to protect," [Colorado County Clerks] association president Scott Doyle said. "One of them is preserving the sanctity of ballots. The cornerstone of our democracy is based on those ballots. It's what we stand for as clerks."

"The clerks are using the false argument about 'secrecy of ballots' as a scare tactic or sympathy evoking tool to try to get a trusting public to side with them in their effort to block public verification of elections," Al Kolwicz of the Colorado Voter Group said in an email. "Why exactly clerks oppose public verification is unknown."

Officials in Saguache County stand accused of more than 30 misdemeanors. [Myers was recalled this year by a 60% vote.]

THE COLORADO INDEPENDENT, 25 MARCH 2011,

`http://coloradoindependent.com/80819/grand-jury-has-its-hands-full-with-saguache-election-case`

Waukesha County, WI: Oops!

Wisconsin Election Surprise: David Prosser Gains 7,500 Votes After 'Human Error' In Waukesha County, by Amanda Terkel

In a dramatic turn of events on Thursday, the Waukesha County clerk announced that the vote total announced for Tuesday's Wisconsin Supreme Court race had been mistaken – and that the corrected numbers changed the outcome of the entire election.

There were 3,456 missing votes for Democratic-backed challenger JoAnne Kloppenburg and 11,059 for incumbent GOP-backed Justice David Prosser. Kloppenburg has previously been beating Prosser by just 200 votes of the roughly 1.5 million cast statewide.

In the city of New Berlin, the total for one ward was recorded as 37 votes for Prosser, but it was actually 237, she said. In the town of Lisbon, a "typing error" resulted in both candidates losing votes. The most significant error, however, occurred in the city of Brookfield.

"The spreadsheet from Brookfield was imported into a database that was provided by the Government Accountability Board, but it inadvertently was not saved," Nickolaus said. "As a result, when I ran the report to show the aggregate numbers that were collected from all the municipalities, I assumed that the city of Brookfield was included. It was not. The city of Brookfield cast 14,315 votes on April 5 – 10,859 votes went for Justice David Prosser, 3,456 went for JoAnne Kloppenburg."

. . . prior to the election, Nickolaus "was heavily criticized for her decision to keep the county results on an antiquated personal computer, rather than upgrade to a new data system being utilized statewide."

"Nickolaus cited security concerns for keeping the data herself . . . "

HUFFINGTON POST, 7 APRIL 2011,

`http://www.huffingtonpost.com/2011/04/07/david-prosser-wisconsin-supreme-court_n_846431.html`

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

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 to do?

- HAVA pushed the country to electronic voting systems without serious consideration of vulnerability and verifiability.
- Any means of counting votes can make errors.
- Because of error, the wrong candidates can appear to win.
- What can be done?
- Two basic responses: certify the equipment, audit
- Claim: auditing is the better tool, but current audit laws are inadequate and only 75% of voters create audit trail

Certification of voting systems

- Tests may include source code review, environmental testing, accuracy testing, drop-tests, etc.
- EAC oversees federal certification.
- Many states require federal certification, state certification, or both.
- Certification expensive–can cost $millions
- Certification time-consuming–can take years.
- Not a guarantee that the equipment will be used properly.
- Not a guarantee that the equipment will work properly when it matters: In elections.
- Hence, not a guarantee that outcomes are right.

What's the right question?

1. Under laboratory conditions, can the vote tabulation system—as delivered from the manufacturer—count votes with a specified level of accuracy?
2. As maintained, deployed, and used in the current election, did the vote tabulation system find the true winners?

Certification addresses Q 1. Leads to things like jurisdictions combing eBay for Zip drives.

Q 2 seems more important. Audits address Q 2.

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[.]

Others

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

Florida does not allow auditing before results are final; limits the amount of auditing.

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.

Counting the whole audit trail won't give right answer unless it's adequately intact.

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.

What should an audit do?

- Check the evidence that the outcome is right.
- If there's convincing evidence, bless the results.
- If there's compelling evidence that the outcome is wrong, correct the results.
- Otherwise, admit that the evidence trail is weak. Outcome should be decided by other means.

What is wanting?

- Law/regulations should require LEOs to give *convincing evidence* that outcomes are right.
- Does not necessarily require radical transparency—but requires a good audit trail.
- Certifying equipment isn't enough: How was the equipment used?
- Election should generate hard evidence, checked for integrity.
- Audit trail needs to be scrutinized to confirm or correct the outcome.
- "I'm good at my job" is widely true, but is not convincing evidence: stuff happens. Often.
- Why certify equipment but not procedures, especially curation of the audit trail?

Foundations

Strongly Software-Independent Voting System (Rivest & Wack)

A voting system is strongly software-independent if an undetected error or change to its software cannot produce an undetectable change in the outcome, and we can find the correct outcome without re-running the election.

Risk-limiting Audit

Large, known chance of a full hand count if the outcome is wrong, thereby correcting the outcome.

Risk is maximum chance of failing to correct an apparent outcome that is wrong, no matter what caused the outcome to be wrong.

Evidence-based elections

$$\text{Evidence} = \text{Auditability} + \text{Auditing}.$$

Resilient Canvass Framework

Known minimum chance that the overall system (human, hardware, software, procedures) gives the correct election outcome—when it gives an outcome.

- Use voting system that creates a voter-verifiable audit trail.
- Conduct a compliance audit to ensure that—as actually used in this election—the system is strongly software-independent.
- If so, conduct a risk-limiting audit. If not, do not declare an outcome.

Resilience: Overall election and canvass process should correct its own errors before reporting, or report it can't guarantee that it corrected its errors (e.g., because evidence that the audit trail is intact is too weak).

Ingredients for resilient canvass framework

- Voters create complete, durable, accurate audit trail.
- LEO curates the audit trail adequately.
- Compliance audit to ensure that the audit trail is adequately intact.
  Was the system, as used, strongly software independent?
  If not, don't declare an outcome.
- Risk-limiting audit: Examine ballots by hand until there's strong evidence that counting the rest won't change the outcome.
  "Explaining" or "resolving" errors isn't enough.

Compliance Audits and Materiality Audits

Effective compliance audit

Determine whether the audit trail is trustworthy enough to determine who won.

If not, do not declare an outcome (nb: danger of DOS attacks).

Effective materiality audit

Correct the outcome if it is wrong.

Requires intact audit trail–need to pass compliance audit first.
Might require counting the entire audit trail by hand.

Compliance audit: Check creation & curation of audit trail

- Did election use equipment that should create an accurate audit trail and adhere to procedures that should keep the audit trail sufficiently accurate to reflect the outcome according to how voters actually voted?
- Should include ballot accounting, checks of seals, chain of custody, surveillance tapes, forensic dismantling of voting machines, etc.
- If compliance audit generates convincing affirmative evidence that a full hand count of the audit trail would show the outcome according to how votes were cast, proceed to risk-limiting audit.
- This evidence is qualitative, like legal evidence: convincing to hypothetical “reasonable person.”
- If insufficient evidence that the outcome is right, don’t declare election outcome.

Materiality audit: check outcome against audit trail

- Did the vote tabulation system count the votes accurately enough to determine who won?
- Relies on the audit trail, which the compliance audit has checked for integrity.
- If hand-to-eye check of sample of ballots generates convincing evidence that a full hand count of the audit trail would show the same outcome that the VTS reported, stop.
- Evidence is quantitative statistical evidence.
- If insufficient evidence, expand the sample and count more votes by hand. Keep expanding until there's convincing evidence or until there has been a full hand count.

What's the question?

- Detection paradigm: If the outcome is wrong, ensure a big chance of finding at least one error.
- But audits almost invariably find at least one error. What then?
- What do we want audits to accomplish?
- One possibility: correct wrong electoral outcomes.
- Risk-limiting paradigm: If the outcome is wrong, ensure a big chance of correcting it.

Risk-limiting audits

- Historically, much debate over how large a sample to start with. Sideways.
- Crucial question: When to *stop* auditing [not how big a sample to start with].
- Answer: If there's compelling evidence that outcome is right, stop; else, audit more.
  Measure evidence by *P*-value.
- Eventually, either have strong evidence that the outcome is right, or the whole contest has been counted by hand and correct outcome is known.
- Sequential test of the null hypothesis that the outcome is wrong. "Risk" is chance of type I error: concluding a wrong outcome is right. Can control rigorously. No possibility of a type II error.

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.

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.
- Comparison audit:
  1. Commit to vote data at some level of aggregation.
  2. Check that the committed data produces the same results as claimed. Should be perfect.
  3. Sample the committed data and check until there is strong evidence that it is accurate enough to find the right election outcome.

Tradeoffs

- Ballot polling audit
  - Virtually no set-up costs
  - Requires nothing of voting system
  - Preserves voter anonymity except possibly for sampled ballots
  - Counting burden comparable to precinct-based comparison audit unless margin is small
  - Requires more counting than ballot-level comparison audit
  - Does not check tabulation: lucky cancellation of errors possible
- Comparison audit
  - Heavy demands on voting system for reporting and export
  - Requires LEO to commit to subtotals
  - Requires ability to retrieve ballots that correspond to CVRs or subtotals
  - May compromise voter privacy (small-batch or ballot-level reporting) & enable coercion through pattern voting
  - Most efficient (ballot-level) may require re-scanning all ballots
  - Checks tabulation (but not for *transitive audits* unless subtotals are cross checked as well)
  - Ballot-level comparison audits require least hand counting

Risk-Limiting Audits

- 16 pilot audits in CA, CO, and OH; another 14 planned.
- EAC funding for pilots in CA and CO and Cuyahoga County, OH
- CO has law; CA has pilot law
- simple measures
- measures requiring super-majority
- multi-candidate contests
- vote-for-*n* contests,
- multiple contests audited simultaneously with one sample
- contest sizes: 200 ballots to 121,000 ballots
- counting burden: 16 ballots to 7,000 ballots
- cost per audited ballot: nil to about $0.55.

California AB 2023 (Saldaña, sponsored by SoS Bowen)

(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`

California AB 2023 backstory

I testified to both houses of California legislature, worked with individual counties and CACEO, etc.

Happy to tell stories later.

Definitions

- *Audit trail* or *ballot*: indelible record of how voters cast their votes, e.g., voter-marked paper ballot or VVPAT.
- *Outcome* of a contest: set of winners, not the exact vote counts.
- *Apparent outcome*: winner or winners according to the voting system.
- *Correct outcome*: winner or winners that a full hand count of the audit trail would find.
- Apparent outcome is *wrong* if it isn't the outcome a full hand count of the audit trail would show.

Risk-limiting audits

- *Risk-limiting audit*: pre-specified minimum chance of correcting apparent outcome if apparent outcome is wrong. (Endorsed by ASA, CC, VV, LWV, CEIMN, . . .)
- *Risk*: largest possible chance an apparent outcome that's wrong won't be caught and corrected—no matter why it's wrong.
- *Simultaneous risk-limiting audit*: pre-specified minimum chance of correcting all incorrect apparent outcomes in the election.
- *Simultaneous risk*: largest possible chance that one or more wrong outcomes won't be caught and corrected—no matter why they are wrong.

Assessing Evidence

- How strong is the evidence that the outcome is correct, given how the sample was drawn, the margin, etc.?
- What is the biggest chance that—if the outcome is wrong—the audit would have found what it did?
- (Maximum) *P*-value of the hypothesis that the apparent outcome of one or more contests is wrong.

Ballot-polling audits and Comparison Audits

- Comparison audit:
  1. LEO “commits” to vote data at some level of aggregation.
  2. Audit checks that the committed data produces the same results as claimed. Should be perfect.
  3. Audit samples and checks the committed data until there is strong evidence that the data are accurate enough to produce the right election outcome (or until the true outcome is known).
- Ballot polling audit: Sample/examine ballots until there is strong evidence that looking at the rest would confirm the outcome (or until the true outcome is known).

Tradeoffs

- Comparison audit
  - Heavy demands on voting system for reporting and export
  - Requires LEO to commit to auditable subtotals
  - Requires ability to retrieve ballots that correspond to CVRs or subtotals
  - May compromise voter privacy (small-batch or ballot-level reporting) & enable coercion through pattern voting
  - Most efficient (ballot-level) may require re-scanning all ballots
  - Checks tabulation (but not for *transitive audits* [Calandrino, Halderman, & Felten] unless subtotals are cross-checked)
  - Ballot-level comparison audits require least hand counting
- Ballot polling audit
  - Requires more counting than *ballot-level* comparison audit
  - Does not check tabulation: Outcome could be right b/c errors cancel
  - Virtually no set-up costs
  - Requires nothing of voting system
  - Generally, need a ballot manifest to draw sample
  - Preserves voter anonymity except possibly for sampled ballots
  - Counting burden comparable to precinct-based comparison audit, unless margin is very small

Counting errors versus counting votes

Johnson (2004): statistical recount versus statistical error count. Like two-sample *t*-test versus paired *t*-test.

If constrained to examine batches of a given size, much more efficient statistically (in counting effort) to count errors in those batches than to count votes in those batches.

But if:

- you can only examine precinct-level batches for error
- exporting precinct-level data is hard/complex/time-consuming
- you can examine individual ballots to count votes

then counting votes can be much more efficient overall.

Getting CVRs for Individual Ballots is Hard!

- Federally certified voting systems do not provide CVRs.
- Even getting precinct-level data from today's voting systems into a usable form can take hours of hand editing . . . and then the batch size is too large for efficient audits.
- Generally need LEOs to re-scan ballots, need to program ballot definitions, etc.
  Serious obstacles to ballot-level comparison audits.
- Need ballot manifests for any kind of risk-limiting audit—comparison or ballot-polling.

Ballot-Polling Audit: Intuition

- Like opinion poll or exit poll, but sample until observed winner's percentage (i.e., sample percentage), discounted by "margin of error," is above 50% (for 2-candidate contest).
- If winner's *true* percentage of valid votes is more than 50%, she won.
- If the true margin is in fact small, confirming outcome might require looking at a lot of ballots; if it's big, don't expect to need to see many randomly selected ballots to have strong evidence that the winner got more than 50%.
- E.g., chance the first 4 ballots selected all would show votes for the reported winner if the reported winner didn't get more than 50% of the vote is 6.25% (less than 10%).
- If the true margin is in fact *negative* (i.e., if the reported winner really lost), very unlikely that sample percentage, discounted by "margin of error," will be over 50%.

Ballot-Polling Audit: Coin-Tossing

Want to check whether candidate 0 really beat candidate 1.

Draw a ballot at random from the ballots cast in the contest.

Condition on the event that the ballot has a valid vote for 0 or 1.

“Heads” is vote for 0; “tails” is vote for 1 (no votes for both, for now).

Like a coin toss: if 0 beat 1, $\pi_{heads} > 1/2$. If not, $\pi \leq 1/2$.

Fixed sample size or sequential?

- When can we stop inspecting ballots at random?
- Could decide to look at, say, 50 and conclude 0 won if she got at least 30 of the votes.
- What if true margin is huge? First 10 might be for 0.
- What if true margin is small? Might get fewer than 30 for 0.
- Want to draw until the *excess* of votes for 0 is strong evidence 0 won.
- How big an excess?

Wald's sequential probability ratio test (SPRT)

Two hypotheses, 0 and 1. Draw observations $X_1, X_2, \ldots$ sequentially.

$$\text{Likelihood ratio: } L_d \equiv \frac{f(x_1, \ldots x_d||1)}{f(x_1, \ldots, x_d||0)}.$$

SPRT:

0. $d \leftarrow 0$
1. Draw an observation; $d \leftarrow d + 1$
2. If $L_d < B$, reject hypothesis 1 and stop. If $L_d > A$, reject hypothesis 0 and stop. Else, go to step 1.

A and B control significance level and power.

Theorem (Wald, 1945)

$A \leq (1 - \beta)/\alpha \leq 1/\alpha$ and $B \geq \beta/(1 - \alpha)$.

Sketch proof of Wald's Theorem

Suppose $B < L_i < A$ for all $i < d$, but $L_d > A$.

By construction, chance of this is at least A times larger if 1 than if 0, for any d.

Chance under 0 should be $\leq \alpha$ and chance under 1 should be $\geq 1 - \beta$; hence $A \leq (1 - \beta)/\alpha$.

Similar proof for B.

Note that this means $1/L_d$ is a conservative P-value for hypothesis 0.

Wald for Bernoulli π

Test $\pi = 1/2$ against alternative $\pi = \pi_0 > 1/2$, independent trials.
$X_i = 1$ if ith ballot with valid vote for either 0 or 1 shows vote for 0.

$$L_d \equiv \frac{\pi_0^{\sum_{i=1}^d x_i}(1-\pi_0)^{d-\sum_{i=1}^d x_i}}{(1/2)^d}.$$

0. Pick α. $L \leftarrow 1$.
1. Draw a ballot at random.
2. If ballot shows vote for 0, $L \leftarrow 2\pi_0 L$.
   If ballot shows vote for 1, $L \leftarrow 2(1-\pi_0)L$.
3. If $L > 1/\alpha$, stop: strong evidence that 0 won.
   Else, go to step 1.

Random walk on a log scale; drift depends on true π.
Can always abort and perform a full hand count.
Modifications for sampling without replacement.

C-candidate, k-winner contest

Test that every winner $w \in \mathcal{W}$ beat every loser $\ell \in \mathcal{L}$.
$k(C-k)$ null hypotheses: loser ℓ beat winner w.

Test all w/ same sample, but one test statistic per pair: $\{L_{w\ell}\}$.

Define $s_{w\ell} \equiv s_w/(s_w + s_\ell)$, fraction of votes w was reported to have received among ballots reported to show a vote for w or ℓ or both.

Can be calculated from standard reported election results.

Define $\pi_{w\ell}$ to be actual fraction of votes w received among ballots that show a vote for *exactly one* of $\{w, \ell\}$.

Sufficient Condition

$\forall w \in \mathcal{W}, \ell \in \mathcal{L}$:

- If w reportedly beat ℓ, $s_{w\ell} > 50\%$.
- If w actually beat ℓ, $\pi_{w\ell} > 50\%$.

Wald for C-candidate, k-winner contest

1. Set $L_{w\ell} = 1$ for all $w \in \mathcal{W}$ and $\ell \in \mathcal{L}$.
2. Draw a ballot uniformly at random w/ replacement from those cast in contest.
3. If the ballot shows a valid vote for a reported winner w, then for each ℓ in $\mathcal{L}$ that did not receive a valid vote on that ballot, multiply $L_{w\ell}$ by $2s_{w\ell}$. Repeat for all such w.
4. If ballot shows a valid vote for a reported loser ℓ, then for each w in $\mathcal{W}$ that did not receive a valid vote on that ballot multiply $L_{w\ell}$ by $2(1 - s_{w\ell})$. Repeat for all such ℓ.
5. For all (w, ℓ) with $L_{w\ell} \geq 1/\alpha$, conclude that w beat ℓ.
   Don't update those $L_{w\ell}$ further.
6. If have concluded that all $w \in \mathcal{W}$ beat all $\ell \in \mathcal{L}$, stop: Reported results stand.

   Else, return to step 2.

Again, can abort at any time and perform full hand tally.

Theorem: Limits risk to at most α.

Multiplicity in pairwise testing for k-winner contest

Stopping short of a full hand count is an error only if at least one of the null hypotheses is true.

Procedure stops short of full hand count only if all $k(C-k)$ null hypotheses are rejected.

Consider the set of null hypotheses that are true. Chance of erroneously rejecting *all* of those is at most the smallest chance of erroneously rejecting any individually.

Hence, testing every (winner, loser) pair individually at level α makes chance of stopping short of a full hand count if any of the $C-k$ apparent losers actually won is at most α.

Moreover, works simultaneously for any number of contests, using the same sample.

Grouping losers

Could combine subsets of winners or of losers to reduce the number of tests.

E.g., winner has 60%, losers have 25% and 15%. Combine losers into a single fictitious losing candidate with 40%.

Theorem: grouping does not reduce expected sample size.

Steampunk audit

Equipment needed: dice, pencil and paper (or a sliderule).

Calculations very transparent.

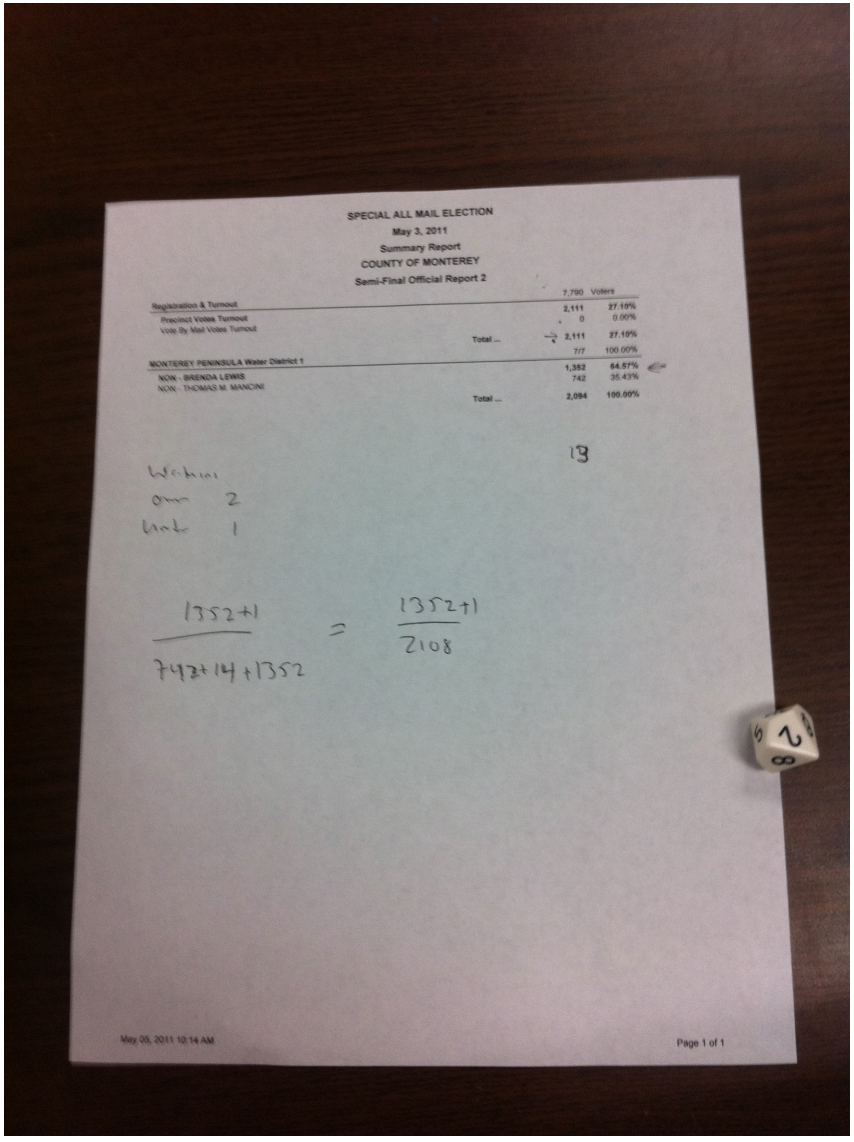
Process very observable: What votes does this ballot show?

Ballot-polling audit Monterey Peninsula Water District 1

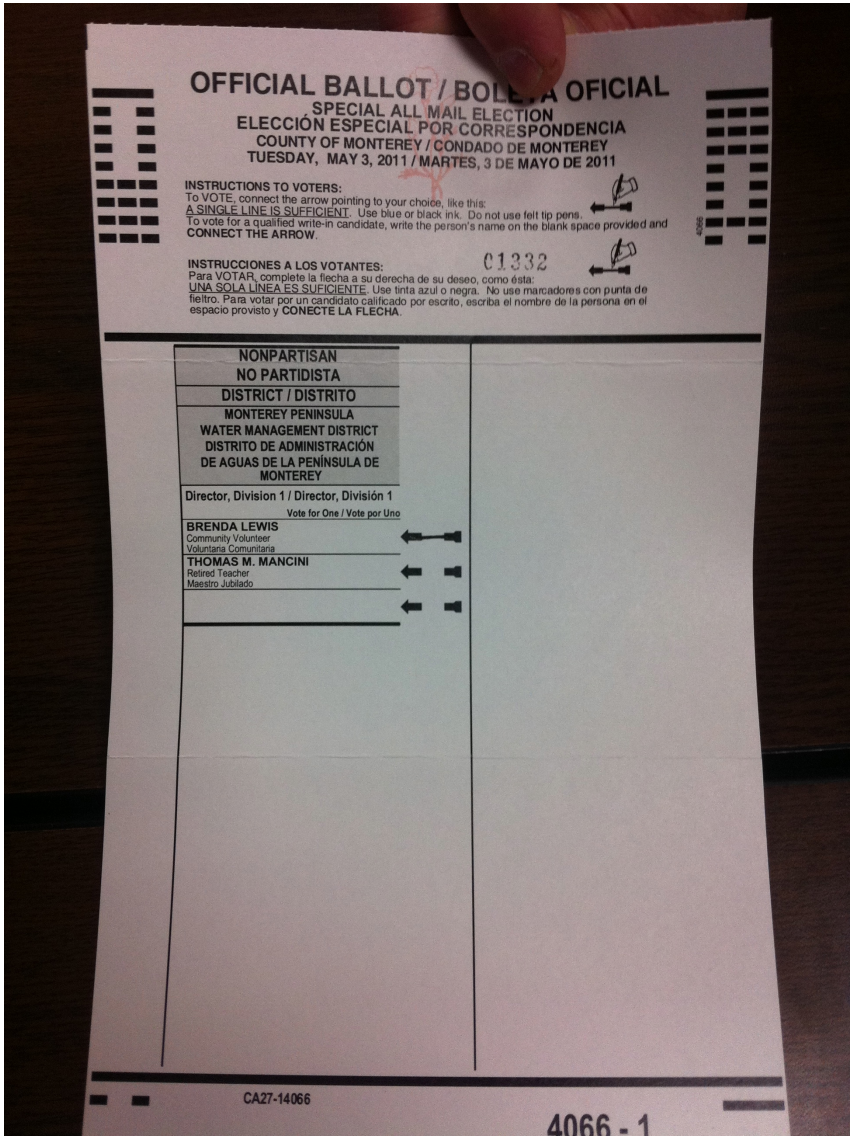
- Conducted in Monterey County in May, 2011, before certification
- 10% risk limit ($\alpha = 0.1$)
- Expected number of ballots to examine: 58
- Actual: 92 draws (89 distinct ballots)
- Monterey County staff Bates' stamped every ballot
- Thanks to RoV Linda Tulett & staff!

Monterey County 2011

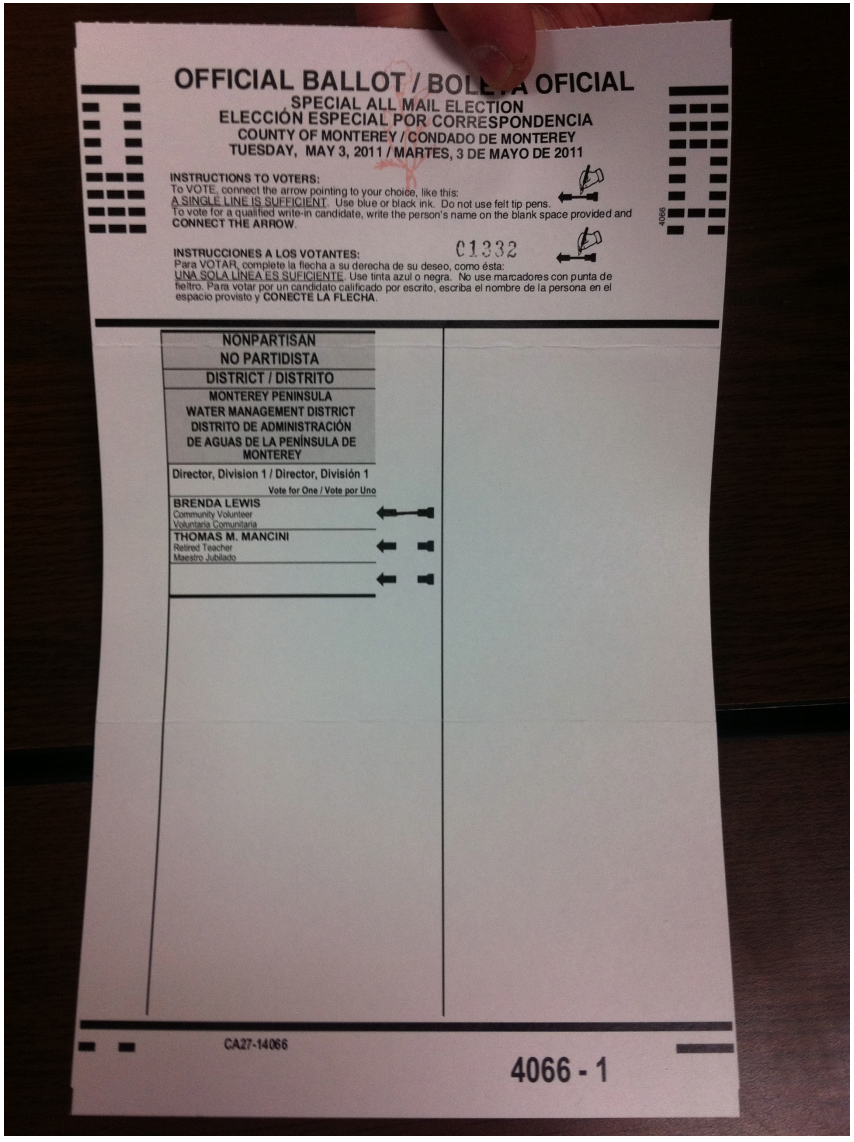
Monterey County 2011



Monterey County 2011



Monterey County 2011



Monterey County 2011

2008 Presidential Contest in CA

Expected sample size to confirm Obama won
Vote share 61.1%:

- $\approx$ 100 ballots from whole state
- $\approx$ 25 from LA County
- $\approx$ 75 total from largest 12 counties (including LA)
- $\approx$ 1 total from the smallest 14 counties.

If Obama's share had been 52%:

- $\approx$ 2,900 from whole state ($\approx$ 0.02% of ballots)
- $\approx$ 725 from LA county
- $\approx$ 2175 total from largest 12 counties (including LA)
- $\approx$ 29 total from smallest 14 counties

Expected Workload: Two Candidates

Winner's True Share	Quantiles					Mean
	25^{th}	50^{th}	75^{th}	90^{th}	99^{th}	
70%	12	22	38	60	131	30
65%	23	38	66	108	236	53
60%	49	84	149	244	538	119
58%	77	131	231	381	840	184
55%	193	332	587	974	2,157	469
54%	301	518	916	1,520	3,366	730
53%	531	914	1,619	2,700	5,980	1,294
52%	1,188	2,051	3637	6,053	13,455	2,900
51%	4,725	8,157	14,486	24,149	53,640	11,556
50.5%	18,839	32,547	57,838	96,411	214,491	46,126

Means and percentiles of #ballots with valid votes to inspect for 10% risk limit. Estimated using 10^7 replications.

Workload at 10% Risk Limit

255 state presidential contests between 1992 and 2008
median statewide expected sample size to confirm the plurality winner in each state is

307 ballots

(On the assumption that the outcomes were right.)

Selecting ballots at random

For transparency, want initial mechanical source of randomness (Cordero, Wagner, & Dill).

Dice courtesy of Ron Rivest.

Use as Seed in Good PRNG

SHA-256 of seed catenated with sample number (Rivest)

Random sampling

Pseudo-Random Sample of Ballots

Seed: 73567556725160627585

Number of ballots: 7116

Current sample number: 623

Draw this many ballots: 623 draw sample reset

Ballots selected: ☑ show sequence numbers ☐ show hash values

```
sequence_number, ballot
1,2086
2,2462
3,3320
4,4719
5,4813
6,3838
7,2655
8,2747
9,3059
```

Ballots selected, sorted:

```
19,34,37,38,51,90,96,96,99,101,109,114,150,156,163,175,187,187,195,197,198,244,280,281,301,316,
372,395,403,404,407,417,429,444,450,451,471,477,480,481,482,491,514,542,545,550,554,577,585,58
5,596,597,613,614,615,629,645,647,657,685,692,692,694,739,750,763,768,792,795,798,819,832,841,
842,857,862,871,874,876,884,901,906,923,923,934,937,937,958,963,973,978,1018,1049,1050,1071,1
081,1097,1105,1125,1126,1130,1165,1205,1210,1218,1219,1224,1226,1284,1288,1291,1318,1327,13
57,1370,1372,1388,1406,1422,1425,1432,1433,1434,1446,1447,1457,1484,1494,1496,1507,1512,152
3,1524,1540,1572,1574,1575,1576,1611,1614,1626,1634,1638,1642,1644,1665,1677,1685,1718,1735,
1761,1764,1774,1788,1791,1793,1816,1827,1851,1855,1893,1921,1978,1989,2010,2017,2034,2056,2
056,2058,2062,2069,2083,2086,2100,2112,2152,2189,2192,2206,2208,2210,2213,2224,2249,2266,22
91,2295,2302,2331,2332,2390,2391,2395,2398,2401,2422,2436,2462,2463,2474,2495,2513,2514,252
```

Ballots selected, sorted, duplicates removed:

```
19,34,37,38,51,90,96,99,101,109,114,150,156,163,175,187,195,197,198,244,280,281,301,316,372,395
,403,404,407,417,429,444,450,451,471,477,480,481,482,491,514,542,545,550,554,577,585,596,597,6
13,614,615,629,645,647,657,685,692,694,739,750,763,768,792,795,798,819,832,841,842,857,862,871
,874,876,884,901,906,923,934,937,958,963,973,978,1018,1049,1050,1071,1081,1097,1105,1125,1126
,1130,1165,1205,1210,1218,1219,1224,1226,1284,1288,1291,1318,1327,1357,1370,1372,1388,1406,1
422,1425,1432,1433,1434,1446,1447,1457,1484,1494,1496,1507,1512,1523,1524,1540,1572,1574,15
75,1576,1611,1614,1626,1634,1638,1642,1644,1665,1677,1685,1718,1735,1761,1764,1774,1788,179
1,1793,1816,1827,1851,1855,1893,1921,1978,1989,2010,2017,2034,2056,2058,2062,2069,2083,2086,
2100,2112,2152,2189,2192,2206,2208,2210,2213,2224,2249,2266,2291,2295,2302,2331,2332,2390,2
391,2395,2398,2401,2422,2436,2462,2463,2474,2495,2513,2514,2520,2549,2556,2558,2563,2578,25
```

Ballot Manifest

Find ballots using a ballot manifest

Ballot look-up tool

Ballot manifest: Each line must have a batch label, a comma, and one of the following:
(i) the number of ballots in the batch
(ii) a range specified with a colon (e.g., 131:302), or
(iii) a list of ballot identifiers within parentheses, separated by spaces (e.g., (996 998 1000)).
Each line should have exactly one comma.

```
001_211161_01,23
002_211162_02,9
003_211561_03,32
004_211561_03,50
005_211561_03,50
006_211562_04,14
007_211562_04,50
008_211562_04,50
009_211562_04,50
010_211563_05,12
011_211751_06,27
012_211761_07,2
013_211761_07,50
014_211761_07,50
015_211761_07,50
016_211761_07,50
017_211771_08,2
018_221161_09,16
019_221161_09,50
020_221161_09,50
021_221161_09,50
022_221162_10,30
023_221162_10,50
024_221162_10,50
025_221162_11,50
```

Ballots to look up (separated by commas):

```
19,34,37,38,51,90,96,96,99,101,109,114,150,156,163,175,187,187,195,197,198,244,280,281,301,316,
372,395,403,404,407,417,429,444,450,451,471,477,480,481,482,491,514,542,545,550,554,577,585,58
5,596,597,613,614,615,629,645,647,657,685,692,692,694,739,750,763,768,792,795,798,819,832,841,
842,857,862,871,874,876,884,901,906,923,923,934,937,937,958,963,973,978,1018,1049,1050,1071,1
081,1097,1105,1125,1126,1130,1165,1205,1210,1218,1219,1224,1226,1284,1288,1291,1318,1327,13
57,1370,1372,1388,1406,1422,1425,1432,1433,1434,1446,1447,1457,1484,1494,1496,1507,1512,152
3,1524,1540,1572,1574,1575,1576,1611,1614,1626,1634,1638,1642,1644,1665,1677,1685,1718,1735,
1761,1764,1774,1788,1791,1793,1816,1827,1851,1855,1893,1921,1978,1989,2010,2017,2034,2056,2
056,2058,2062,2069,2083,2086,2100,2112,2152,2189,2192,2206,2208,2210,2213,2224,2249,2266,22
91,2295,2302,2331,2332,2390,2391,2395,2398,2401,2422,2436,2462,2463,2474,2495,2513,2514,252
```

Look-up

look up ballots

Sorted lookup table:

sorted_number, ballot, batch_label, which_ballot_in_batch
1, 19, 001_211161_01, 19
2, 34, 003_211561_03, 2
3, 37, 003_211561_03, 5
4, 38, 003_211561_03, 6
5, 51, 003_211561_03, 19
6, 90, 004_211561_03, 26
7, 96, 004_211561_03, 32
8, 96, 004_211561_03, 32
9, 99, 004_211561_03, 35
10, 101, 004_211561_03, 37
11, 109, 004_211561_03, 45
12, 114, 004_211561_03, 50
13, 150, 005_211561_03, 36
14, 156, 005_211561_03, 42
15, 163, 005_211561_03, 49
16, 175, 006_211562_04, 11
17, 187, 007_211562_04, 9
18, 187, 007_211562_04, 9
19, 195, 007_211562_04, 17
20, 197, 007_211562_04, 19
21, 198, 007_211562_04, 20
22, 244, 008_211562_04, 16
23, 280, 009_211562_04, 2
24, 281, 009_211562_04, 3

Gotchya!

Better ballot accounting

Ballot manifests are *not* a solved problem.

It's easy to deal with errors in ballot manifest if there's an upper bound on the number of ballots in each container (Bañuelos & Stark).

But sometimes there isn't a good upper bound—esp. with multipage ballots.

What if margin is small, or if contests don't overlap?

Ballot-polling has modest workload until margins get small.

Then, comparison audits have an advantage that can make up for their higher up-front costs.

How can we test the hypothesis that one or more reported outcomes are wrong?

Comparison audits: MACRO

Error: Hand-count disagrees with reported count; hand-count presumed correct.

Overstatement: correcting the error would narrow at least one margin. Increase the required sample—decrease confidence.

Understatement: correcting the error would widen every margin. Decrease required sample—increase confidence—but by less.

More confidence if sample shows no misstatements than if understatements balance overstatements.

Sufficient condition for all outcomes to be right:

For every (winner, loser) pair, net overstatement of the margin between them is less than 100% of the reported margin between them.

MACRO

For $w \in \mathcal{W}_\chi$, $\ell \in \mathcal{L}_\chi$, define

$$e_{pw\ell} \equiv \begin{cases} \frac{(v_{wp} - v_{\ell p}) - (a_{wp} - a_{\ell p})}{V_{w\ell}}, & \text{if batch } p \text{ contains contest } \chi \\ 0, & \text{otherwise.} \end{cases}$$

If any apparent outcome is wrong,

$$\exists \chi \in \{1, \ldots, X\} \text{ s.t. } \exists (w \in \mathcal{W}_\chi,\ \ell \in \mathcal{L}_\chi) \text{ with } \sum_{p=1}^{N} e_{pw\ell} \geq 1. \tag{1}$$

Test based on sufficient condition

$$e_p \equiv \max_{\chi} \max_{w \in \mathcal{W}_\chi,\, \ell \in \mathcal{L}_\chi} e_{pw\ell}.$$

Bound: (sum of max) $\geq$ (max of sum).

Simple sufficient condition: All outcomes must be correct if

$$E \equiv \sum_{p=1}^{N} e_p < 1.$$

Maximum across-contest relative overstatement of pairwise margins (MACRO)

Controlling the familywise error rate

M null hypotheses,

$$\{\text{the outcome of contest } \chi \text{ is incorrect}\}_{\chi=1}^{M}.$$

If $E < 1$, the entire family of M null hypotheses is false:
All apparent outcomes are right.

Test of hypothesis $E \geq 1$ at significance level α is a test of the M hypotheses with familywise error rate no larger than α.

Bounding the overstatement error in each batch

A priori bounds are crucial.

If number of valid ballots cast in batch p for contest χ is at most $b_{\chi p}$ then

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

Hence,

$$e_p \leq \max_{\chi \in \{1,\ldots,X\}} \max_{w \in \mathcal{W}_\chi, \ell \in \mathcal{L}_\chi} \frac{v_{wp} - v_{\ell p} + b_{\chi p}}{V_{w\ell}} \equiv u_p.$$

$U \equiv \sum_p u_p$, upper bound on total MACRO.

Sampling Designs

- Most jurisdictions that have audits use stratified cluster sampling.
- For most certified systems, limited to some kind of cluster sample (c.f., Alameda, Humboldt, Merced, Monterey, Napa, Orange, San Luis Obispo, Stanislaus, Yolo, audits).
- 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, Feige?

Taint & PPEB Sampling

taint of batch p

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

Independent draws. In each draw,

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

PPS, used in financial auditing.

Taint of ith draw is T_i. $\{T_i\}$ are iid. $\mathbf{E}T_i = E/U$.

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

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

Wald, again

Let F be the cdf of a nonnegative variable with mean μ.

Then $\int_0^\infty (x/\mu) dF = 1$; i.e., $(x/\mu)F$ is a cdf.

Imagine testing the hypothesis that observations $X_1, X_2, \ldots$ come from F versus coming from $(x/\mu)F$.

The likelihood ratio after d observations is

$$L_d = \prod_{i=1}^{d} (\mu/x_i).$$

Hence, by Wald, can reject if $\prod_{i=1}^{d} (\mu/x_i) > 1/\alpha$.

Define $X_i = 1 - T_i$. Then X_d is nonnegative; relevant μ is $1 - 1/U$.

Likelihood ratio is $L_d = \prod_{i=1}^{d} (1 - 1/U)/(1 - T_i)$; P-value $1/L_d$.

Refinements possible.

Sequential risk-limiting audit using Kaplan-Markov-Wald bound

1. Calculate bounds $\{u_p\}$, U. Set $d = 1$. Pick $\alpha \in (0, 1)$ and $D > 0$.

1. Draw a batch using PPEB. Audit batch if it has not already been audited.

3. Find $T_d \equiv t_p \equiv e_p/u_p$, taint of the batch p drawn at stage d.

4. Compute

$$P_d \equiv \prod_{i=1}^{d} \frac{1 - 1/U}{1 - T_i}. \quad \text{See November 2010 WIRED, p.56} \quad (2)$$

5. If $P_d < \alpha$, report apparent outcomes and stop. If $d = D$, audit remaining batches, report then-known outcomes and stop.. Else, $d \leftarrow d + 1$ and go to 2.

This sequential procedure is risk-limiting

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

If any outcome is wrong,

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

Remarkably efficient if batches are not too big.

Super-simple method: ballot-level comparison audit

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., OpenCount, TEVS) *a la* Calendrino et al. (2007)

Advantages of super-simple method

- Audit entire collection of contests with one simple random sample of ballots.
- Very simple calculation determines when to stop.
- Chance of correcting all wrong outcomes is guaranteed to be at least as high as claimed.
- Transparent, easy to observe.
- Only have to count to 1 (for plurality contests): does ballot have vote for a candidate, or not?
  (A ballot can agree with CVR or have overstatement or understatement of 1 or 2 votes.)

Super-simple++

Special case of the previous method: uses upper bound on overstatement in each ballot, algebraic inequalities.
Results in drawing ballots with equal probability, with replacement.

Pick risk limit α and 2 parameters:

- limit D on draws before performing full hand count.
- $\gamma \geq 100\%$. Controls tradeoff between pain of 1-vote overstatements and 2-vote overstatements.

m is "diluted" margin: margin in votes divided by ballots, not by valid votes.
o_1, o_2, u_1, u_2 are numbers of 1 and 2-vote overstatements and understatements in the sample.

Super-simple++ audit

$\alpha = 0.1$, $\gamma = 1.03905$.

1. Pick D, maximum draws before full hand count. s is winner's share of the valid votes according to the vote tabulation system. Set $T = 1$, $d = 0$.
2. Select a ballot at random, uniformly, from ballots cast in the contest. $d \leftarrow d + 1$.
3. Compare ballot to CVR; note whether correct, understatement, overstatement
4. If $d \geq \frac{4.8+1.4(o_1+5o_2-0.6u_1-4.4u_2)}{m}$, stop audit: reported results stand
   Else if $d < D$, return to step 2.
5. Perform full hand count; hand-count results trump reported results.

Theorem: limits risk to α. For this "tuning," 1-vote understatement offsets 60% of 1-vote overstatement and 2-vote understatement offsets 85% of 2-vote overstatement.

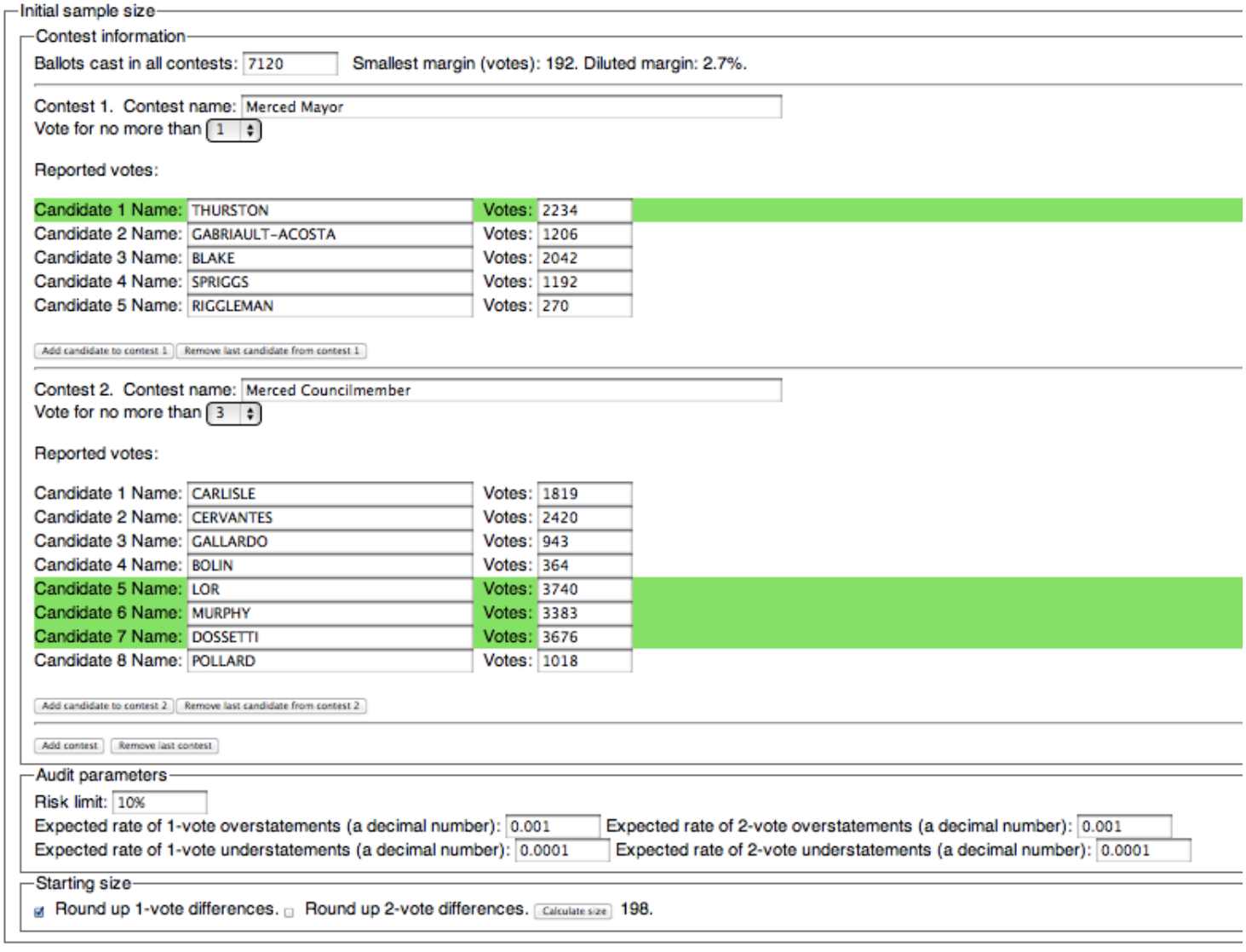
auditTools.htm

Need simple, friendly tools for auditing, e.g.:

`statistics.berkeley.edu/~stark/Vote/auditTools.htm`

Used for audits in Alameda, Humboldt, Merced, Napa, Stanislaus, Ventura.

auditTools in action



Initial sample size

Contest information

Ballots cast in all contests: 7120 Smallest margin (votes): 192. Diluted margin: 2.7%.

Contest 1. Contest name: Merced Mayor

Vote for no more than 1

Reported votes:

Candidate 1 Name:	THURSTON	Votes:	2234
Candidate 2 Name:	GABRIAULT-ACOSTA	Votes:	1206
Candidate 3 Name:	BLAKE	Votes:	2042
Candidate 4 Name:	SPRIGGS	Votes:	1192
Candidate 5 Name:	RIGGLEMAN	Votes:	270

Add candidate to contest 1 | Remove last candidate from contest 1

Contest 2. Contest name: Merced Councilmember

Vote for no more than 3

Reported votes:

Candidate 1 Name:	CARLISLE	Votes:	1819
Candidate 2 Name:	CERVANTES	Votes:	2420
Candidate 3 Name:	GALLARDO	Votes:	943
Candidate 4 Name:	BOLIN	Votes:	364
Candidate 5 Name:	LOR	Votes:	3740
Candidate 6 Name:	MURPHY	Votes:	3383
Candidate 7 Name:	DOSSETTI	Votes:	3676
Candidate 8 Name:	POLLARD	Votes:	1018

Add candidate to contest 2 | Remove last candidate from contest 2

Add contest | Remove last contest

Audit parameters

Risk limit: 10%

Expected rate of 1-vote overstatements (a decimal number): 0.001 Expected rate of 2-vote overstatements (a decimal number): 0.001

Expected rate of 1-vote understatements (a decimal number): 0.0001 Expected rate of 2-vote understatements (a decimal number): 0.0001

Starting size

Round up 1-vote differences. Round up 2-vote differences. Calculate size 198.

Random sampling

Pseudo-Random Sample of Ballots

Seed: 12082217
Number of ballots: 7120
Current sample number: 198
Draw this many ballots: 198 draw sample reset

Ballots selected: ☑ show sequence numbers ☐ show hash values

```
sequence_number, ballot
1,2660
2,5463
3,5334
4,2208
5,3459
6,6223
7,2407
8,5245
9,1899
```

Ballots selected, sorted:

```
35,82,98,99,197,220,241,254,256,369,389,416,422,447,501,573,638,738,760,831,932,940,964,986,1006,1
027,1067,1197,1208,1234,1285,1298,1410,1446,1464,1476,1495,1509,1548,1568,1621,1647,1745,1778,1
877,1879,1899,1947,1973,2023,2061,2133,2173,2208,2241,2318,2339,2398,2400,2407,2514,2557,2654,2
660,2666,2725,2744,2760,2847,2866,2894,3119,3123,3197,3223,3227,3232,3233,3299,3366,3370,3405,3
444,3459,3585,3588,3598,3624,3629,3637,3718,3758,3774,3802,3839,3875,3906,3977,4168,4177,4223,4
243,4261,4286,4321,4357,4382,4410,4426,4427,4429,4449,4517,4528,4536,4542,4571,4668,4712,4715,4
748,4749,4755,4779,4803,4805,4812,4814,4817,4828,4899,4922,4976,4988,5073,5116,5119,5138,5194,5
210,5240,5245,5305,5334,5414,5429,5463,5523,5534,5554,5658,5681,5691,5730,5740,5787,5854,5878,5
904,5980,5998,6001,6029,6032,6032,6043,6052,6078,6113,6161,6166,6223,6233,6258,6291,6379,6421,6
428,6446,6518,6549,6567,6599,6607,6628,6644,6697,6716,6784,6818,6853,6877,6908,6972,7001,7017,7
```

Ballots selected, sorted, duplicates removed:

```
35,82,98,99,197,220,241,254,256,369,389,416,422,447,501,573,638,738,760,831,932,940,964,986,1006,1
027,1067,1197,1208,1234,1285,1298,1410,1446,1464,1476,1495,1509,1548,1568,1621,1647,1745,1778,1
877,1879,1899,1947,1973,2023,2061,2133,2173,2208,2241,2318,2339,2398,2400,2407,2514,2557,2654,2
660,2666,2725,2744,2760,2847,2866,2894,3119,3123,3197,3223,3227,3232,3233,3299,3366,3370,3405,3
444,3459,3585,3588,3598,3624,3629,3637,3718,3758,3774,3802,3839,3875,3906,3977,4168,4177,4223,4
243,4261,4286,4321,4357,4382,4410,4426,4427,4429,4449,4517,4528,4536,4542,4571,4668,4712,4715,4
748,4749,4755,4779,4803,4805,4812,4814,4817,4828,4899,4922,4976,4988,5073,5116,5119,5138,5194,5
210,5240,5245,5305,5334,5414,5429,5463,5523,5534,5554,5658,5681,5691,5730,5740,5787,5854,5878,5
904,5980,5998,6001,6029,6032,6043,6052,6078,6113,6161,6166,6223,6233,6258,6291,6379,6421,6428,6
446,6518,6549,6567,6599,6607,6628,6644,6697,6716,6784,6818,6853,6877,6908,6972,7001,7017,7024,7
```

Repeated ballots:

```
Ballot, multiplicity
6032,2
```

Finding ballots using a ballot manifest

Ballot look-up tool

Ballot manifest: (batch label, ballots) pairs separated by commas, one pair per line

```
Merced1-cvr.txt,162
Merced11-cvr.txt,284
Merced13&15-cvr.txt,423
Merced14-cvr.txt,163
Merced16-cvr.txt,257
Merced17-cvr.txt,172
Merced18-cvr.txt,237
Merced2-cvr.txt,249
Merced20&21-cvr.txt,756
Merced22&29-cvr.txt,415
Merced23&26-cvr.txt,465
Merced24&25-cvr.txt,504
Merced27&32-cvr.txt,534
Merced28-cvr.txt,484
Merced3&30-cvr.txt,257
Merced31&33-cvr.txt,312
Merced4&12-cvr.txt,394
Merced5&9&10-cvr.txt,357
Merced6&19-cvr.txt,326
Merced7&8-cvr.txt,369
```

Ballots to look up (separated by commas):

```
35,82,98,99,197,220,241,254,256,369,389,416,422,447,501,573,638,738,760,831,932,94
0,964,986,1006,1027,1067,1197,1208,1234,1285,1298,1410,1446,1464,1476,1495,1509,
1548,1568,1621,1647,1745,1778,1877,1879,1899,1947,1973,2023,2061,2133,2173,2208,
2241,2318,2339,2398,2400,2407,2514,2557,2654,2660,2666,2725,2744,2760,2847,2866,
2894,3119,3123,3197,3223,3227,3232,3233,3299,3366,3370,3405,3444,3459,3585,3588,
3598,3624,3629,3637,3718,3758,3774,3802,3839,3875,3906,3977,4168,4177,4223,4243,
4261,4286,4321,4357,4382,4410,4426,4427,4429,4449,4517,4528,4536,4542,4571,4668,
4712,4715,4748,4749,4755,4779,4803,4805,4812,4814,4817,4828,4899,4922,4976,4988,
5073,5116,5119,5138,5194,5210,5240,5245,5305,5334,5414,5429,5463,5523,5534,5554,
5658,5681,5691,5730,5740,5787,5854,5878,5904,5980,5998,6001,6029,6032,6032,6043
```

look up ballots

Sorted lookup table:

```
sorted_number, ballot, batch_label, which_ballot_in_batch
1, 35, Merced1-cvr.txt, 35
2, 82, Merced1-cvr.txt, 82
3, 98, Merced1-cvr.txt, 98
4, 99, Merced1-cvr.txt, 99
5, 197, Merced11-cvr.txt, 35
6, 220, Merced11-cvr.txt, 58
7, 241, Merced11-cvr.txt, 79
8, 254, Merced11-cvr.txt, 92
9, 256, Merced11-cvr.txt, 94
```

Should more ballots be audited?

Stopping sample size and escalation

Ballots audited so far: 198

1-vote overstatements:	0	Rate: 0
2-vote overstatements:	0	Rate: 0
1-vote understatements:	0	Rate: 0
2-vote understatements:	0	Rate: 0

Estimated stopping size

Calculate Audit complete

If no more differences are observed: 178.

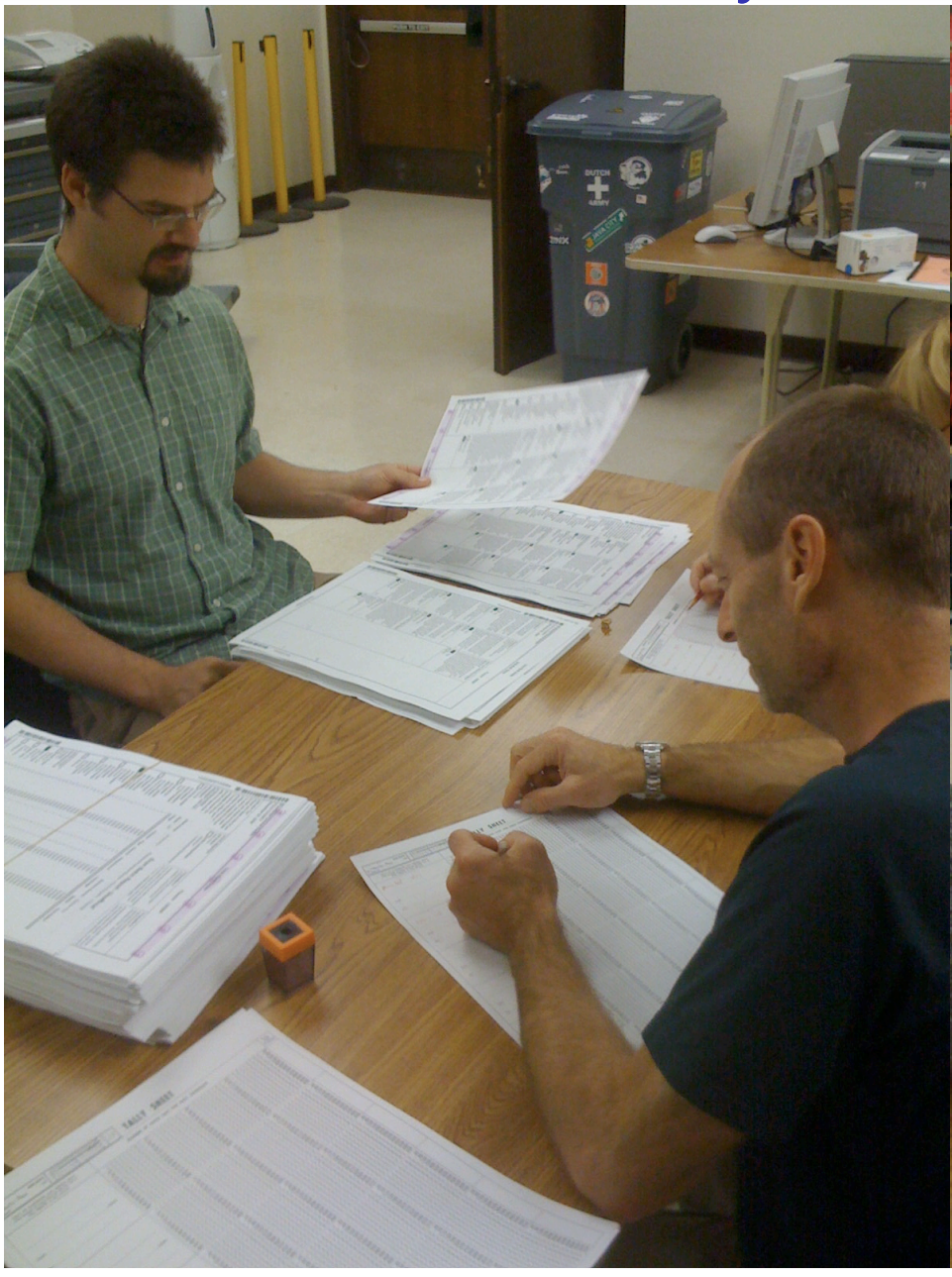
If differences continue at the same rates: 178.

Estimated additional ballots if difference rates stay the same: 0.

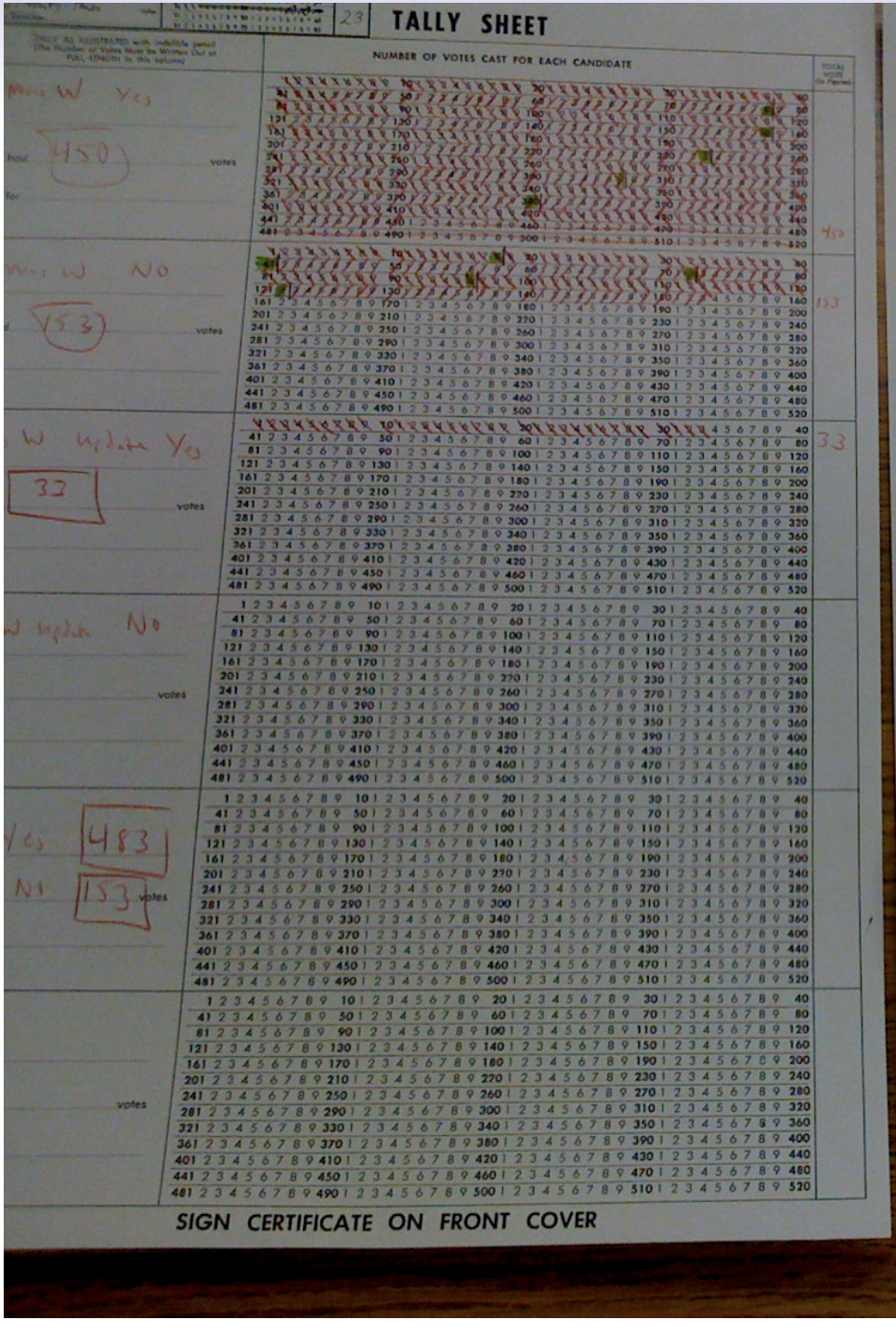
Secret sauce

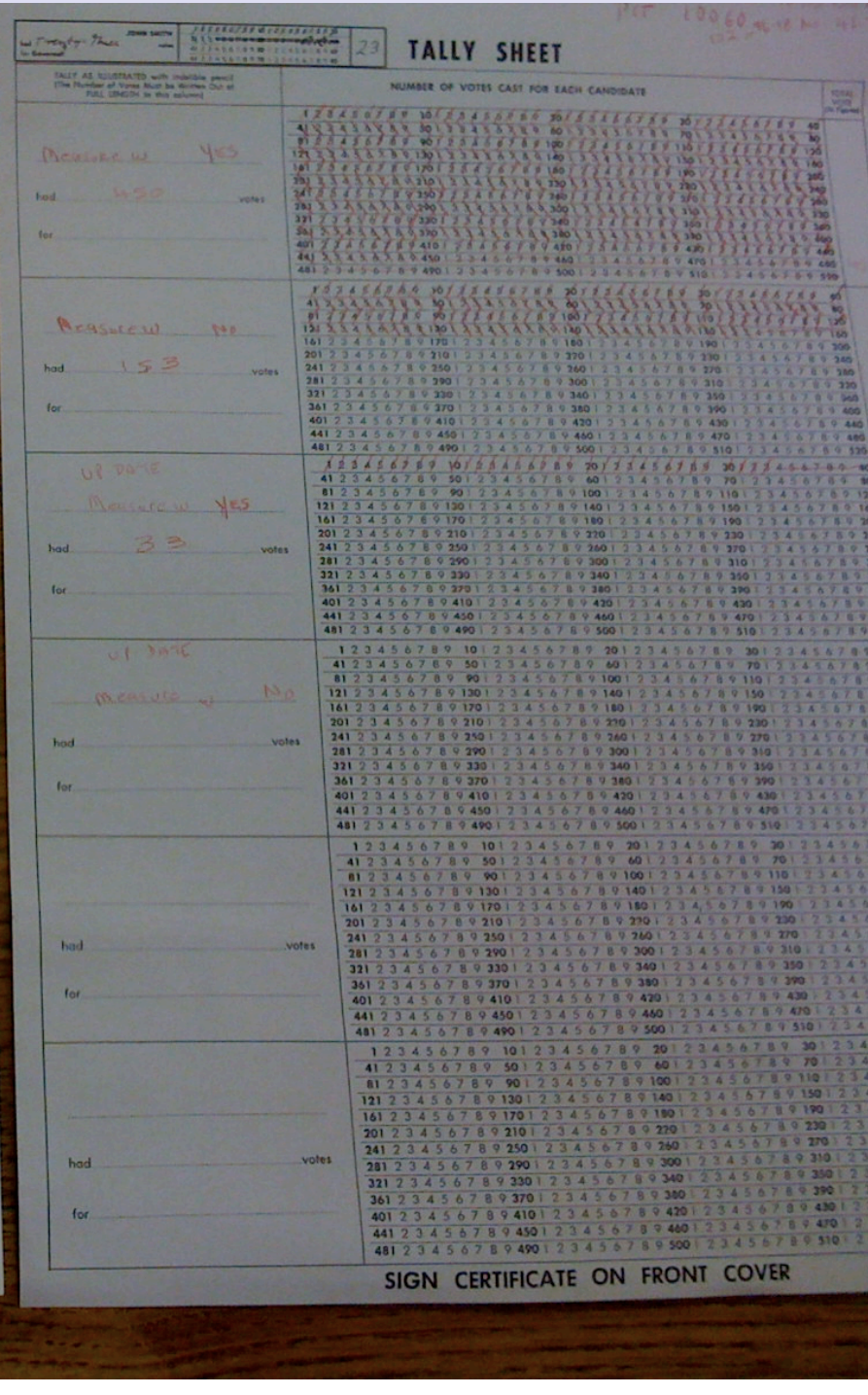
- To implement ballot-level comparison audits, have to associate individual cast vote records (CVRs) with individual physical ballots.
  Impossible with current U.S. federally certified systems.
- “Transitive” auditing using an unofficial vote tabulation system that does produce CVRs—such as those of OpenCount or TEVS—and confirming transitively that the apparent outcome is correct, might be the best interim option. (See Calendrino et al. 2007)
  If official system says “Lincoln won” and unofficial system says “Lincoln won,” then if unofficial system is right, so is official system.
- Performed transitive audits in Alameda, Merced, Stanislaus, Ventura.
  Napa, Orange, Yolo upcoming.

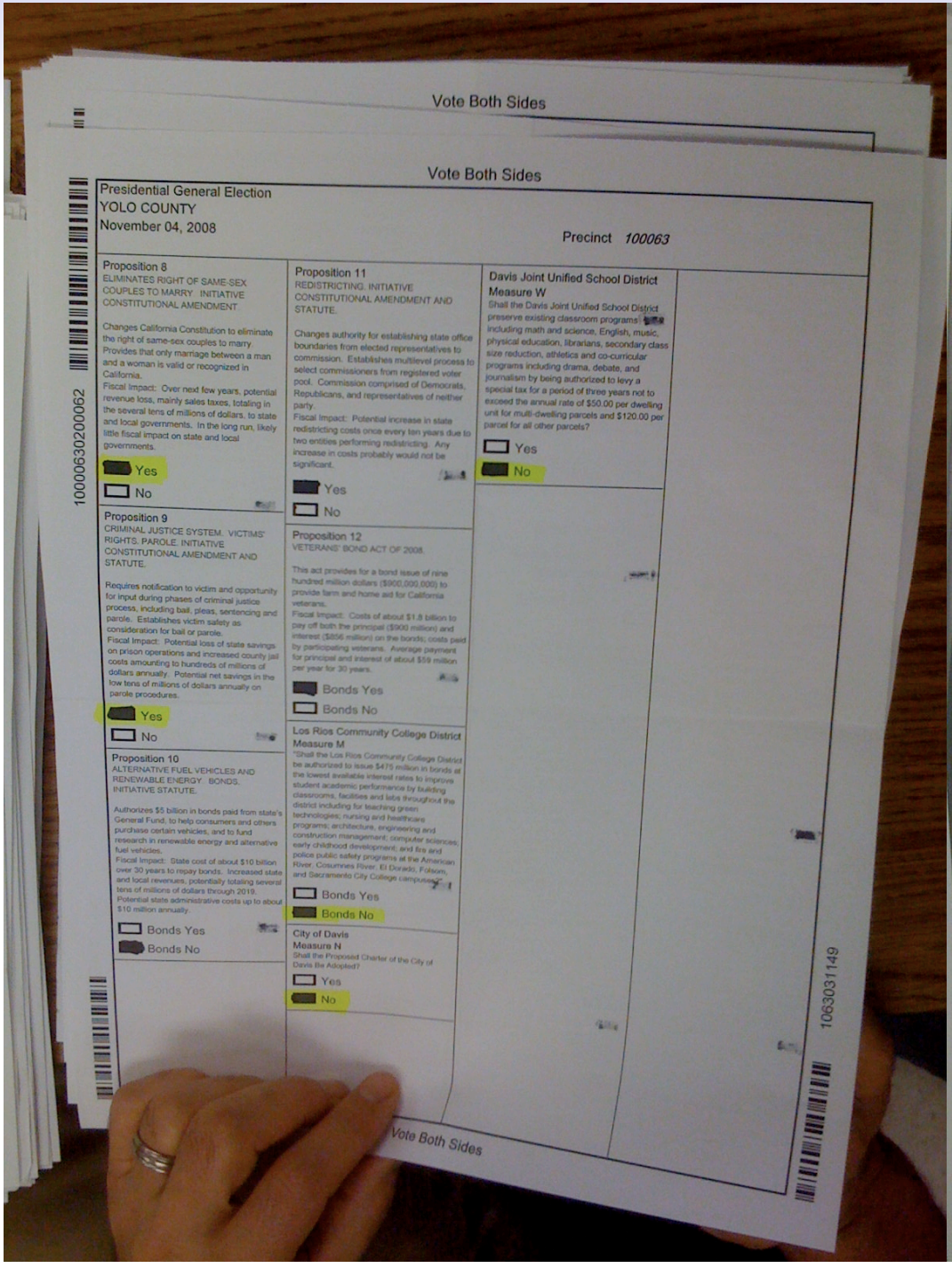
2008 Yolo County, CA Measure W Audit

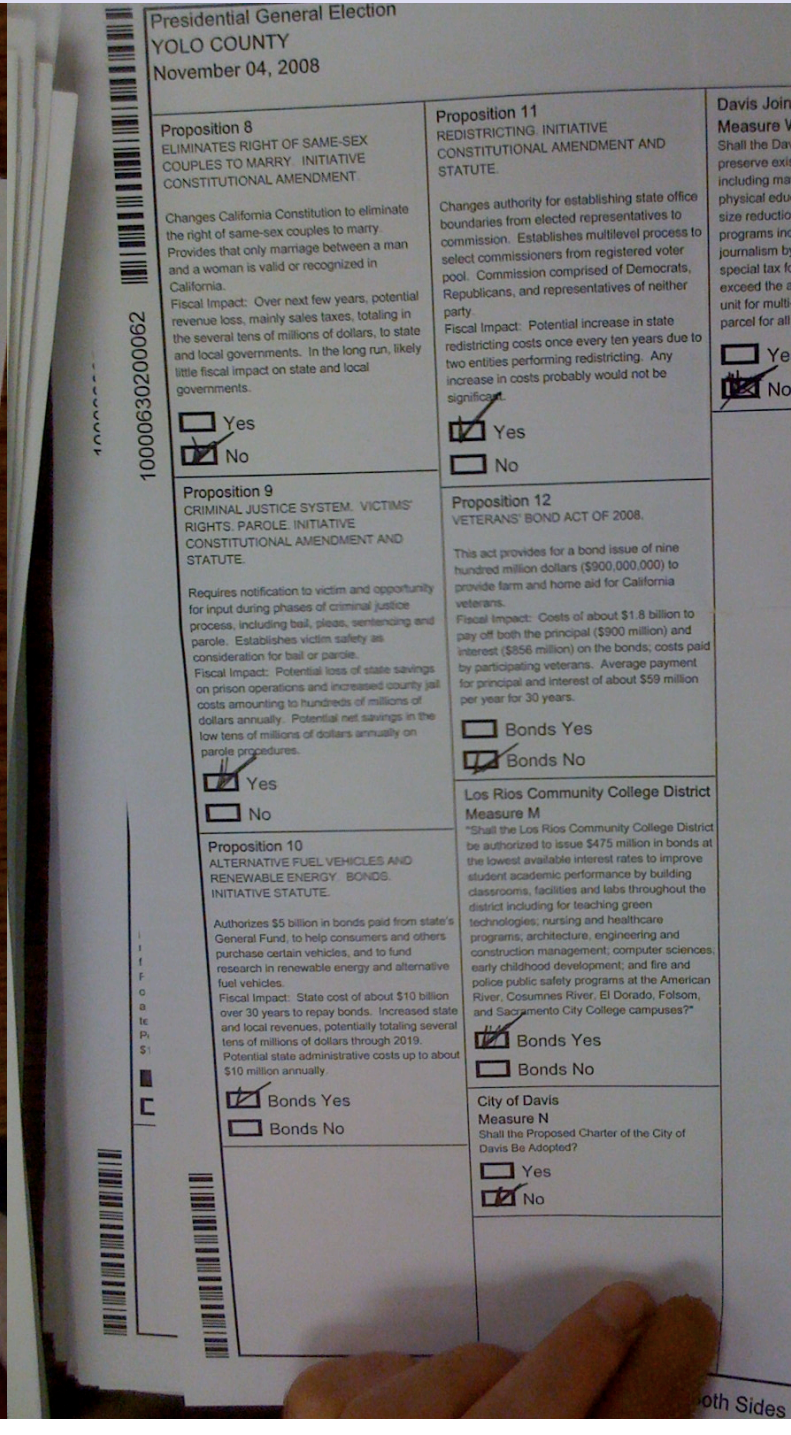


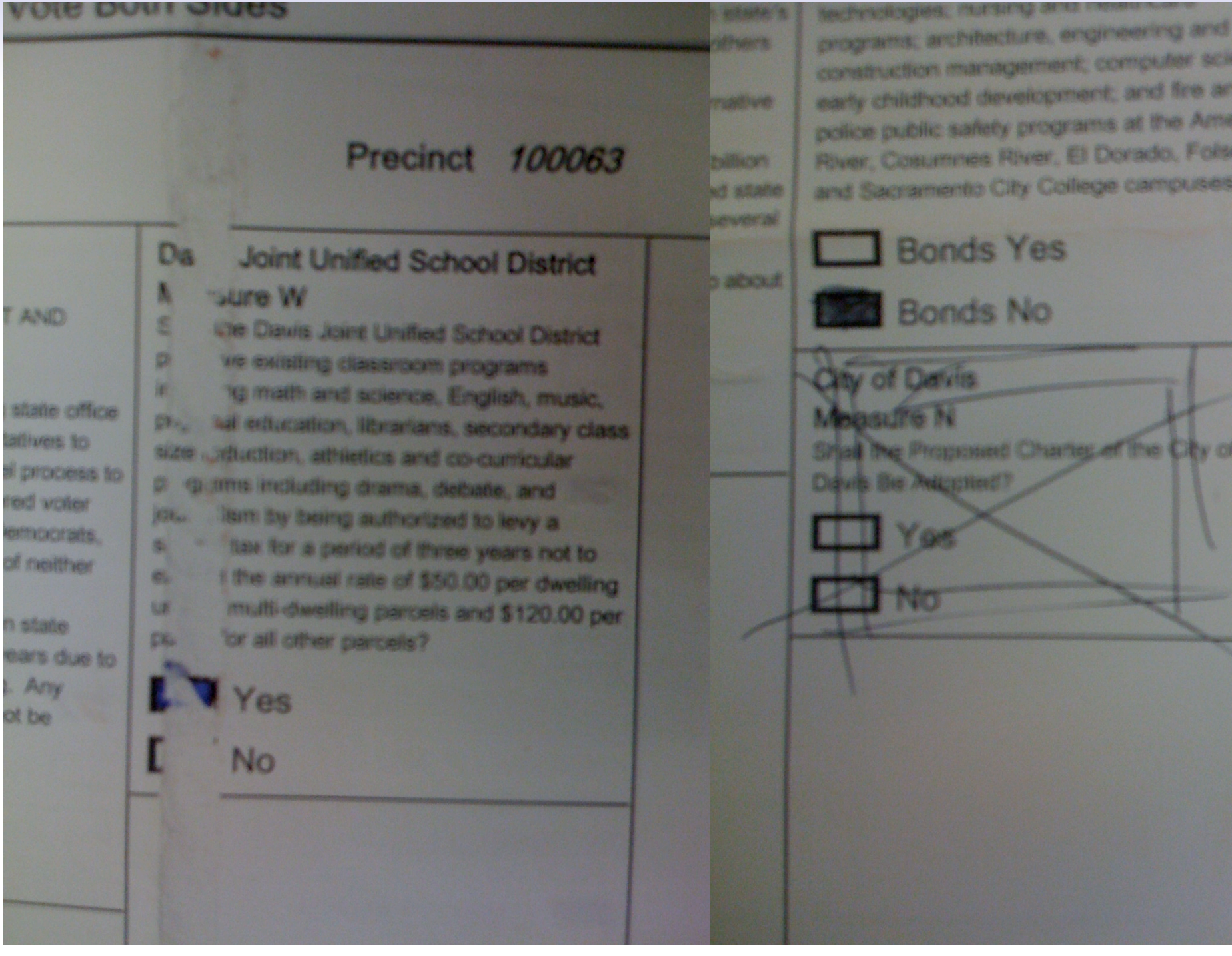

Hefty
Hefty


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

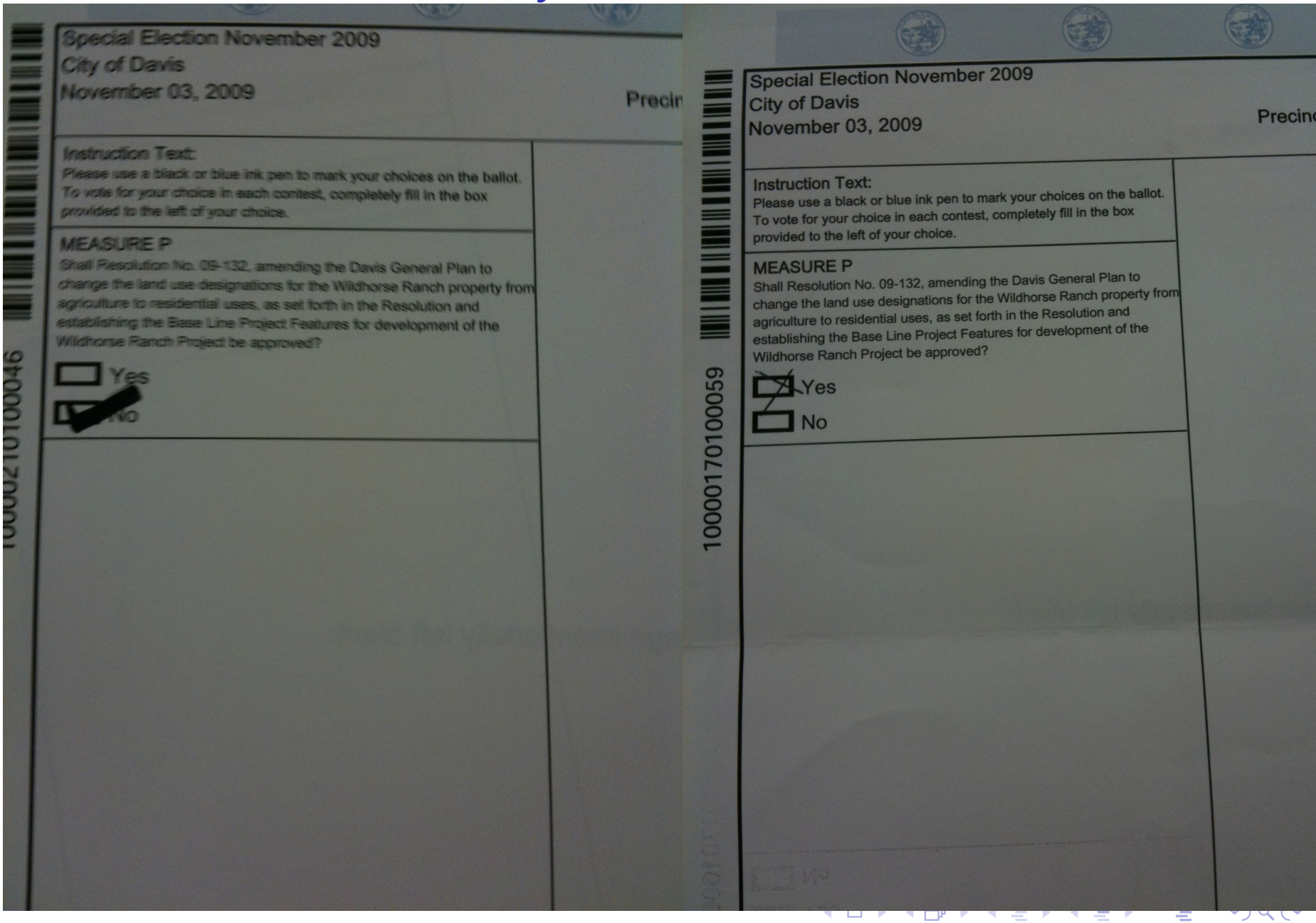

TALLY SHEET
NUMBER OF VOTES CAST FOR EACH CANDIDATE
Measure W Yes
had 450 votes
for
Measure W No
had 153 votes
for
Update
Measure W Yes
had 33 votes
for
Update
Measure W No
had votes
for
had votes
for
had votes
for
SIGN CERTIFICATE ON FRONT COVER

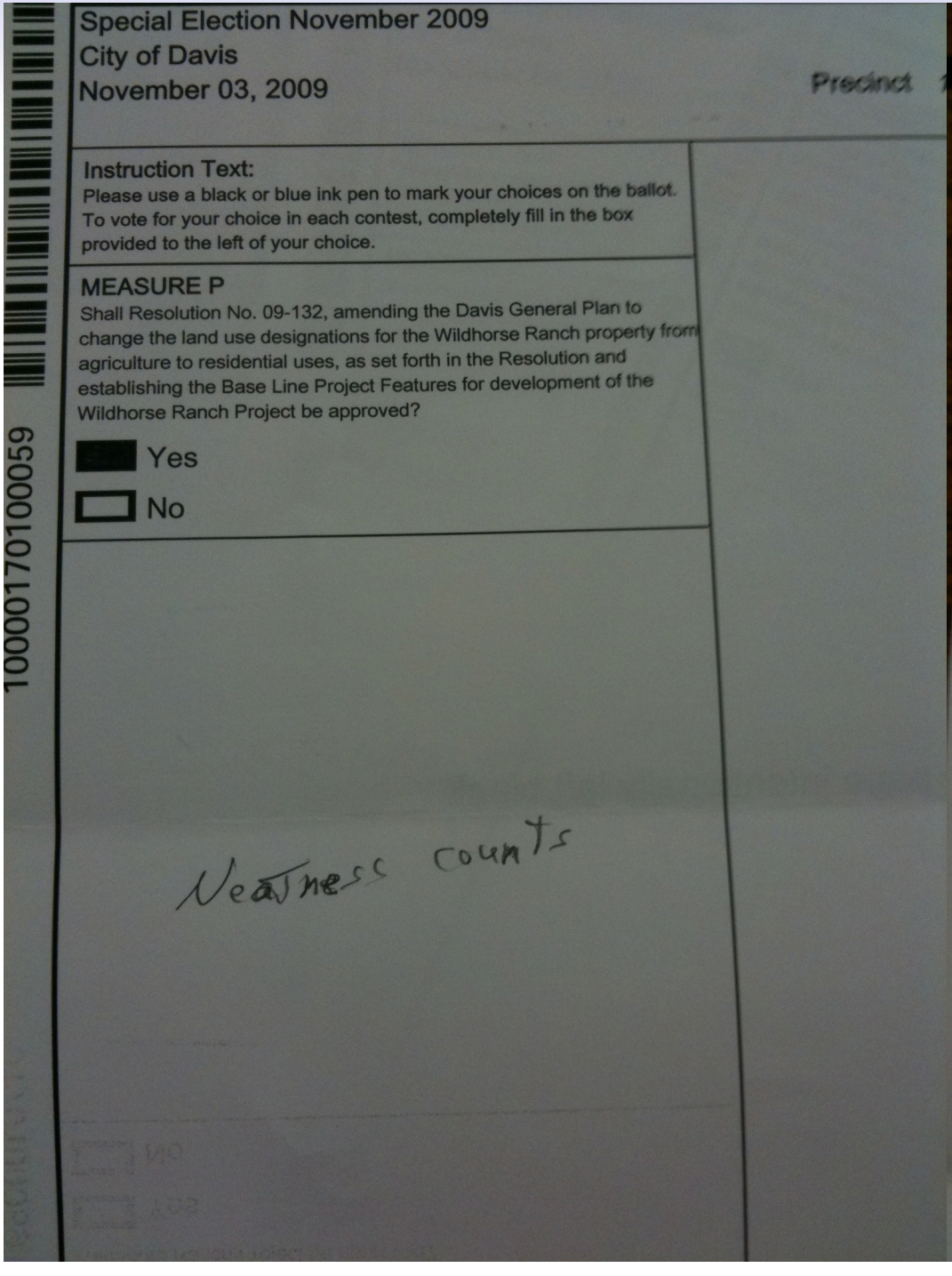

Vote Both Sides
Vote Both Sides
Presidential General Election
YOLO COUNTY
November 04, 2008
Precinct 100063
Proposition 8
ELIMINATES RIGHT OF SAME-SEX COUPLES TO MARRY. INITIATIVE CONSTITUTIONAL AMENDMENT
Changes California Constitution to eliminate the right of same-sex couples to marry. Provides that only marriage between a man and a woman is valid or recognized in California.
Fiscal Impact: Over next few years, potential revenue loss, mainly sales taxes, totaling in the several tens of millions of dollars, to state and local governments. In the long run, likely little fiscal impact on state and local governments.
Yes
No
Proposition 9
CRIMINAL JUSTICE SYSTEM. VICTIMS' RIGHTS. PAROLE. INITIATIVE CONSTITUTIONAL AMENDMENT AND STATUTE.
Requires notification to victim and opportunity for input during phases of criminal justice process, including bail, pleas, sentencing and parole. Establishes victim safety as consideration for bail or parole.
Fiscal Impact: Potential loss of state savings on prison operations and increased county jail costs amounting to hundreds of millions of dollars annually. Potential net savings in the low tens of millions of dollars annually on parole procedures.
Yes
No
Proposition 10
ALTERNATIVE FUEL VEHICLES AND RENEWABLE ENERGY. BONDS. INITIATIVE STATUTE.
Authorizes $5 billion in bonds paid from state's General Fund, to help consumers and others purchase certain vehicles, and to fund research in renewable energy and alternative fuel vehicles.
Fiscal Impact: State cost of about $10 billion over 30 years to repay bonds. Increased state and local revenues, potentially totaling several tens of millions of dollars through 2019. Potential state administrative costs up to about $10 million annually.
Bonds Yes
Bonds No
Proposition 11
REDISTRICTING. INITIATIVE CONSTITUTIONAL AMENDMENT AND STATUTE.
Changes authority for establishing state office boundaries from elected representatives to commission. Establishes multilevel process to select commissioners from registered voter pool. Commission comprised of Democrats, Republicans, and representatives of neither party.
Fiscal Impact: Potential increase in state redistricting costs once every ten years due to two entities performing redistricting. Any increase in costs probably would not be significant.
Yes
No
Proposition 12
VETERANS' BOND ACT OF 2008.
This act provides for a bond issue of nine hundred million dollars ($900,000,000) to provide farm and home aid for California veterans.
Fiscal Impact: Costs of about $1.8 billion to pay off both the principal ($900 million) and interest ($856 million) on the bonds; costs paid by participating veterans. Average payment for principal and interest of about $59 million per year for 30 years.
Bonds Yes
Bonds No
Los Rios Community College District Measure M
"Shall the Los Rios Community College District be authorized to issue $475 million in bonds at the lowest available interest rates to improve student academic performance by building classrooms, facilities and labs throughout the district including for teaching green technologies; nursing and healthcare programs; architecture, engineering and construction management; computer sciences; early childhood development; and fire and police public safety programs at the American River, Cosumnes River, El Dorado, Folsom, and Sacramento City College campuses?"
Bonds Yes
Bonds No
City of Davis Measure N
Shall the Proposed Charter of the City of Davis Be Adopted?
Yes
No
Davis Joint Unified School District Measure W
Shall the Davis Joint Unified School District preserve existing classroom programs including math and science, English, music, physical education, librarians, secondary class size reduction, athletics and co-curricular programs including drama, debate, and journalism by being authorized to levy a special tax for a period of three years not to exceed the annual rate of $50.00 per dwelling unit for multi-dwelling parcels and $120.00 per parcel for all other parcels?
Yes
No
10000630200062
1063031149
Vote Both Sides


Presidential General Election
YOLO COUNTY
November 04, 2008
Proposition 8
ELIMINATES RIGHT OF SAME-SEX COUPLES TO MARRY. INITIATIVE CONSTITUTIONAL AMENDMENT
Changes California Constitution to eliminate the right of same-sex couples to marry. Provides that only marriage between a man and a woman is valid or recognized in California.
Fiscal Impact: Over next few years, potential revenue loss, mainly sales taxes, totaling in the several tens of millions of dollars, to state and local governments. In the long run, likely little fiscal impact on state and local governments.
Yes
No
Proposition 9
CRIMINAL JUSTICE SYSTEM. VICTIMS' RIGHTS. PAROLE. INITIATIVE CONSTITUTIONAL AMENDMENT AND STATUTE.
Requires notification to victim and opportunity for input during phases of criminal justice process, including bail, pleas, sentencing and parole. Establishes victim safety as consideration for bail or parole.
Fiscal Impact: Potential loss of state savings on prison operations and increased county jail costs amounting to hundreds of millions of dollars annually. Potential net savings in the low tens of millions of dollars annually on parole procedures.
Yes
No
Proposition 10
ALTERNATIVE FUEL VEHICLES AND RENEWABLE ENERGY. BONDS. INITIATIVE STATUTE.
Authorizes $5 billion in bonds paid from state's General Fund, to help consumers and others purchase certain vehicles, and to fund research in renewable energy and alternative fuel vehicles.
Fiscal Impact: State cost of about $10 billion over 30 years to repay bonds. Increased state and local revenues, potentially totaling several tens of millions of dollars through 2019. Potential state administrative costs up to about $10 million annually.
Bonds Yes
Bonds No
Proposition 11
REDISTRICTING. INITIATIVE CONSTITUTIONAL AMENDMENT AND STATUTE.
Changes authority for establishing state office boundaries from elected representatives to commission. Establishes multilevel process to select commissioners from registered voter pool. Commission comprised of Democrats, Republicans, and representatives of neither party.
Fiscal Impact: Potential increase in state redistricting costs once every ten years due to two entities performing redistricting. Any increase in costs probably would not be significant.
Yes
No
Proposition 12
VETERANS' BOND ACT OF 2008.
This act provides for a bond issue of nine hundred million dollars ($900,000,000) to provide farm and home aid for California veterans.
Fiscal Impact: Costs of about $1.8 billion to pay off both the principal ($900 million) and interest ($856 million) on the bonds; costs paid by participating veterans. Average payment for principal and interest of about $59 million per year for 30 years.
Bonds Yes
Bonds No
Los Rios Community College District Measure M
"Shall the Los Rios Community College District be authorized to issue $475 million in bonds at the lowest available interest rates to improve student academic performance by building classrooms, facilities and labs throughout the district including for teaching green technologies; nursing and healthcare programs; architecture, engineering and construction management; computer sciences; early childhood development; and fire and police public safety programs at the American River, Cosumnes River, El Dorado, Folsom, and Sacramento City College campuses?"
Bonds Yes
Bonds No
City of Davis Measure N
Shall the Proposed Charter of the City of Davis Be Adopted?
Yes
No
10000630200062


Vote Both Sides
Precinct 100063
Joint Unified School District
Davis Joint Unified School District
existing classroom programs
math and science, English, music,
education, librarians, secondary class
athletics and co-curricular
including drama, debate, and
authorized to levy a
for a period of three years not to
annual rate of $50.00 per dwelling
multi-dwelling parcels and $120.00 per
all other parcels?
Yes
No
technologies; nursing and
programs; architecture, engineering and
construction management; computer sci
early childhood development; and fire a
police public safety programs at the Am
River, Cosumnes River, El Dorado, Fols
and Sacramento City College campuses
Bonds Yes
Bonds No
City of Davis
Measure N
Shall the Proposed Charter of the City o
Davis Be Adopted?
Yes
No

2009 Yolo County, CA Measure P Audit





Special Election November 2009
City of Davis
November 03, 2009

Precinct

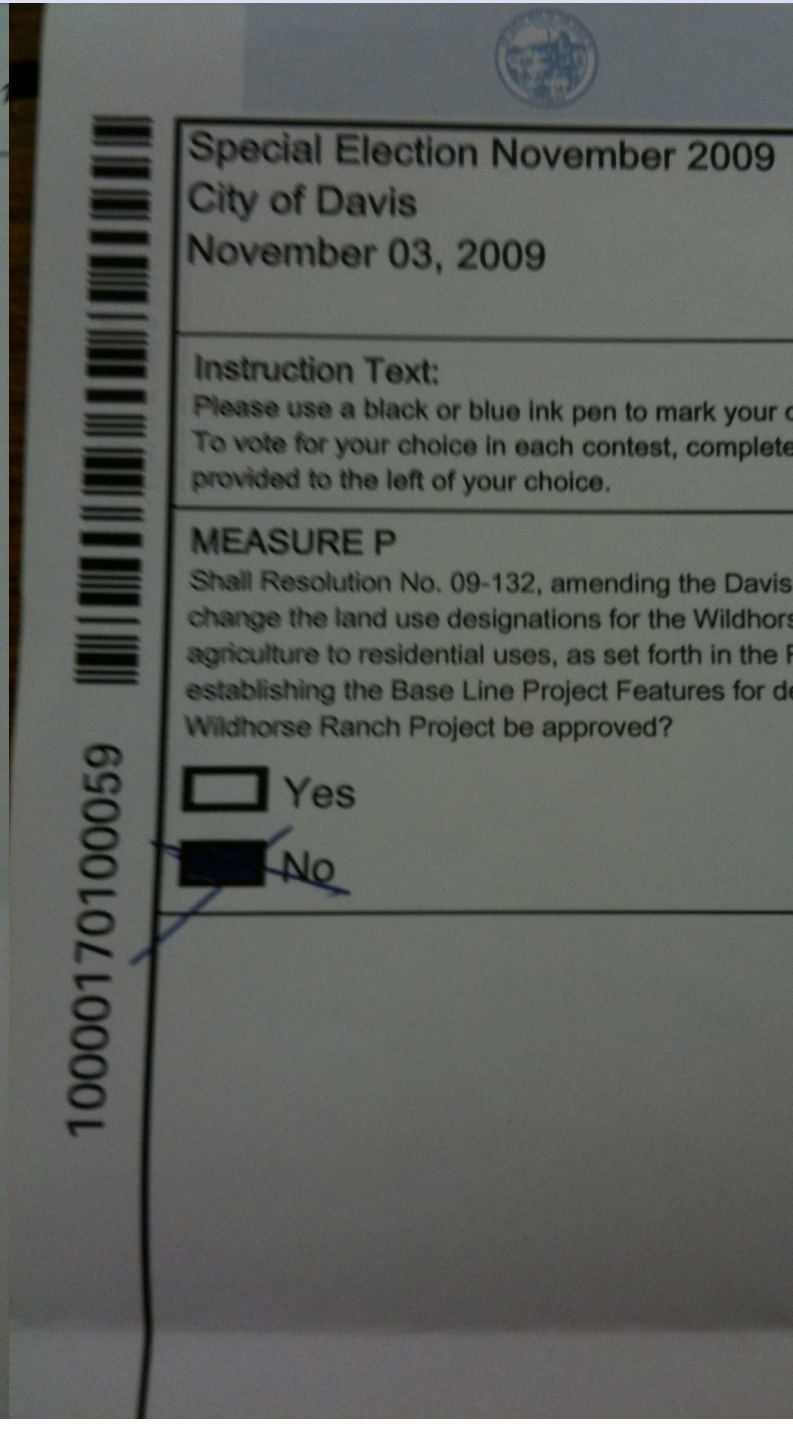
Instruction Text:
Please use a black or blue ink pen to mark your choices on the ballot. To vote for your choice in each contest, completely fill in the box provided to the left of your choice.

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 c
To vote for your choice in each contest, complete
provided to the left of your choice.

MEASURE P
Shall Resolution No. 09-132, amending the Davis
change the land use designations for the Wildhors
agriculture to residential uses, as set forth in the
establishing the Base Line Project Features for d
Wildhorse Ranch Project be approved?

☐ Yes
■ No

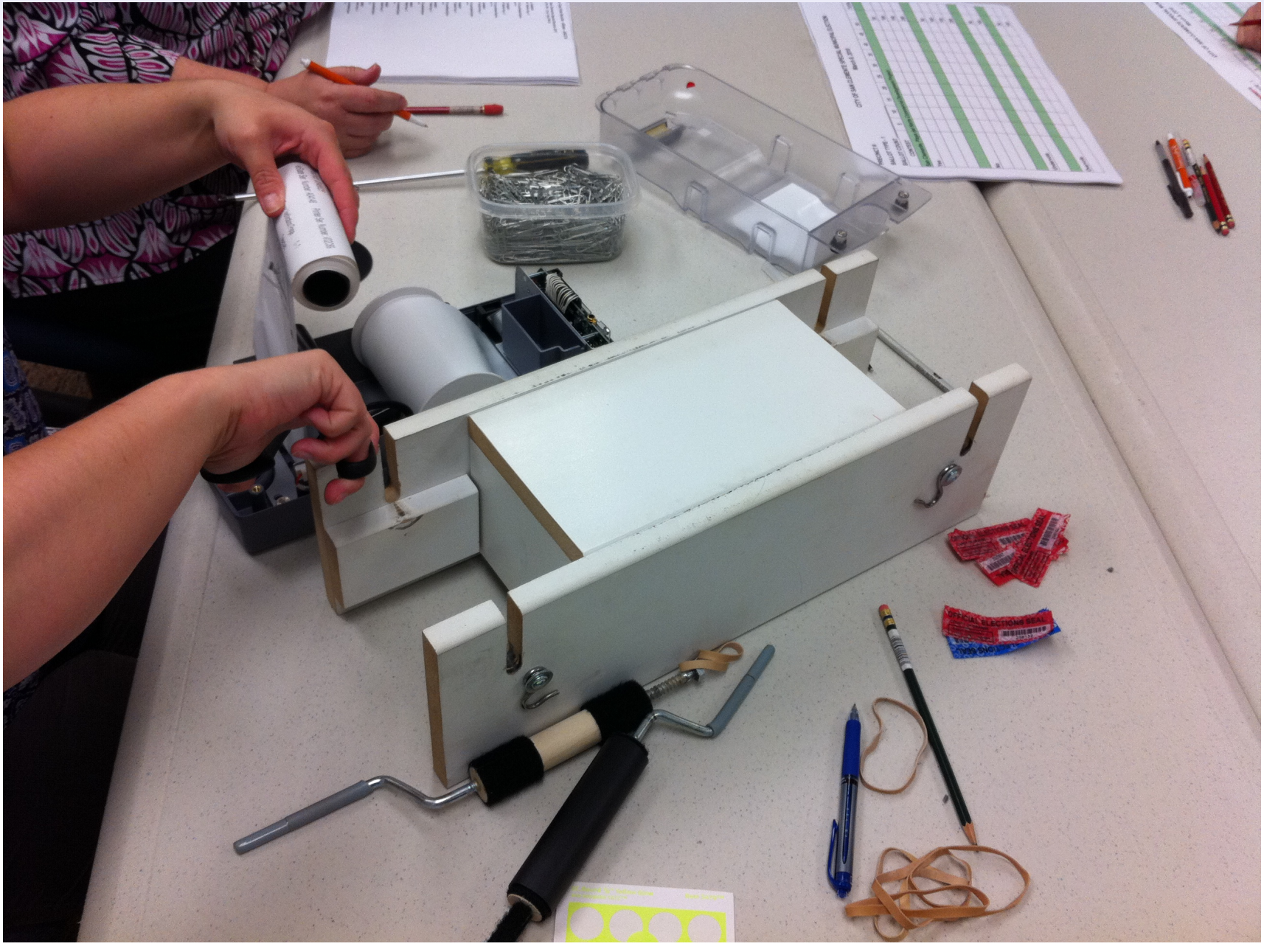
10000170100059

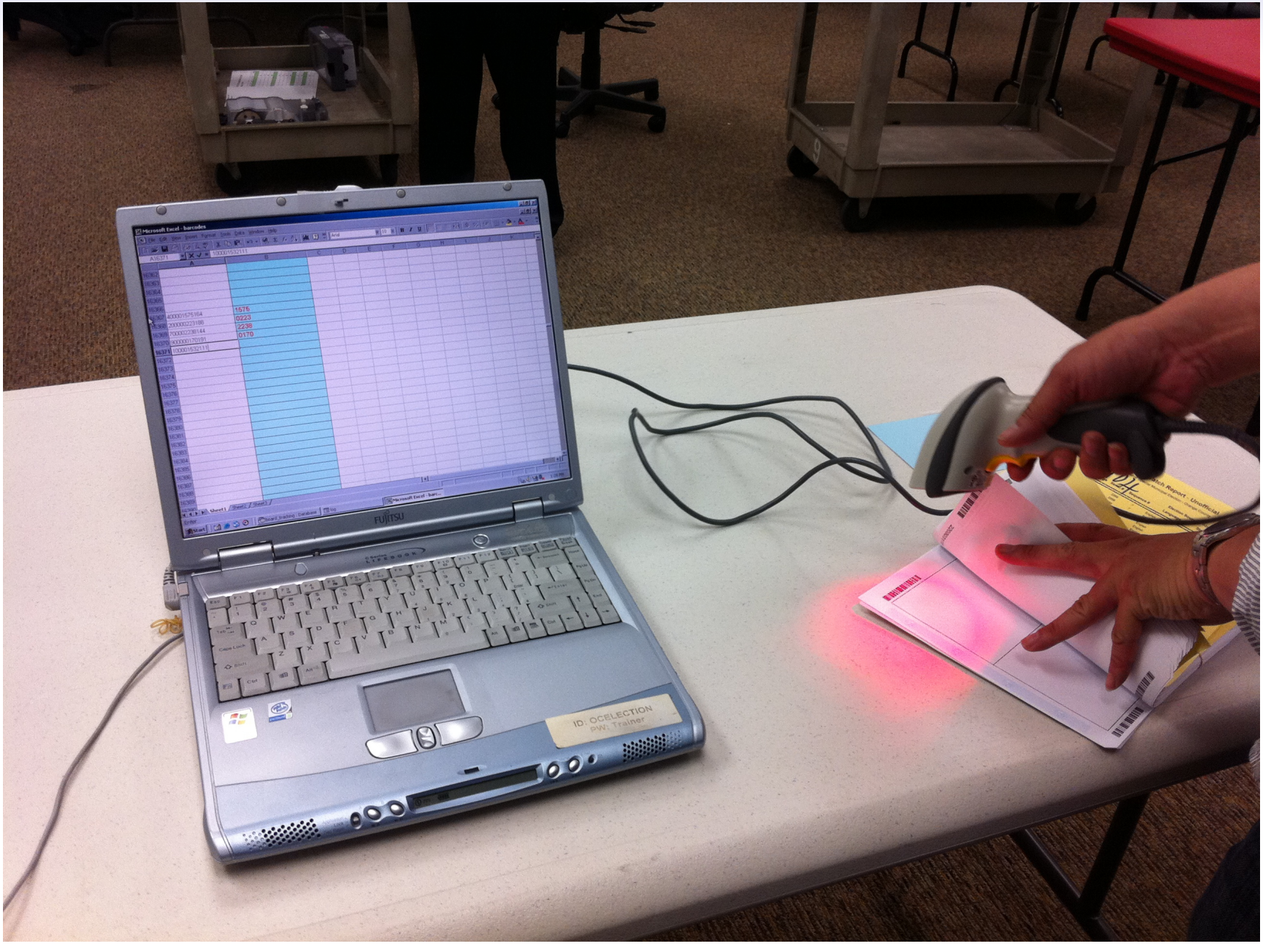
2011 Orange County, first audit under AB 2023

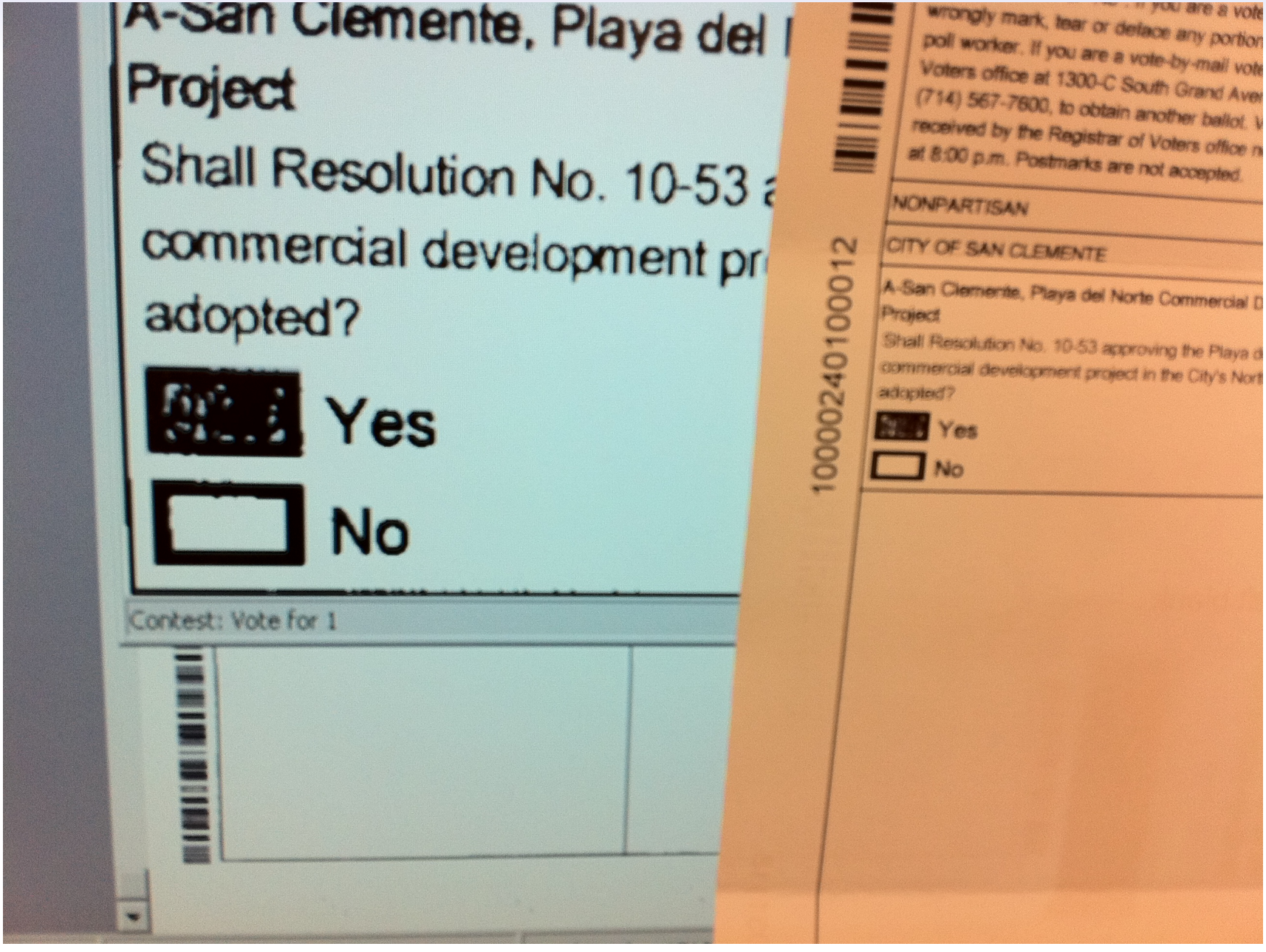

SCANNER #4
BATCH #1-12
SCANNER #4
BATCH #13-15, 17-24
DELIVERED #6, 16
SCANNER #6
BATCH #22-30
SCANNER #6
BATCH #41-54
SCANNER #6
BATCH #55-93
RULES FOR DROPPING OFF VOTE-BY-MAIL BALLOTS
• VERIFY ENVELOPE IS FILLED OUT SIGNED AND DATED
• ONLY A RELATIVE OR A PERSON RESIDING WITH VOTER MAY DROP OFF BALLOT IN PLACE OF VOTER
STATEWIDE GENERAL ELECTION
TO BE OPENED BY POLL SITE INSPECTOR ONLY
Neal Kelley • Orange County Registrar of Voters
POLL WORKER PASS
CARRY YOUR POLL WORKER PASS CARD AT ALL TIMES - FULL DETAILS AT OCVOTE.COM/PASS
TO BE OPENED BY POLL SITE INSPECTOR ONLY
Neal Kelley • Orange County Registrar of Voters
POSTAL SERVICE

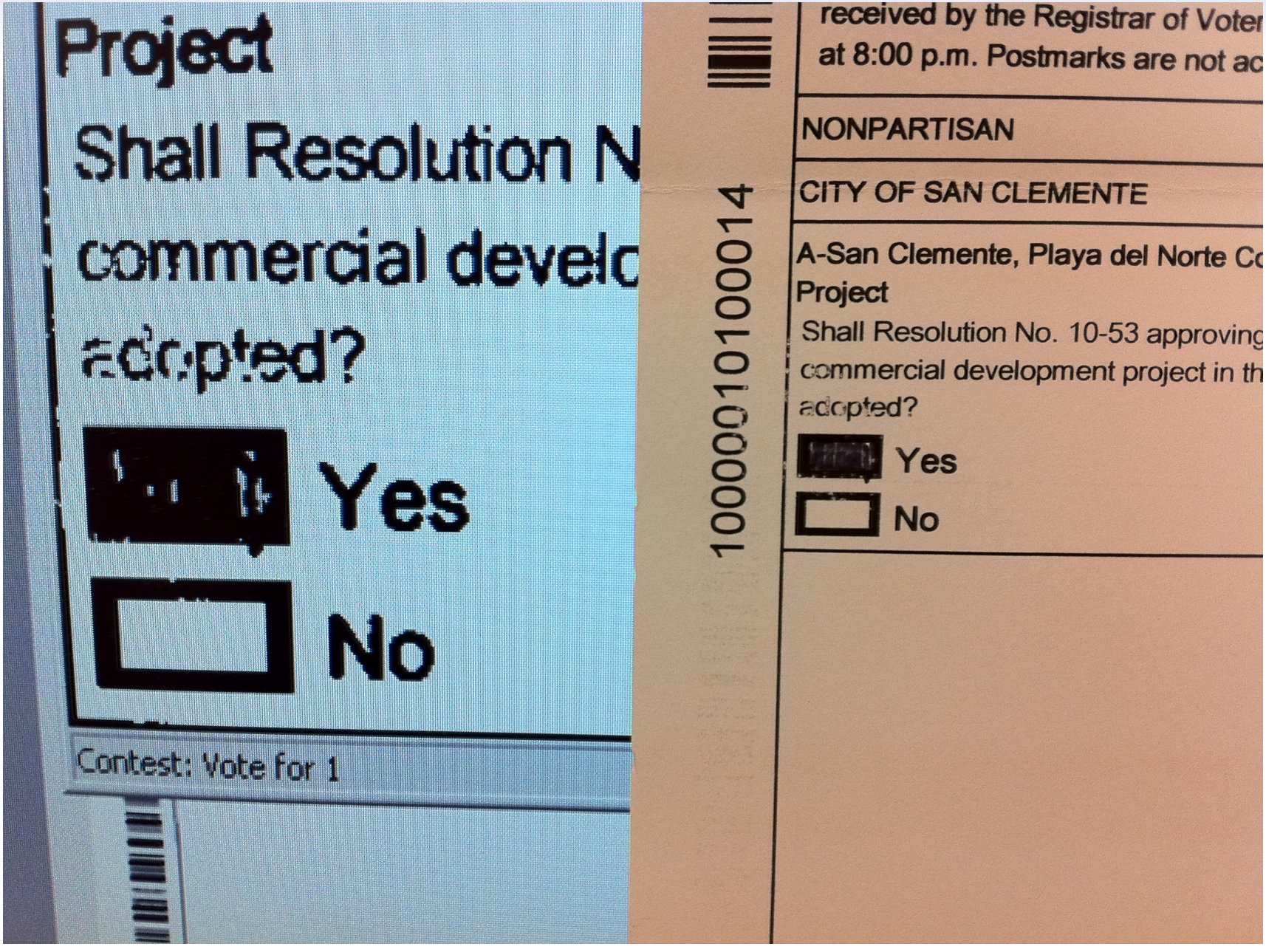

TAMPER
HART
IENTE SPECIAL MUNICIPAL ELECTION
March 8, 2010

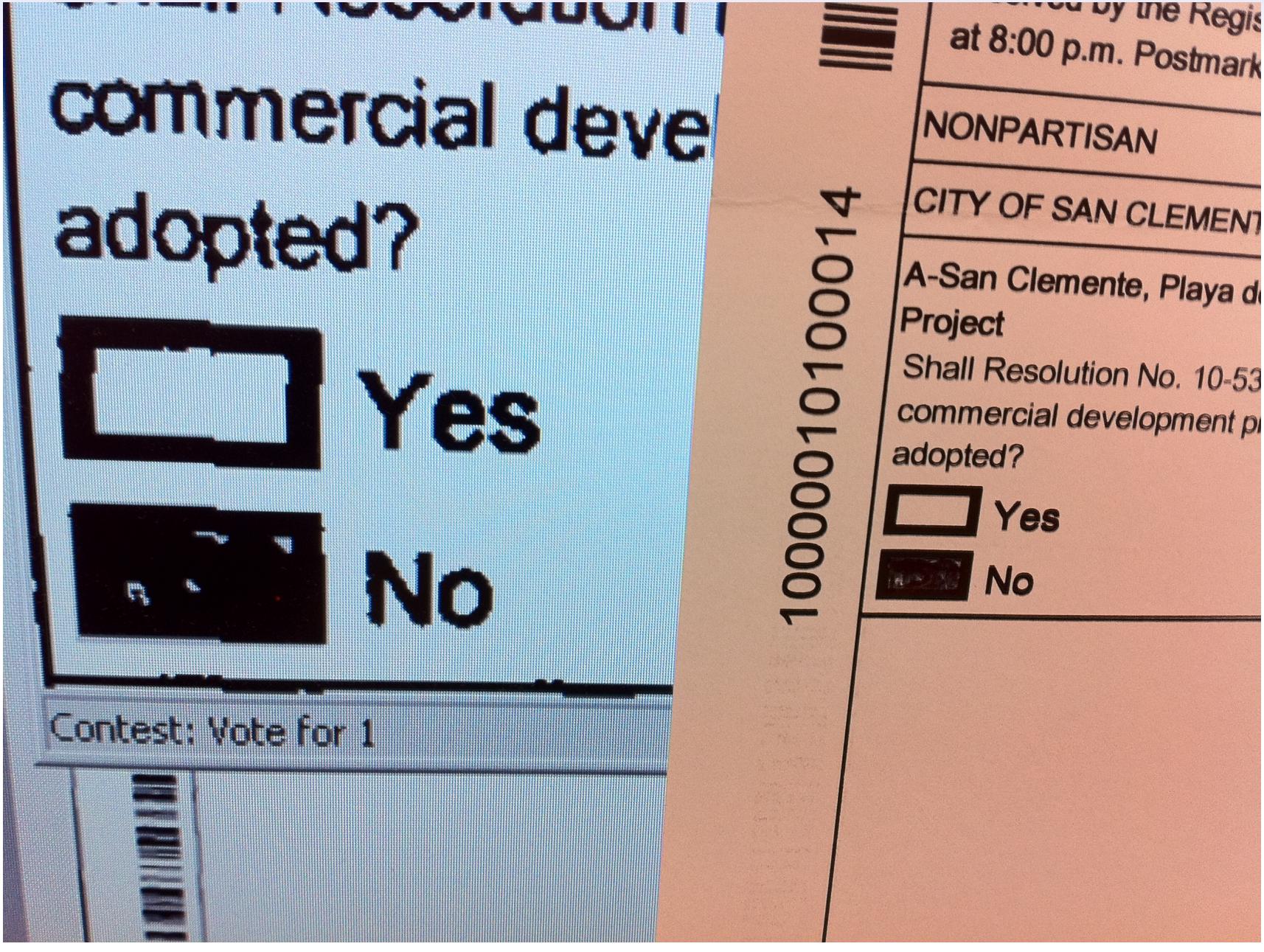

Notice
OPEN DOOR SLOWLY

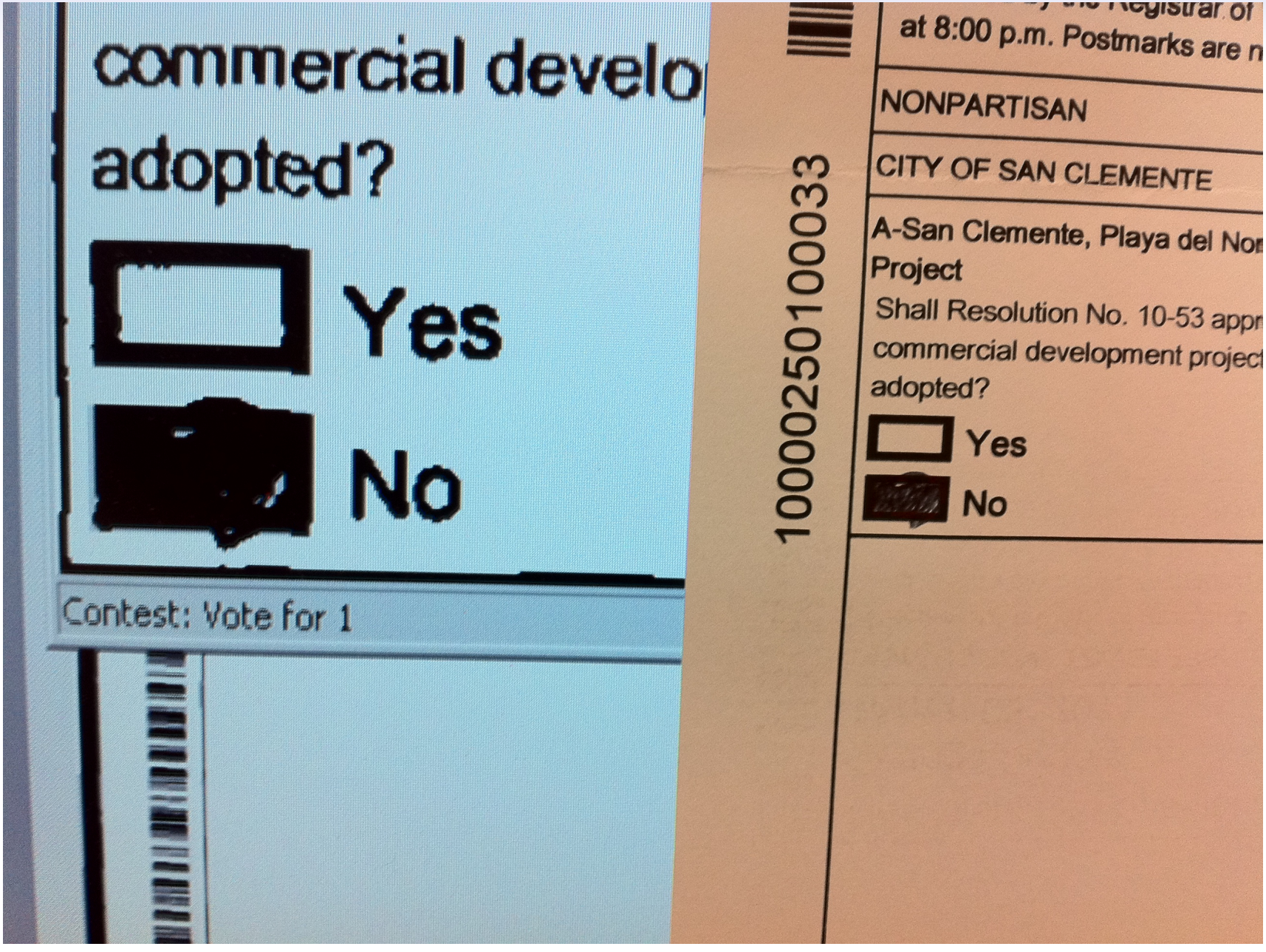

OFFICIAL ELECTION SEAL

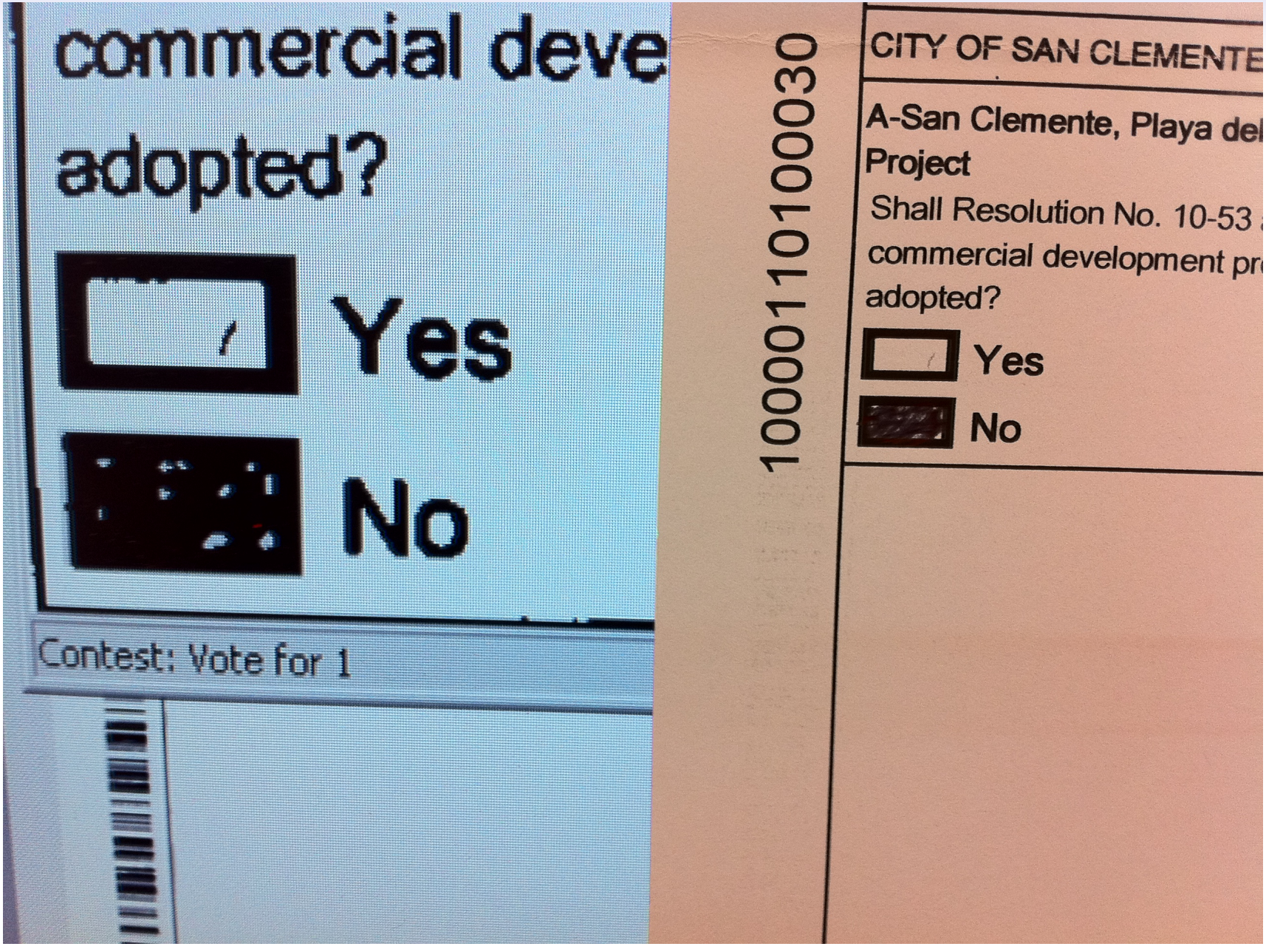

FUJITSU
ID: OCELECTION
PW: Trainer

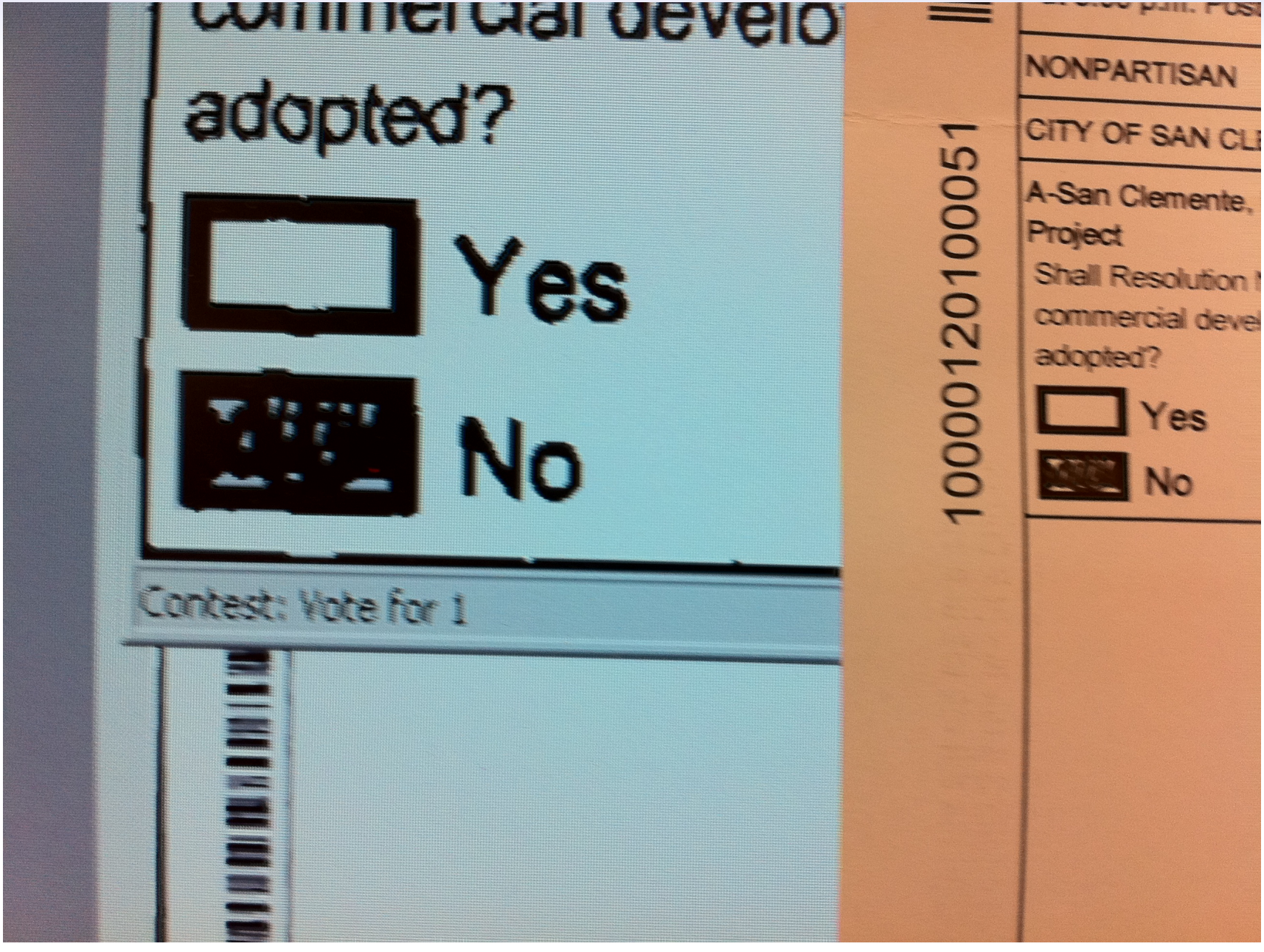

A-San Clemente, Playa del
Project
Shall Resolution No. 10-53
commercial development pr
adopted?
Yes
No
Contest: Vote for 1
wrongly mark, tear or deface any portion
poll worker. If you are a vote-by-mail vote
Voters office at 1300-C South Grand Aven
(714) 567-7600, to obtain another ballot.
received by the Registrar of Voters office
at 8:00 p.m. Postmarks are not accepted.
NONPARTISAN
CITY OF SAN CLEMENTE
A-San Clemente, Playa del Norte Commercial
Project
Shall Resolution No. 10-53 approving the Playa
commercial development project in the City's Nort
adopted?
Yes
No
10000240100012

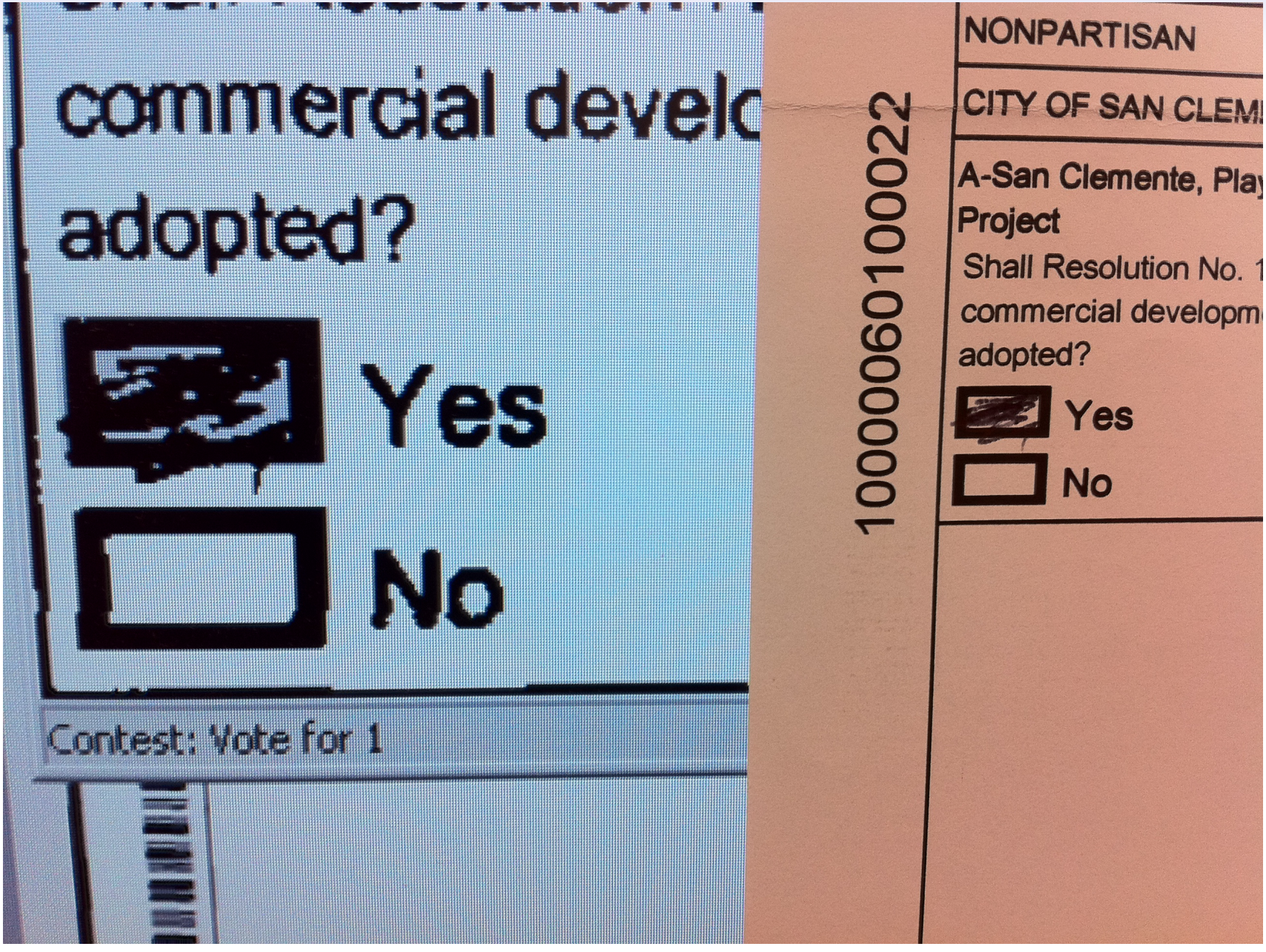

Project
Shall Resolution N
commercial develo
adopted?
Yes
No
Contest: Vote for 1
received by the Registrar of Voter
at 8:00 p.m. Postmarks are not ac
NONPARTISAN
CITY OF SAN CLEMENTE
A-San Clemente, Playa del Norte Co
Project
Shall Resolution No. 10-53 approving
commercial development project in th
adopted?
Yes
No
10000010100014

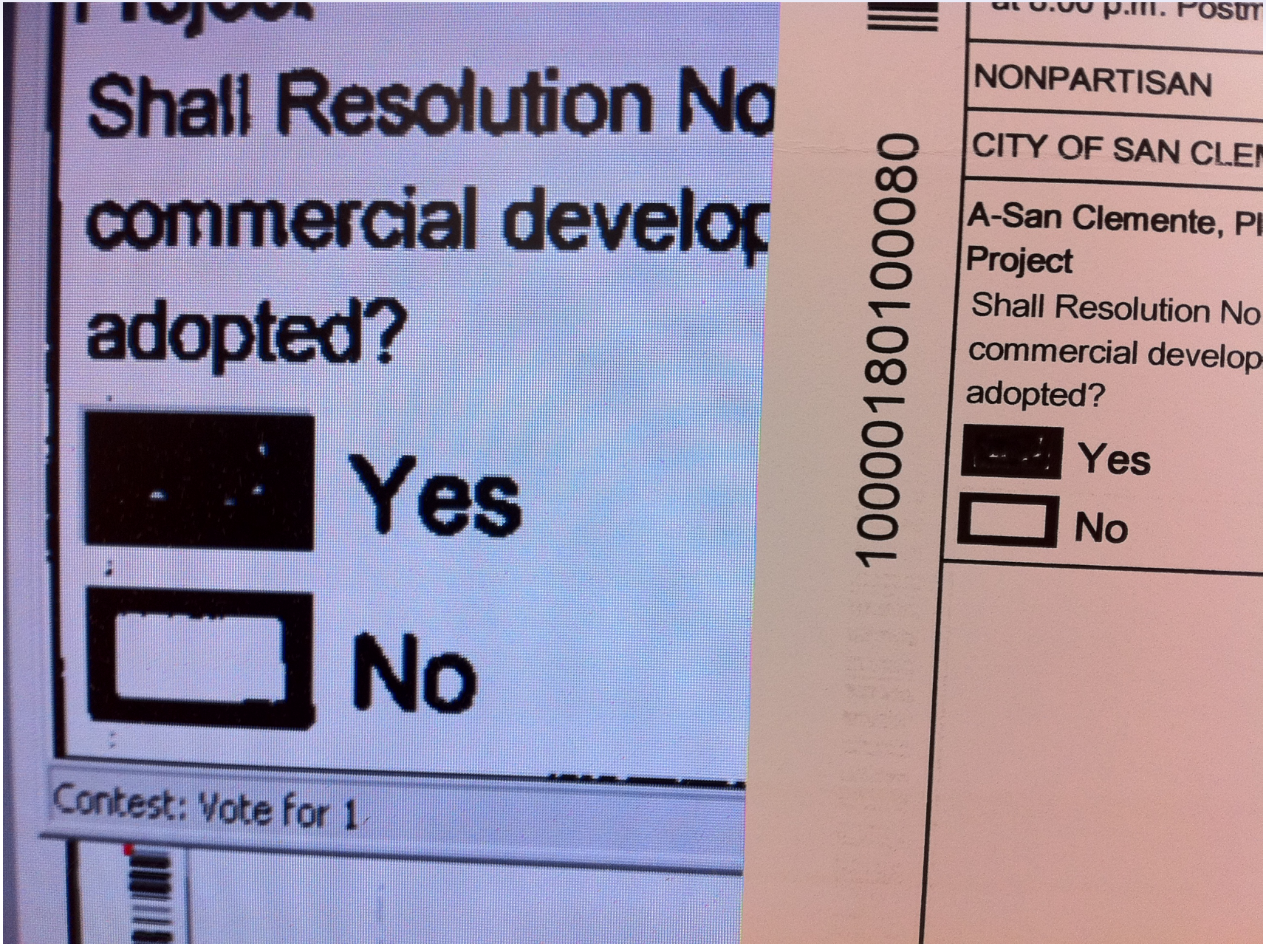

commercial deve
adopted?
Yes
No
Contest: Vote for 1
at 8:00 p.m. Postmark
NONPARTISAN
CITY OF SAN CLEMENT
A-San Clemente, Playa d
Project
Shall Resolution No. 10-53
commercial development p
adopted?
Yes
No
10000010100014


commercial develo
adopted?
Yes
No
Contest: Vote for 1
at 8:00 p.m. Postmarks are n
NONPARTISAN
CITY OF SAN CLEMENTE
A-San Clemente, Playa del Nor
Project
Shall Resolution No. 10-53 appr
commercial development project
adopted?
Yes
No
10000250100033


commercial deve
adopted?
Yes
No
Contest: Vote for 1
100001101000030
CITY OF SAN CLEMENTE
A-San Clemente, Playa del
Project
Shall Resolution No. 10-53
commercial development pr
adopted?
Yes
No


adopted?
Yes
No
Contest: Vote for 1
10000120100051
NONPARTISAN
CITY OF SAN CL
A-San Clemente,
Project
commercial devel
adopted?
Yes
No


commercial develo
adopted?
Yes
No
Contest: Vote for 1
10000060100022
NONPARTISAN
CITY OF SAN CLEM
A-San Clemente, Play
Project
Shall Resolution No. 1
commercial developm
adopted?
Yes
No


Shall Resolution No
commercial develop
adopted?
Yes
No
Contest: Vote for 1
10000180100080
NONPARTISAN
CITY OF SAN CLEM
A-San Clemente, Pl
Project
Shall Resolution No
commercial develop
adopted?
Yes
No

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.

Orange County 2011 Audit design and sample

Left provisionals in machine ballot counts for error bounds. 5523 total.

One VBM-only precinct with 119 ballots. 158 election-day paper ballots. 38 rejected provisional ballots

Used deck of cards to pick 9-digit seed: shuffled cards well, counted Ace as 1, etc., 10 as 0, and ignored face cards, dealt until we had 9 digits. Used R implementation of Mersenne Twister.

Sample gave 12 eSlate machines with a total of 446 ballots, and 21 individual ballots. Total sample size 467 ballots (expected size was 384.8 ballots). One of the eSlates had already been audited as part of the statutory 1% audit.

1% Statutory Audit

Votes in one precinct counted by hand. No errors found.

Chance the 1% audit would find no errors even if the outcome is wrong could be over 88%.

Statutory audit does little to limit risk, even if it required a full hand count if errors were found.

Risk-limiting Audits: Costs

San Clemente Measure A, 3/8/2011

1% Statutory Audit: $257.68
Scales as the size of the contest: a contest twice as large would cost about twice as much to audit.

Risk-limiting: $483.79 (does not include my time or airfare)
Would have cost essentially the same for any contest with the same percentage margin, no matter how large the contest.

SOBA: Preserve voter anonymity, better verifiability

Way to audit that:

- Has a big chance of correcting the outcome if the outcome is wrong (risk-limiting).
- Enables the public to have strong evidence that the outcome is right, without having to trust (many) others.
- Preserves voter anonymity.
- Is efficient, affordable, and currently feasible—modulo re-scanning costs.

Motivation

- Risk-limiting audits now widely considered best practice.
- Comparison audit of individual ballots requires least counting.
- Auditing individual ballots increases transparency.
- Simultaneously auditing all contests on each selected ballot can increase efficiency.
- Publishing data at the ballot level can compromise voter privacy.
- But if the raw data aren't published, public might not trust the results or the audit.
- Can we keep the benefits of simultaneous comparison auditing at the ballot level and have data transparency without compromising privacy?
- E2E could do it, but requires changes, heavy crypto, "critical mass" of voters.
- Is there a bolt-on solution that doesn't require much change to voting systems or procedures, and that relies less on mathy stuff?

Goal of SOBA

Personally verifiable privacy-preserving *P*-resilient canvass framework.

WTF?

More Definitions

- *Canvass framework*: the vote-tabulation system together with other human, hardware, software, and procedural components of the canvass, including compliance audit and other audits.
- Canvass framework is *resilient with probability P* or *P-resilient* if the probability that the outcome it gives is the correct outcome is at least *P*, even if its software has an error, shortcoming, or undetected change: System tends to recover from (some) faults. (Strong software independence [Rivest & Wack], plus procedures that exploit that independence.)
- *P*-resilience can mean requiring a re-vote if the audit trail can't be shown to be in good shape.

and more . . .

- Canvass framework is *personally verifiable P-resilient* if it is *P*-resilient and a single individual could, as a practical matter, observe enough of the process to have convincing evidence that the canvass framework is in fact *P*-resilient.
- *Personally verifiable privacy-preserving P-resilient* canvass framework: personally verifiable *P*-resilient and it does not sacrifice privacy unnecessarily.

Neither *personally verifiable* nor *privacy-preserving* is mathematically precise; *P*-resilience is.

“Personally verifiable” and “privacy-preserving” can be defined separately from “P-resilience.”

SOBA++

- Adds a special risk-limiting audit to a strongly software-independent voting system that has had a compliance audit.
- Publishes results by ballot by contest: anybody can verify outcomes.
- Does not allow public to reconstruct whole-ballot CVRs, to protect privacy.
- Uses cryptographic commitment to allow auditors and observers to reconstruct the ballots selected for audit.
- Audit checks accuracy of CVRs *and* of the cryptographic commitment.

Aside: cryptographic commitments

- Ensures that the ballot identifier is secret but indelible, so every ballot is properly reflected in the electronic results.
- Select and publish commitment function $H()$.
- To commit that a given CCVR comes from ballot b, LEO selects secret "salt" u and computes $y = H(b, u)$. Publishes shrouded ID (SID) y.
- If ballot b is selected for audit, LEO can reveal u and b: Anyone can check whether $y = H(b, u)$.

Commitment function key properties: *binding* (*collision-resistant*), and *hiding* (*one-way*).

- *Binding*: infeasible to find any pair $(b', u') \neq (b, u)$ for which $H(b', u') = H(b, u)$. Helps ensure nobody can claim more than one CCVR for a given contest comes from the same ballot.
- *Hiding*: infeasible for anyone with access only to the SIDs to learn anything about which ballot is involved in each commitment.

Salt should be random number with at least 128 digits.

SOBA preparations

X contests, N_χ ballots cast in contest χ, N ballots in all, M voting opportunities in all.

- Compliance audit, including ballot accounting: determine $\{N_\chi\}$, N, M.
- Find apparent outcomes of the X contests.
- Construct CVR for each ballot (perhaps by unofficial scan: transitive auditing); assign unique ID to each ballot.
- Disaggregate CVRs into X per-contest sets of *CCVRs*; Publish X CCVR files. N_χ lines in file X, each gives CCVR and SID. Sort by SID.
- Publish *ballot style file*. N lines. Each line lists contests on ballot and a unique ballot ID (e.g., #17,097, or 275th in 39th deck).
- Construct (but don't publish) *lookup file*. M lines, 3 entries per line: SID, corresponding unshrouded ID b, and "salt" u
- Select and disclose H, risk limit, PRNG.

What can go wrong?

The CCVRs might fail to be sufficiently accurate because

- At least one CCVR and the ballot it purports to represent do not match because human and machine interpretations of voter intent differ (for instance, because the voter marked the ballot improperly). This is a failure of the generation of CCVRs.
- At least one CCVR does not in fact correspond to any ballot. It is an "orphan." This is a failure of the mapping between ballots and CCVRs.
- More than one CCVR for the same contest is mapped to the same ballot. It is a "multiple." This is also a failure of the mapping between ballots and CCVRs.
- There is no CCVR corresponding to some voting opportunity on a ballot.

Audit checks these things *while* checking the accuracy of the CCVRs, with the same sample.

SOBA Audit at 10% risk limit

1. Verify that, for each contest χ, there are not more than N_χ entries in the CCVR file for contest χ.
2. Verify that, for each contest χ, the CCVR file shows the same outcome (not count!) as the reported outcome. If not, hand count any discrepant contests.
3. Verify that the $M = N_1 + \cdots + N_X$ shrouded ballot identifiers in all X CCVR files are unique.

4. Verify that, for each contest χ, there are N_χ entries in the ballot style file that list the contest.
5. Verify that the ballot identifiers in the ballot style file are unique.

If 1, 3, 4, or 5 fails, LEO needs to correct before risk-limiting stage of audit can start.

6. Set audit parameters:
   - 6.1. Find *diluted margin m* from CCVRs: smallest apparent margin in votes for any contest, divided by N.
   - 6.2. Select maximum number of draws D before conducting a full hand count; set $d = 0$.
   - 6.3. Select a *seed* s. Observers could contribute to s or roll dice.
7. Select a pseudo-random number between 1 and N. Find that row in the ballot style file; retrieve corresponding ballot. $d \leftarrow d + 1$.
8. If ballot cannot be found, treat ballot as valid vote for all losers in all contests. Compare CVR with ballot for all contests on the ballot. If ballot has a contest the style file doesn't show, treat CCVR as vote for apparent winner. If style file says ballot has a contest ballot doesn't, treat ballot as valid vote for all losers in that contest.
9. o_1, o_2, u_1, u_2 are numbers of 1 and 2-vote overstatements and understatements in sample so far. Stop audit if

$$d \geq \frac{4.8 + 1.4(o_1 + 5o_2 - 0.6u_1 - 4.4u_2)}{m}. \tag{3}$$

Else if $d = D$, conduct full hand count.

Else go to step 7.

Research directions

- IRV/RCV, NPV
- “False winner rate”
- Extending KM/Wald to stratified cluster samples
- Sharper test given sampling design (Shacham et al. use KL distance for ballot-level)
- Optimal tests if sampling design is up for grabs. Concentration inequalities? Feige?
- Auditing E2E encrypted systems (Wallach, Pereira, et al.)
- Simpler, simpler, simpler

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—possibly not certified system.

Reducing counting effort is mostly about reducing batch sizes.

Hopes and plans

- Move to evidence-based requirements instead of equipment-based requirements.
- Work with elections officials at the state and local level, integrity advocates, vendors, computer scientists, political scientists, statisticians, financial auditors, attorneys, to draft model legislation for election auditing. (White paper forthcoming in a matter of weeks; result of 1-year collaboration.)
- Clarify tradeoff of risks and costs. What kinds of errors are we (as a society) willing to tolerate? With what frequency? What are we willing to pay? How long are we willing to make the canvass?
- Work with computer scientists, usability experts, and others to build voting systems that support efficient audits. (E.g., STAR-Vote w/ Wallach, Benaloh, Byrne, Kortum, Pereira.)
- Do the work to put theory into practice, to create resilient canvass frameworks.

GOTA: Get out the Audit!

Ballot-polling audits are possible for the November 2012 presidential election in any jurisdiction that has VVPRs—and has knows how many and where they are.

Workload *not* large in most states; preparations minimal.
Equipment needed: dice, pencil, and paper.
(Alternatively, dice and simple web-based tools.)

Compliance audit needs attention—ensure audit trail adequately accurate.
Coordination across jurisdictions needs attention—logistics and transparency.

Verified Voting Foundation is working to get ballot-polling audits in several states for November 2012 presidential election.

Let's Get out the Audit!

Connections to other things

- Tax audits, financial audits. Working with NM Department of Taxation.
- Healthcare audits.
- Auditing science: Reproducible research versus reproduced research.
  Want to know that everything required to reproduce the work was published, and that following the steps really does reproduce the results.