Orange County, CA
Pilot Risk-Limiting Audit
December 7, 2018

Verifiedvoting

By Stephanie Singer and Neal McBurnett,
for the Verified Voting Foundation.

This report was made possible with financial support from Microsoft.

Introduction

In June 2018, the Orange County Registrar of Voters conducted a pilot audit of all countywide election contests. The goal of the audit pilot was to collect useful data on the process of conducting a risk-limiting audit, as well as collecting evidence that the outcomes of the election were true and correct. This audit practiced some new approaches to risk-limiting audits (RLAs), while also identifying areas in which new techniques could make RLAs more efficient and effective. RLAs are one element of "evidence-based elections,"[1] and should be combined with audits of other aspects of an election.

The election agency conducted a risk-limiting audit of three out of five countywide election contests: County Assessor, Auditor-Controller, and Clerk-Recorder.[2] The audit provided good evidence for the correctness of the tabulation outcomes of these three contests and established maximum risk levels based on evidence for the outcomes of two other contests, subject to caveats as discussed below. The cost to carry out all activities described in the Audit Process section below was under $4,000. The audit did not rely on digital images.[3]

Supplementary Materials

Supplementary materials are available in an online GitHub repository[4] for reference and exploration. These materials include open-source software, input data for the audit, and Jupyter notebooks[5] with worked examples and additional analysis allowing users to reproduce the calculations.

---

[1] Evidence-Based Elections, Philip B. Stark & David Wagner, IEEE Security & Privacy (Volume: 10 , Issue: 5 , Sept.-Oct. 2012) http://statistics.berkeley.edu/~stark/Preprints/evidenceVote12.pdf

[2] Although in California law the word "contest" refers to a challenge to an election outcome, in this report we use "contest" in a different, widely accepted way -- to refer to any single decision to be made by the electorate, such as a single ballot question, or who will fill one particular office.

[3] Except that, to a limited extent, digitally faxed images were used to create paper records of some UOCAVA voters' intent, without verification of those images by the voters.

[4] https://github.com/nealmcb/ocrla-2018p

[5] M., Randles, Bernadette; V., Pasquetto, Irene; S., Golshan, Milena (2017-07-27). "Using the Jupyter Notebook as a tool for open science: An empirical study" https://escholarship.org/uc/item/5w52878j

Table of Contents

Context

Risk-Limiting Audits

One purpose of the pilot was to use and assess a "risk-limiting" audit process. A risk-limiting audit has a predetermined chance of correcting a tabulation error that alters the apparent winner(s).

> Risk-limiting audits are designed to ensure strong audit evidence that a reported outcome is correct – if the outcome is indeed correct. ("Outcome" refers to consequence, such as who won. The "correct outcome" means whatever a full hand count would show.[6]) If a full hand count would show a different outcome than the initial tabulation, RLAs have a large chance of leading to a full hand count that corrects the reported outcome.[7] The corresponding predetermined small maximum chance that an RLA of an incorrect outcome will not lead to a full hand count is called the risk limit. Beyond providing high assurance in specified contests, RLAs can reduce overall audit burden by allocating more resources to closer contests where more checking is needed to validate outcomes. (Adapted from Principles and Best Practices for Post-Election Audits[8])

How large the "large chance" is depends on what "risk limit" is chosen. The June 2018 pilot in Orange County used a risk-limit of 20% in three contests. With no tabulation audit, there would be little chance of catching a mistaken outcome. With the currently common practice in California of auditing a fixed percentage of precincts but omitting late-arriving ballots, many tabulation errors could be detected, but errors specific to the omitted ballots would not be detected.[9] With a risk-limiting audit and a 20% risk limit, there is at least an 80% chance of not just catching but actually correcting a mistaken outcome.

Orange County Registrar of Voters piloted an RLA of a small contest in 2011 with the help of Philip Stark.[10] The 2018 pilot audit scaled the procedure up to more contests and many more ballots, and documented it in more detail.

---

[6] Appropriately done hand counts, typically via a "sort and stack" procedure, with verifiable stack counting methods are the most robust and accurate way to tabulate a single contest. See e.g. *Proposal for rules related to definition of "full hand count" for purposes of Colorado RLA and integration with existing statutory recount*. Harvie Branscomb and John McCarthy, 2017. https://www.sos.state.co.us/pubs/rule_making/written_comments/2017/20170718BranscombMcCarthy.pdf

[7] RLAs are designed never to overturn a correct outcome. Only a full hand count can change an outcome.

[8] https://www.verifiedvoting.org/wp-content/uploads/2018/11/Principles_Best-Practices_Tabulation-Audits-20181023.pdf

[9] *An Introduction to Risk-Limiting Audits and Evidence-Based Elections* Prepared for the Little Hoover Commission, 2 July 2018, Philip B. Stark https://www.stat.berkeley.edu/~stark/Preprints/lhc18.pdf

[10] https://community.amstat.org/blogs/steve-pierson/2011/05/09/california-and-colorado-piloting-risk-limiting-audits

Orange County June 2018 Tabulation

Equipment

Orange County uses the Hart InterCivic Voting System 6.2.1. This system provides two mechanisms for casting and counting ballots. Many voters (all who vote by mail, and those who choose paper ballots at vote centers or precinct polling places) mark their choices by hand on paper ballots. Each ballot consists of one or more pieces of paper, called "ballot sheets." Voters return these ballot sheets to the Board of Elections. These sheets are interpreted by the Hart BallotNow (version 3.3.11) central count scanners. Military and overseas voters (UOCAVA) may submit faxed images of the paper on which they mark their choices.[11] Note that each face -- front and back -- of a hand-marked paper ballot sheet has two preprinted identification numbers, one in the left margin and one in the right margin. These numbers are not unique – they repeat every 2500.

---

[11] Note that such faxed images received at the elections office cannot be verified by the voter.

Other voters who cast ballots in vote centers or precinct polling places cast their votes by interacting with computers called eSlates (version 4.2.13). The eSlates are Direct Recording Electronic (DRE) systems with a Voter-Verifiable Paper Audit Trail (VVPAT). Before finally casting a vote, each voter can examine and verify, through a window, a paper record that is intended to reflect the contents of the voter's ballot. All the paper records from each single eSlate machine in each single election are stored on a single continuous roll of paper.

Details of the eSlate system, including pictures and videos are available at VerifiedVoting.org.[12] Pictures and slides describing Stark's audit of a contest in 2011 which used the same Hart voting system can be found online.[13]

Preliminary Tabulation Results

Six contests appeared on every ballot in Orange County in June 2018. These were "nonpartisan" primary contests on the Orange County ballot. One (for Treasurer-Tax Recorder) had only one candidate, making auditing unnecessary. The other five were:
- Auditor-Controller
- Assessor
- Clerk-Recorder
- Sheriff-Coroner
- District Attorney-Public Administrator.

Article 8[14] of the California Elections Code sets out the rules for a "nonpartisan primary" contest. In such a contest, each voter can cast a vote for at most one candidate, and only one person will eventually be elected to the seat. If one candidate wins an outright majority -- more than 50% of all valid votes cast for candidates in the contest -- then that candidate wins the office based on the primary alone. Note that a candidate does not have to receive votes on a majority of ballots counted in the contest, but just a majority of the ballots with a valid vote in that contest.[15] Otherwise, no one wins the office outright and the top two vote-getters will both be on the ballot in the general election, which will determine the winner of the office. No candidates could receive write-in votes for any of these contests under audit on the June 2018 ballot, as no write-in candidates had pre-qualified as required by California law (Election Code Article 3 Section 15341).

Given the top-two rules, there were two kinds of outcomes. In some contests the initial tabulation showed an outright majority for one candidate. In this case the audit had to

---

[12] See https://www.verifiedvoting.org/resources/voting-equipment/hart-intercivic/eslate/.
[13] See Close enough for government [to] work: Risk-Limiting Post-Election Audits. Philip B. Stark Department of Statistics, UC Berkeley 12 April 2011 Joint Berkeley-Stanford Statistics Colloquium Stanford University, starting at page 91. https://www.stat.berkeley.edu/~stark/Seminars/stanford11.pdf
[14] See: http://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=ELEC&division=8.&title=&part=1.&chapter=1.&article=8
[15] Rejected ballots, ballots where there were no choices or too many choices marked by the voter and ballots with write-in votes for unqualified candidates were excluded.

address the possibility that the candidate in question received only 50% or fewer of the votes according to the paper records.

In other contests the tabulation showed two reported winners, neither with an outright majority. In this case the audit had to address several possibilities:
- Did any reported loser actually get more votes than one or more of the reported winners?
- Did either of the reported winners actually win a majority of the votes?

Auditor-Controller

There were two candidates, Smart and Woolery. The initial tabulation showed
- 369,704 votes for Woolery
- 127,768 votes for Smart.

According to this initial tabulation, Woolery won an outright majority. To provide evidence supporting this outcome, the audit had to address the possibility that Woolery did not in fact win an outright majority of the votes.

Assessor

There were three candidates, Epstein, Parrish and Ramirez. The initial tabulation showed
- 358,634 votes for Parrish
- 90,631 votes for Ramirez
- 70,831 votes for Epstein.

According to this initial tabulation, Parrish won an outright majority. To provide evidence supporting this outcome, the audit had to address the possibility that Parrish did not in fact win an outright majority of the votes.

Clerk-Recorder

There were two candidates, Nguyen and Rocco. The initial tabulation showed
- 407,638 votes for Nguyen
- 107,926 votes for Rocco.

According to this initial tabulation, Nguyen won an outright majority. To provide evidence supporting this outcome, the audit had to address the possibility that Nguyen did not in fact win an outright majority of the votes.

Sheriff-Coroner

There were three candidates, Barnes, Nguyen and Harrington. The initial tabulation showed
- 265,059 votes for Barnes
- 166,848 votes for Nguyen
- 104,623 votes for Harrington.

According to this initial tabulation, Barnes and Nguyen had earned spots on the November ballot for Sheriff-Coroner. To provide evidence supporting this outcome, the audit had to address all four of the following possibilities:
- Barnes won an outright majority
- Nguyen won an outright majority

- Harrington got more votes than Nguyen
- Harrington got more votes than Barnes.

District Attorney-Public Administrator

There were four candidates, Albert-Sheridan, Murdock, Rackauckas and Spitzer. The initial tabulation showed
- 209,148 votes for Rackauckas
- 191,346 votes for Spitzer
- 121,818 votes for Murdock
- 20,890 votes for Albert-Sheridan.

According to this initial tabulation, Rackauckas and Spitzer had earned spots on the November ballot for District Attorney-Public Administrator. To provide evidence supporting this outcome, the audit had to address all six of the following possibilities:
- Rackauckas won an outright majority
- Spitzer won an outright majority
- Murdock got more votes than Rackauckas
- Murdock got more votes than Spitzer
- Albert-Sheridan got more votes than Rackauckas
- Albert-Sheridan got more votes than Spitzer

Major Audit Design Decisions

Direct Visual Examination of Paper Records

The auditors based their assessment of voter intent on paper records that the voters had the opportunity to verify (except for UOCAVA ballots based on faxed transmission of images). Risk-limiting audit experts agree that, like any other computer-generated or computer-mediated artifact, digital images cannot be assumed to be trustworthy records.[16]

Algorithm

Limitations of the voting systems in use constrained the choice of algorithm for the risk-limiting audit. The most efficient known type of risk-limiting audit, a ballot-level comparison audit,[17] requires a voting system that produces cast vote records ("CVRs") linking the computer's interpretation of the votes cast on each ballot to the corresponding voter-verifiable paper record of voter intent.

Because the CVRs generated by the Ballot Now system cannot be matched to the corresponding paper ballots, a ballot-level comparison audit was not possible. However,

---

[16] As noted by the National Academies consensus study report, ballot images and other electronic evidence "can be altered by compromised or faulty hardware or software". National Academies of Science, Engineering, and Medicine, Securing the Vote: Protecting American Democracy (National Academies Press, 2018), available at https://www.nap.edu/catalog/25120/securing-the-vote-protecting-american-democracy p. 94.

[17] Note that "ballot-sheet-level comparison" would be more accurate.

another type of risk-limiting audit -- a ballot-polling audit -- was feasible based on the paper records alone. Orange County chose to use the well-established BRAVO algorithm[18] for the June 2018 post-election risk-limiting audit pilot. The BRAVO algorithm specifies a procedure for conducting a ballot-polling risk-limiting audit, including a "stopping rule" that determines when the audit results provide sufficiently strong evidence for the outcome. To apply the BRAVO algorithm to the question of whether a particular candidate, say, Parrish, won an outright majority, one can pool all the votes for all of Parrish's opponents and ask whether Parrish got more votes than the pool.

The BRAVO algorithm is an example of "frequentist" statistics. Bayesian methods, as introduced by Rivest and Shen,[19] provide a more flexible approach to auditing, but it is not known whether they are formally "risk-limiting", or how to compare their risk measures to RLA risk measures. See the supplementary materials for a simple worked example and simulation, courtesy of Ron Rivest, of a Bayesian "upset probability" calculation for the District Attorney-Public Administrator contest.

Audit Unit

A basic design choice for any tabulation audit is the "audit unit." What will be selected at random for examination – ballots, precincts, or something else? The choice of audit unit can be constrained by the process – for example, in Orange County separate sheets of single ballots could end up in separate tabulation batches, so choosing ballots as audit units was problematic. And the choice of audit unit affects the workload of the audit – generally, having fewer ballots per audit unit reduces the workload.

The audit unit in the June 2018 Orange County audit was a "ballot sheet." Most voters hand-marked paper ballots, which consisted of up to three separate sheets of paper. At the vote centers and precinct polling places, voters could choose instead to cast ballots on eSlate machines. Each eSlate machine produced a long reel-to-reel thermal paper roll containing records of all the ballots cast on that machine in the June primary election. A "ballot sheet" on the eSlate was the portion of that roll corresponding to a single complete ballot, i.e., the part of the roll reflecting the choices made and verified by one single voter.

Choosing entire ballots as the unit of the audit didn't seem practicable without introducing bias into the selections. Almost all the hand-marked paper ballots consisted of three separate ballot sheets. It would have been unduly burdensome to track how these ballot sheets should be combined into ballots without trusting the voting system, as sheets from a single ballot that found their way into separate scanning batches were not catalogued or counted together.

The need for a smaller audit unit increased the number of audit units to be examined. Each ballot contained each contest, but each ballot sheet did not. In a ballot-polling audit,

---

[18] Mark Lindeman, Philip B. Stark and Vincent S. Yates, "BRAVO: Ballot-polling Risk-limiting Audits to Verify Outcomes" published online at https://www.usenix.org/system/files/conference/evtwote12/evtwote12-final27.pdf.

[19] Ron Rivest and Emily Shen, "A Bayesian Method for Auditing Elections," Usenix Electronic Technology Workshop, 2012. https://www.usenix.org/system/files/conference/evtwote12/rivest_bayes_rev_073112.pdf

ballot sheets which do not contain the contest being audited don't help move the audit forward. As detailed in the supplementary materials, only about one in 2.78 ballot sheets contained any given contest. This is also reflected in the "fully-diluted margin" for each pair of candidates, i.e., the difference in their vote totals divided by the number of audit units from which the sample is drawn. This factor of 2.78 applied to an increase in the work of pulling audit units, but did not increase the time spent by the auditors to interpret and record the interpretation of voter marks.

Note that the statistical integrity of the audit does not depend on the grouping of ballot sheets into ballots. In other words, there is no requirement that the division of ballots into sheets be the same for all ballots. The validity of the audit requires only that, for each contest, every sheet containing that contest have the same probability of contributing to the audit calculations. Sheets not containing a particular contest contribute nothing to the BRAVO calculations for that contest.

Workload

Requiring that an audit satisfy a smaller risk limit provides stronger evidence of the correctness of the outcome, but also requires more sampling, i.e., pulling and examining more audit units. Because the audit was a pilot, we took steps to avoid examining a burdensome number of ballot sheets:
- Setting a risk limit only for contests with margins wide enough to make examination of more than a few hundred ballot sheets unlikely
- Choosing a relatively large risk limit of 20%

While the audits of the two closest contests were not required to achieve a pre-defined risk limit, these contests were "opportunistically" audited: the auditors recorded their interpretations of voters' marks in those contests, and the resulting risk level attained by the audit was measured and recorded.

Audit Timing

The timing of the pilot satisfied an important best practice for post-election tabulation audits: it was completed before the election agency finally certified the election results.

Audit Process

Preparatory Steps During Tabulation

Hand-marked paper ballot sheets were tabulated in batches. The catalog of all batches of ballot sheets, along with the number of ballot sheets in each batch, is called the "Ballot Sheet Manifest." Verifying the Ballot Sheet Manifest without relying on the computer tabulation system is an essential ingredient of the risk-limiting audit. During tabulation, in preparation for this audit, each batch was weighed on an Adams GFC 330a precision

scale calibrated to report the number of ballot sheets.[20] Weighing the batches added a negligible amount of time to the scanning process. Next, the batch was inserted into a Hart InterCivic BallotNow scanner, which produced a Scan Batch Report[21] indicating, among other information, the number of sheets in the batch and the precinct and "serial number" -- the last few digits of the identification number preprinted in the left margin of each page -- of each ballot.[22] The Scan Batch Report, because it was generated by the voting system, could not serve as a trusted element of the audit. To confirm the number of sheets given in the report, OCRV staff compared it to the number given by the precision scale. No batch differed by weight from the count of sheets in the Scan Batch Report by more than 2 sheets.[23] The information crucial for the Ballot Sheet Manifest was captured by hand into an Excel spreadsheet.[24]

---

[20] Ideally this report would include hard data about the accuracy of this scale. Anecdotally, election agency staff report that in their experience the scale is accurate to within one or two ballot sheets.

[21] See completescanbatch_pri2018.pdf at https://github.com/nealmcb/ocrla-2018p/blob/master/data/completescanbatch_pri2018.pdf

[22] Specifically, the Scan Batch Report listed the number of images created from the sheets in the batch, minus the number of "rejected" images. The difference was the number of "accepted" images, which should be exactly twice the number of physical sheets in the batch, because each sheet had information on two sides, captured in two separate images. For each image the last few digits of the preprinted number from the left margin were listed in the "Serial Number" column.

[23] Sometimes the first scale reading would differ by more than 2 sheet-weights from the first count from the Scan Batch Report. In each instance the discrepancy was resolved (e.g. not all of the sheets in the batch were scanned) and the resulting reweighing and rescanning matched to within 2 sheet-weights.

[24] A sample spreadsheet with the appropriate format is batch_collection_rla.xlsx at https://github.com/nealmcb/ocrla-2018p/blob/master/data/batch_collection_rla.xlsx

Each batch was then placed on a table along with its Scan Batch Report. The batches were arranged on the table in a manner that would allow easy retrieval of the randomly-selected ballots.

Ballot Sheet Manifest

The official Ballot Sheet Manifest, a comma-separated-variable (csv) computer text file containing a catalog of all ballot sheets cast in the election, is available for download from the Orange County Registrar of Voters website.[25] It consists of a list of 7096 batches, along with the number of sheets in each batch. There were 1,447,871 ballot sheets in the manifest. There were three types of batches listed in the Ballot Sheet Manifest, corresponding to the three types of voting available to Orange County voters:
- Hand-Marked Paper Ballots (Vote-by-Mail, voted at Vote Center or voted at Precinct Polling Place)
  - 4,559 batches of paper ballot sheets
  - Mean: 268 sheets. median: 282 sheets
- Early voting on eSlates at Vote Center
  - 1,561 eSlate rolls
  - Mean: 2.7 sheets (ballots), median: 1 sheet (ballot)
- Voting on eSlates at Precinct Polling Place
  - 976 precinct polling places
  - Mean: 227.2 sheets (ballots), median 220 sheets (ballots)

---

[25] https://www.ocvote.com/fileadmin/live/pri2018/rla/Phase2/combined-manifest.csv
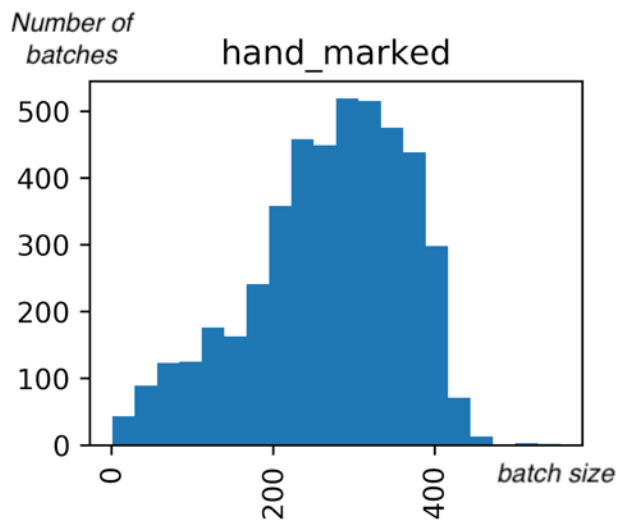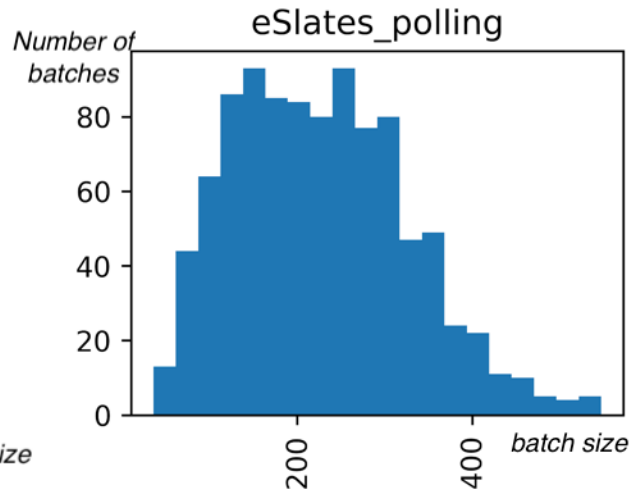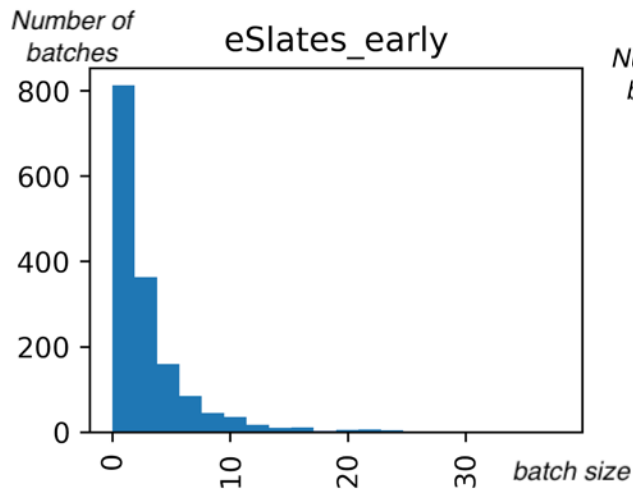
A histogram showing number of batches vs batch size for each batch type.

Rejected provisional ballots were handled differently depending on whether they were on hand-marked paper ballots or eSlates. Rejected provisional hand-marked paper ballots were not included in the manifest. Rejected provisional ballots recorded on the eSlates were included in the manifest, to make the ballot-counting process less cumbersome for the auditors.[26] Since rejected provisional ballots contain no legitimate votes for any candidate, the audit algorithm allows their inclusion, as long as they are treated as blank ballots during the audit.

Full integrity of a post-election audit of the tabulation requires that the ballot manifest be created or confirmed independently of the voting system being audited. Orange County

---

[26] When a person casts a provisional ballot on the eSlate computer system, the corresponding portion of the roll of paper is labeled "Provisional" and contains a number linking the paper record to the person who cast the ballot. The purpose is to allow inclusion or exclusion of those votes after the voter's eligibility has been checked. A side effect is that the voter's choices could be linked to the voter's identity, increasing the risk of a compromise of voter privacy.

met this requirement only for the portion of the manifest cataloguing hand-marked paper ballot sheets, and only imprecisely, as the scales used to confirm the number of sheets in each batch were expected to perform not exactly but to within one or two sheets. The system made it far too cumbersome to export all the scan batch reports to check them automatically against the scale measurements to tighten up these results.

In contrast, the manifest of the contents of the eSlate rolls was taken from the voting system. The design of the eSlate voting system did not seem to allow an efficient way to confirm its count of ballots independently, short of scrolling through thousands of rolls of thermal paper by hand, while keeping track of cast ballots vs spoiled ballots vs rejected provisional ballots. The sections devoted to each voted ballot were of varying lengths. So weighing a paper roll could not confirm the number of tabulated ballots on that paper roll.

This gap -- the lack of independent confirmation of the number of tabulated ballots on each eSlate paper roll -- significantly undercuts the strength of the evidence provided by the audit.[27]

Random Selection of Ballot Sheets

The random selection of ballot sheets was determined by a random 20-digit "seed" number, "81330464974734480366", obtained by rolling 10-sided dice in a publicly announced ceremony. That seed, along with the ballot sheet manifest, was fed into Philip Stark's Tools for Ballot-Polling Risk-Limiting Election Audits[28] web page, which generated the initial random sequence of ballots. Stark's page implements Ron Rivest's "sampler" pseudo-random number generator algorithm.[29] This way of generating a random sequence is commonly used for RLAs and is well documented on Stark's webpage, so observers could verify that the right ballot sheets were selected without relying on the software system on the web page.

---

[27] The authors are not aware of any direct recording electronic ("DRE") systems with voter-verified paper audit trails ("VVPAT") that make it practical to independently construct a manifest, and view this as a major reason to replace this legacy equipment where it still exists. In the meantime, the authors recommend the approach outlined below in section "Creating a ballot sheet manifest without relying on the voting system can be a challenge" for conducting RLAs of such systems.

[28] https://www.stat.berkeley.edu/~stark/Vote/ballotPollTools.htm

[29] https://people.csail.mit.edu/rivest/sampler.py

```
┌─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.                                      │
│ ┌───────────────────────────────────────────────────────────────────────┐    │
│ │ 1-1-1,195                                                           ▲ │    │
│ │ 1-1-2,234                                                           █ │    │
│ │ 1-1-3,147                                                           █ │    │
│ │ 1-1-4,213                                                             │    │
│ │ 1-1-5,243                                                             │    │
│ │ 1-1-6,177                                                             │    │
│ │ 1-1-7,180                                                             │    │
│ │ 1-1-8,180                                                             │    │
│ │ 1-1-9,216                                                             │    │
│ │ 1-1-10,264                                                          ▼ │    │
│ │ 1-1-11,397                                                            │    │
│ └───────────────────────────────────────────────────────────────────────┘    │
│ Ballots to look up (separated by commas):                                     │
│ ┌───────────────────────────────────────────────────────────────────────┐    │
│ │ 14293,17147,41614,173765,267145,269509,291656,380220,413627,442474,458877,547199,5889 │ │
│ │ 56,655931,766857,880924,998381,1018359,1266472,1356617                 │    │
│ └───────────────────────────────────────────────────────────────────────┘    │
│ [look up ballots]                                                             │
│ Sorted lookup table:                                                          │
│ ┌───────────────────────────────────────────────────────────────────────┐    │
│ │ sorted_number, ballot, batch_label, which_ballot_in_batch             │    │
│ │ 1, 14293, 1-1-55, 256                                                  │    │
│ │ 2, 17147, 1-1-64, 290                                                  │    │
│ │ 3, 41614, 1-1-142, 206                                                 │    │
│ │ 4, 173765, 1-2-13, 223                                                 │    │
│ │ 5, 267145, 1-2-346, 264                                                │    │
│ │ 6, 269509, 1-2-354, 291                                                │    │
│ │ 7, 291656, 1-2-442, 66                                                 │    │
│ │ 8, 380220, 1-3-173, 99                                                 │    │
│ │ 9, 413627, 1-3-288, 117                                                │    │
│ │ 10, 442474, 1-3-388, 203                                               │    │
│ │ 11, 458877, 1-3-445, 92                                                │    │
│ │ 12, 547199, 1-4-174, 66                                                │    │
│ │ 13, 588956, 1-4-323, 77                                                │    │
│ │ 14, 655931, 1-5-125, 321                                               │    │
│ │ 15, 766857, 1-6-4, 105                                                 │    │
│ │ 16, 880924, 1-6-468, 72                                                │    │
│ │ 17, 998381, 1-8-267, 53                                                │    │
│ │ 18, 1018359, 1-9-19, 150                                               │    │
│ │ 19, 1266472, 14075, 163                                                │    │
│ │ 20, 1356617, 48256, 242                                                │    │
│ └───────────────────────────────────────────────────────────────────────┘    │
└───────────────────────────────────────────────────────────────────────────────┘
```

The output of Stark's Audit Tool, in groups of 20 selections at a time, was transferred to "Tick Sheets"[30] created in Microsoft Excel and printed out on ledger-sized paper (11 x 14 inches). The Tick Sheets contained columns, one for each ballot sheet to be audited, with space for auditors to note the choices made by the voter in each of the five audited contests.

---

[30] See June 2018 - Tick Sheet- RLA.pdf at https://github.com/nealmcb/ocrla-2018p/blob/master/data/tick-sheet-sample.xls.

PRECINCT # _____
BALLOT COUNT _____
DATE: _____

SHEET 8

Statewide Direct Primary Election - 6/5/18

Risk Limiting Audit

BOARD MEMBERS (PRINT NAME) / INIT — 1, 2, 3, 4

BATCH NAME:	1-1-1	1-6-227	1-8-107	1-9-34	1-6-252	1-9-222	1-2-172	1-4-65	1-4-45	1-5-122	1-5-212	1-5-204	47251	Not Tab	Not Tab	Not Tab	Not Tab	Not Tab	Not Tab	Not Tab
NUMBER:	7	283	328	161	278	100	139	24	295	269	209	63	89	5232	39041	86413	165224	199666	201190	249007
POSITION																				
ASSESSOR - vote for 1																				
RICHARD B. RAMIREZ																				
NATHANIEL FERNANDEZ EPSTEIN																				
CLAUDE PARRISH																				
AUDITOR-CONTROLLER - vote for 1																				
TONI SMART																				
ERIC H. WOOLERY																				
CLERK-RECORDER																				
HUGH NGUYEN																				
STEVE ROCCO																				
DISTRICT ATTORNEY-PUBLIC ADMINISTRATOR - vote for 1																				
TODD SPITZER																				
TONY RACKAUCKAS																				
BRETT MURODCK																				
LEONORE ALBERT-SHERIDAN																				
SHERIFF-CORONER - vote for 1																				
DAVID C. HARRINGTON																				
DUKE NGUYEN																				
DON BARNES																				
NOTES																				

Determining the Initial Sample Size

As detailed in the supplementary materials,[31] of the three contests chosen to drive the audit, the closest margin was in the Assessor contest, where the declared outright winner, Parrish, had 37.91% more votes than the "pool" of other candidates. Given that margin, in order to gather evidence which would meet a 20% risk limit with a ballot-polling RLA, it would be necessary to sample 25 ballots which contained a valid vote in the contest. Given that there were 1,447,871 ballot sheets to sample from, and 520,078 valid votes, only one out of every 2.78 ballots would be expected to contain a valid vote. Multiplying 25 by 2.78 yields a mean sample size of 70.

But the sample size actually needed can vary widely from the mean, depending on the vote counts in the sample drawn. In simulations of the necessary sample size compared to the Average Sample Number (ASN) for a simple 2-candidate contest with no undervote (Table 1 of BRAVO), 25% of the time the audit can conclude with a sample count of just 41% of the ASN, but 10% of the time it may require a sample count of 209% of the ASN. In contests with more than a single margin in question, the variance is even larger.

Other considerations also come in to play. A larger sample size makes it more likely that additional rounds of auditing won't be necessary, and also offers the possibility of hitting the risk limit for contests that are being audited opportunistically.

---

[31] https://github.com/nealmcb/ocrla-2018p/blob/master/data/ocrla-2018p.ipynb

After reviewing the tradeoffs, the registrar chose a generous initial sample size of 160.

Pulling Ballot Sheets

Each hand-counted paper ballot sheet subject to examination for the audit was identified by the name of the batch containing the sheet and the sheet's position within the batch. For example, the third ballot sheet randomly selected for audit was the 203rd sheet in batch 1-3-388. A member of the election agency staff found the batch labeled "1-3-388" of paper ballot sheets -- this would be the 388th batch scanned on Scanner Number 3 for Database Number 1[32] -- and counted through it by hand to find the 203rd sheet[33]. The count was confirmed by comparing the last few digits of the preprinted identification number in the left margin of the paper ballot sheet to the serial number listed for the 203rd sheet on the Scan Batch Report.

Each eSlate ballot sheet subject to examination was identified by the precinct polling place number and the ballot's position within the eSlate rolls. Note that each eSlate ballot sheet was a full ballot -- for eSlates, "ballot" and "ballot sheet" are synonyms. For example, the second ballot sheet selected for audit was the 242nd eSlate ballot in Precinct Polling Place Number 48256. Each precinct polling place had several eSlate machines. Election agency staff found the paper rolls from the eSlate machines from Precinct Polling Place Number 48256 and brought them to the auditors. The paper roll from the eSlate with the lowest identification number was considered the "first roll", the one from the eSlate with the next lowest identification number was considered the "second roll," and so on. The auditors scrolled through the first roll, then the second, and so on, counting the non-spoiled ballots on the rolls by hand, until they reached the 242nd ballot on the set of paper rolls from Precinct Polling Place Number 48256. Finding ballots on the eSlate rolls was time-consuming, even with the aid of specialized scrolling boxes custom-made by the OCRV staff. Provisional ballots were identified by the word "Provisional" on the roll, and each had a code number that could be checked against a list of code numbers for rejected provisional ballots. Any provisional ballot chosen for audit had to be checked against that list to determine whether votes recorded for that ballot should be included in the tally for the audit.

All ballot sheets indicated by the random sequence were found. But the audit needs a process for circumstances in which the selected ballot sheet cannot be found in the paper batches. It can happen that the official manifest count for a batch may be larger than the actual number of paper ballots (including provisional ballots). Consider a hypothetical example where the batch count was 100, but for some reason, the number of paper ballot sheets is only 95. If the random selection calls for ballot sheet number 97 to be audited, there will be no piece of paper for the auditors to examine. Because it is impossible to

---

[32] To avoid unnecessary delay in creating English-language ballots while waiting for ballot translation into other languages, Orange County Registrar of Voters maintains two separate-labeled databases of paper ballots.

[33] This process can be made more efficient by imprinting ID numbers onto the sheets as they are printed. However, imprinting ID sheets might violate the California law prohibiting marking ballots after they are cast. Imprinting is only possible when the tabulating scanners are designed to allow imprinting.

know why the ballot sheet is missing, nor what it might have contained, the audit calculations must "assume the worst": standard RLA algorithms dictate that the missing sheet be interpreted during the audit as containing a vote for every loser (even though the rules for the contest might not allow that many valid votes).[34]
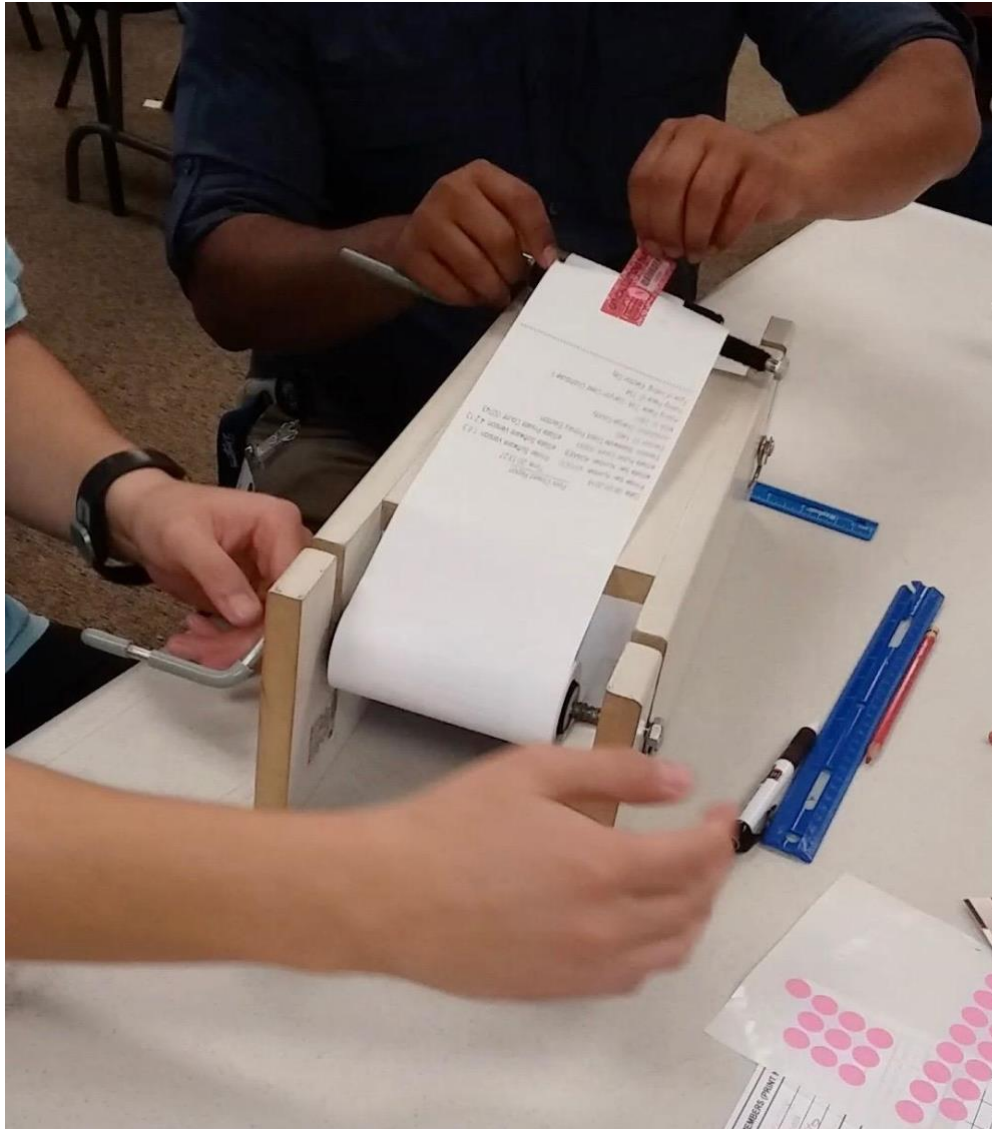
Examining Ballot Sheets & Calculating Results

For the initial sample, the auditors examined the first 160 ballot sheets in the random sequence. Each ballot sheet was examined by a team of four auditors. Two auditors on the team would look at the ballot sheet, while the other two each entered a tick mark for

---

[34] *Limiting Risk by Turning Manifest Phantoms into Evil Zombies* by Bañuelos and Stark
https://www.stat.berkeley.edu/~stark/Preprints/2ez12.pdf

each mark on the ballot onto their own paper tick sheet. Scans of the completed tick sheets, which essentially represent manual cast-vote records, are available in Neal McBurnett's GitHub repository.[35] The auditors' vote interpretations were recorded on paper to preserve the software independence of the audit.

Election agency staff, under the supervision of the auditors, entered the information from these Tick Sheets by hand into a computer database for convenient sharing and processing.[36]

BATCH 1 (WHITE?) BALLOT COUNT _____ DATE: June 28, 2018	Statewide Direct Primary Election - 6/5/18 Risk Limiting Audit		BOARD MEMBERS (PRINT NAME)	INIT
			1 Steven Morena	SM
			2 K Usmeuek	KL
			3 Roxana Castro	R
			4 Marvin Martinez	MM

BATCH NAME:	1-1-85	1-1-64	1-1-142	1-2-13	1-2-846	1-2-354	1-2-442	1-3-173	1-3-288	1-3-358	1-3-445	1-4-174	1-4-823	1-5-125	1-6-4	1-6-488	1-6-207	1-9-19	14075	48256
NUMBER:	256	290	206	223	264	291	66	99	117	203	92	66	77	321	105	72	53	150	163	242
POSITION		2	6		9	10		12	13	15	14	19	18	20	21	15	24	21	33	54
ASSESSOR - vote for 1																				
RICHARD B. RAMIREZ																			
NATHANIEL FERNANDEZ EPSTEIN																				
CLAUDE PARRISH																				
AUDITOR-CONTROLLER - vote for 1																				
TONI SMART																				
ERIC H. WOOLERY																				
CLERK-RECORDER																				
HUGH NGUYEN																				
STEVE ROCCO																				
DISTRICT ATTORNEY-PUBLIC ADMINISTRATOR - vote for 1																				
TODD SPITZER																				
TONY RACKAUCKAS																				
BRETT MURODCK																				
LEONORE ALBERT-SHERIDAN																				
SHERIFF-CORONER - vote for 1																				
DAVID C. HARRINGTON																				
DUKE NGUYEN																				
DON BARNES																				
NOTES																				

The data necessary to carry out the BRAVO calculations was exported from the database via a set of computer routines called the "rla_export tool" and imported into a computer program that produced a status report showing, for each contest:
- The number of candidates
- The number of ballot sheets examined and entered into the database
- The number of votes cast for each candidate according to the election tabulation
- The number of votes for each candidate on ballots that were examined and entered during the audit (the "selected ballot vote count")

---

[35] https://github.com/nealmcb/ocrla-2018p/blob/master/data/tick-sheets-selectionbatches-1-4.pdf and https://github.com/nealmcb/ocrla-2018p/blob/master/data/tick-sheets-selectionbatches-5-8.pdf

[36] Specifically, the computer database was hosted by Free & Fair, and the interface was the open-source RLAtool developed by Free & Fair for the State of Colorado. The version we used is at https://github.com/FreeAndFair/ColoradoRLA/. A related version from Democracy Works is at https://github.com/democracyworks/ColoradoRLA/

- The pairwise margins[37] of the audited votes
  - Each winner vs. each loser
  - Each candidate vs. the pool of all other candidates (to determine whether an outright majority was reached)
- The risk level attained for each pair so far
- The estimated number of additional ballot sheet selections needed to attain a risk limit of 20%.

For sample calculations, see the supplementary materials.[38]

Risk Levels for Each Contest

The statistical results below for each contest are based on the BRAVO algorithm, applied to the vote counts found by the auditors. These results assume that the ballot sheet manifest was correct. For each set of audited vote counts from the sampled ballots, we present the measured "risk level" attained for the contest. Any risk level of 20% or less meets the required 20% risk limit set for the audit.

Details of the calculations, and code to reproduce them, are in the supplementary materials

Auditor-Controller

Reported outcome: outright majority for Woolery. The auditors found
- 46 votes for Woolery
- 17 votes for Smart

Risk level: 1%

Assessor

Reported outcome: outright majority for Parrish. The auditors found
- 40 votes for Parrish
- 11 votes for Ramirez
- 12 votes for Epstein

Risk level: 16%

Clerk-Recorder

Reported outcome: outright majority for Nguyen. The auditors found
- 38 votes for Nguyen
- 15 votes for Rocco.

Risk level: 2%

---

[37] The margin between any two candidates (ignoring all other candidates) is called a "pairwise margin."
[38] https://github.com/nealmcb/ocrla-2018p

Sheriff-Coroner

Reported outcome: Barnes and Nguyen took top two spots, with no one winning an outright majority. The auditors found
- 28 votes for Barnes
- 23 votes for Nguyen
- 10 votes for Harrington

Risk levels
- Top two: 12%
- Barnes majority: 95%
- Nguyen majority: 28%

District Attorney-Public Administrator

The auditors found
- 22 votes for Rackauckas
- 23 votes for Spitzer
- 9 votes for Murdock
- 6 votes for Albert-Sheridan.

Risk levels
- Top two: 10%
- Rackauckas majority: 13%
- Spitzer majority: 22%

Concluding the Audit

The computed risk levels for the three contests selected for an RLA were less than the 20% risk limit after the first round. There was no need to expand the audit to an additional round of sampling.

Software Independence of the Audit

A major responsibility of conducting an evidence-based election is to provide evidence for election outcomes in a manner that is "software-independent", i.e., can be checked independently without relying on the correctness of the software involved in casting, tabulating, and reporting votes.[39] Using additional software for the audit itself introduces a responsibility to check that not just the election systems, but also any computer systems supporting the audit are doing their jobs correctly.

Here are the parts of the audit that relied on computer systems, along the mechanisms for software independence.

---

[39] See Wack, J. and R. Rivest, http://people.csail.mit.edu/rivest/RivestWack-OnTheNotionOfSoftwareIndependenceInVotingSystems.pdf

Preservation of integrity of data

Once the inputs for an audit are generated, they should be committed to with a timestamp in a durable way so that people later on can rest assured that they are working with the right data.

To accomplish this we published hashes of the reported tabulation results and the ballot sheet manifest online, and the tweets noted in the Transparency section are helpful though not really software independent. We recommend printing a full copy of the manifest, which in this case could have been printed 7 columns across, over 14 pages.[40]

Calculation of the sequence of ballot sheets to audit from the 20-digit random seed and the ballot sheet manifest

The calculation of the sequence of ballot sheets to audit from the 20-digit random seed and the ballot sheet manifest was done by entering this information into Philip Stark's Tools for Ballot-Polling Risk-Limiting Election Audits[41] web page. This page is available to anyone to use free of charge. This software implements the "sampler" algorithm[42] for selecting ballots. Software independence is achieved because the algorithm is sufficiently well-documented and widely implemented that it can be checked independently. The same list of samples are also independently reproduced in the supplementary materials, with step-by-step worked examples of the algorithm.

Calculation of risk levels from the auditors' observations

The calculation of risk levels from the auditors' observations used two kinds of software:
- Colorado Department of State RLA Tool (ColoradoRLA version 1.1.0.3, aka RLAtool, open source, hosted by Free & Fair at no cost as a courtesy to Orange County)[43]
- The Python code in our supplementary materials, open source and available to the public at https://github.com/nealmcb/ocrla-2018p

Software independence of the risk level calculations was achieved by the tick sheets -- paper records of the auditors' observations -- by checking some tallies by hand from the tick sheets, by using peer-reviewed calculation methods, by putting the auditing software online so anyone can reproduce them and check the implementations, and by checking some calculations with alternate software.

---

[40] https://github.com/nealmcb/ocrla-2018p/blob/master/data/combined-manifest.pdf
[41] https://www.stat.berkeley.edu/~stark/Vote/ballotPollTools.htm
[42] https://people.csail.mit.edu/rivest/sampler.py
[43] This software was originally created to support simultaneous comparison risk-limiting audits in Colorado counties. An open-source license makes this software available at no charge to anyone who wishes to use it.

The Pilot vs the Ideal

No one expected the pilot to meet all of the highest standards for tabulation audits. A major purpose of the pilot is to identify specific opportunities for improvement in technology, law and process.

Several aspects of the pilot audit met high standards:
- near-total reliance on voter-verified paper ballots
- transparency
- inclusion of all ballot types
- statistical design
- timing

The following point-by-point assessment of the pilot according to the standards detailed in Principles and Best Practices for Election Audits[44] should be understood to be a guide for future ameliorations, not a report card.[45]

- **Examination of Voter-Verifiable Paper Ballots** -- Good. 64% of the ballots were hand-marked paper ballots, which meet the highest standard of voter-verifiability. The 36% of ballots produced on eSlate paper rolls were acceptable, because each voter had an opportunity to review the paper record before choosing to cast the ballot. However, in the absence of strong evidence of voters taking advantage of that opportunity -- and much anecdotal evidence of voters passing up that opportunity -- systems like this (DREs with VVPATs, or Ballot Marking Devices) are less than ideal. An unknown number of non-voter-verified paper ballots were created from information faxed in by some UOCAVA voters.
- **Transparency** -- Very Good. The data to be audited (Ballot Sheet Manifest and the tabulation results) were published online in a standard format before the dice were rolled for the determination of the random seed, and this was documented and timestamped by posting SHA256 hashes to twitter.[46] Important audit processes were documented on the county's RLA web site[47] and open to public observation: the determination of the random seed and the examination of ballot sheets by the auditors. Some other processes were not open to observation: the creation of the ballot sheet manifest from the physical ballot sheets and the pulling of the ballots for audit, both of which occurred in secured areas of the election warehouse not open to the public. The risk limit, the random seed and the choices of contest from the tick sheets were available to public observers, as was the random sequence of ballot sheets.[48]

---

[44] https://www.verifiedvoting.org/wp-content/uploads/2018/11/Principles_Best-Practices_Tabulation-Audits-20181023.pdf
[45] Note that the authors of this report were involved in the design and conduct of the audit.
[46] https://twitter.com/nealmcb/status/1012379767608303616
[47] https://www.ocvote.com/results/risk-limiting-audit/
[48] See *Checking the Paper Record: A Guide for Public Oversight of Tabulation Audits* from Verified Voting at https://www.verifiedvoting.org/wp-content/uploads/2018/09/Checking-The-Paper-Record-Tabulation-Audit-Oversight-Guide.pdf.

- **Separation of Responsibilities** -- Room for improvement. The design of the audit, as well as its official evaluation, were done by the agency that conducted the election and examined the ballots, namely, the Orange County Registrar of Voters. But independent consultants (ourselves), whom they selected, assisted in the design and implementation, and then were engaged by an outside party (Verified Voting) to write up this report.
  The Principles recommend that some other entity or entities be made responsible for establishing the high-level audit policies — such as how to determine the contests and number of audit units to be audited, and how to select the particular audit units. This entity might be the legislature, an existing state agency (e.g., the Department of State or the Auditor's Office), or a new independent commission.
- **Ballot Protection** -- Not Assessed. The protection of the ballot sheets being tabulated and audited from loss, substitution, alteration or addition was not evaluated by the authors of this report.
- **Comprehensiveness** -- Mixed. The comprehensive inclusion of all ballot types in the audit met the highest standard. However, many contests were excluded a priori from auditing.
- **Appropriate Statistical Design** -- Good. The imprecision of the scale used to confirm the counts of ballot sheets, as well as the lack of independent confirmation of the counts of eSlate ballots introduced uncertainty into the ballot sheet manifest that was not reflected in the computed risk levels. Also, there was no adjustment for ballots faxed by military or overseas voters. Despite these gaps in data collection, the audit produced evidence about tabulation accuracy and demonstrated techniques of scientific assessment of that evidence while making efficient use of available resources.
- **Responsiveness to Particular Circumstances** -- Room for improvement. The audit did not have an explicit mechanism to invite input or respond to circumstances that might have come to light affecting particular devices, ballots or contests.
- **Binding on Official Outcomes** -- Good. The audit was done in time to have changed official outcomes had hand counts so indicated (which they did not). As a pilot, however, the activity had no inherent legal power to change outcomes.
- **Investigating Discrepancies and Promoting Continuous Improvement** -- Good. Unlike comparison audits, ballot-polling audits do not provide detailed information about discrepancies. Therefore, they do not provide the opportunity to investigate the cause of individual discrepancies and, thereby, improve the voting system process. But as intended, a great deal was learned in the process of the audit, as detailed in the Lessons Learned section, and the county, state and nation are well-poised to take advantage of the lessons.

Lessons Learned

Even with challenges posed by legacy voting systems, the pilot was affordable

The June 2018 pilot in Orange County showed that with a small budget, an election agency can conduct high-quality audits, collecting measurable evidence to build public confidence in elections scientifically.

A pilot audit also gives insights into the expected costs of routine audits. The cost of the routine part of the June 2018 pilot audit in Orange County -- the salary and wage cost for the work described in the Audit Process section of this report -- was under $4,000. This figure includes neither indirect costs nor benefits. Nor does the figure include constructing a manifest that is independent of the voting system, nor the work of designing and planning the audit, which we expect to quickly become much easier as more guides and capable open-source software comes on line.

For legacy voting systems, starting with flexible pilots using either ballot polling or batch comparison is a good approach

The best approach to auditing in a particular situation can depend on several factors, especially with legacy equipment. In the case of Orange County, the inability of the Hart InterCivic Voting System 6.2.1 to produce usable CVRs ruled out the use of more efficient auditing techniques in the June 2018 Orange County pilot audit.

Ballot polling techniques work with any voting system that has a voter-verifiable paper record.
However, ballot-polling audits require a large and unpredictable sample size for close contests.

For example, in this audit, the closest contest was for the position of Sheriff-Coroner, where candidate Barnes failed to win an outright majority by a margin[49] of 1.2%. The audit results in no way threw doubt on the original tabulation outcome, but neither did they provide much evidence for confirmation. To reduce the risk to below 20% with a ballot polling audit in a race with such a small margin would require examining between 10,000 and 50,000 ballot sheets.

The exact sample size for a ballot-polling audit cannot be determined in advance, as the number of audit units the auditors must examine depends on what the auditors find along the way. If the auditors find an unusually small percentage of votes for Barnes on the first sheets they examine, the possibility that Barnes won an outright majority could be ruled out with a small sample. On the other hand, if the auditors find Barnes receiving more

---

[49] For this ballot-polling audit the fractional margin is calculated as the difference between Barnes's vote count and the sum of the vote counts of all Barnes's opponents, divided by the number of votes for any candidate.

than expected, they would need to sample many more ballot sheets to achieve the risk limit.

Despite these drawbacks, ballot-polling audit pilots can sometimes be a good option for agencies whose existing technology, like that of Orange county, precludes ballot-level comparison techniques. But they can be problematic if all the contests are likely to be close. To make ballot polling pilots manageable, avoid setting risk limits for contests with small margins of victory.

A third technique, "batch comparison" is usually well worth consideration. In risk-limiting batch comparison audits, auditors collect and measure evidence by hand-counting batches of ballots (e.g., by precinct) and comparing the counts to counts reported by the voting system. Batch comparison auditing techniques provide risk reduction and valuable quality control feedback with a fairly predictable workload, regardless of the risk limit chosen. But these are not always feasible with a legacy voting system, and they are less efficient than ballot-level comparison audits.

Even with legacy voting systems, risk-limiting audit pilots are worthwhile

A pilot audit gives insights into implementation details, which depend on an election agency's technology, existing practices and legal context, which vary from jurisdiction to jurisdiction. Orange County Registrar of Voters Neal Kelley reports that the June 2018 pilot audit helped him "understand the process and logistics and work through these in advance of potential expanded RLAs."

The news media has repeatedly praised Orange County's risk-limiting audits as a positive step for election security.[50]

With better voting systems, risk-limiting audits will be even more efficient and effective

The hardest parts of the design and implementation of the Orange County pilot audit in 2018 were due to features of Orange County's existing tabulation system, which was not designed with auditing in mind. There are several voting systems already on the market that make risk-limiting audits much easier, such as the systems used by almost all counties in Colorado, where risk-limiting audits are routine and required by law.

---

[50] "California doesn't need better voting machines — it needs better audits, experts say," Peninsula Press, 11/8/2018, https://blog.sfgate.com/inthepeninsula/2018/11/08/california-doesnt-need-better-voting-machines-it-needs-better-audits-experts-say/ downloaded 11/9/2018 ; "Security researchers, voting vendors clash anew," Politico, 10/29/2018, https://www.politico.com/newsletters/morning-cybersecurity/2018/10/29/security-researchers-voting-vendors-clash-anew-393221 downloaded 11/9/2018; "The hacking threat to the midterms is huge. And technology won't protect us," Vox, https://www.vox.com/2018/10/25/18001684/2018-midterms-hacked-russia-election-security-voting downloaded 11/9/2018.

Let's look at what the Orange County audit would have looked like with a modern system designed for efficient auditing.

Comparison audits are more efficient because they measure error, rather than absolute proportions. A system with auditable CVRs would allow ballot-level comparison audits, which could have met a 20% risk limit in the Auditor-Controller, Assessor and Clerk-Recorder contests by auditing just 27 ballots (assuming that no discrepancies were found). All five contests could have been audited -- including the very tight Sheriff-Coroner contest which had a fully-diluted margin[51] of just 0.41% -- by examining just 816 ballot sheets (again assuming no discrepancies were found). In addition to being more efficient, comparison audits are easier to manage since the sample size is highly predictable.

Cost effective, required, routine risk-limiting audits of election tabulations are possible where voting systems provide CVRs that allow the auditors to link each electronic record of a voter's choices to a corresponding paper record of those choices.

It is possible to start a risk-limiting audit before the end of the tabulation

The audit described in the body of this document began after tabulation was complete. The original plan was to start before the tabulation was finished. The planned protocol (described in <u>Appendix: Protocol for RLA Started Before End of Tabulation</u>) required not underestimating the number of ballot sheets yet to be tabulated. Because the estimated number of ballot sheets turned out not to be large enough, this approach could not be applied. There are other recently-developed approaches based on "ticket numbers" that make the problem of
starting an audit before all ballots have been tabulated tractable even when the number of audit units cannot be reliably estimated.[52]

Good software to support many forms of risk-limiting audits is available free of charge, though some expertise is currently required for implementation

Thanks to Philip Stark, Neal McBurnett, the State of Colorado, Free & Fair and Democracy Works, much software to support risk-limiting audits has been licensed as open source, which allows it to be used, enhanced and shared by anyone free of charge. As of June 2018 some expertise was required to host, implement and adapt some of this software, but the landscape is changing quickly and for the better. We expect that hosted software pre-adapted to many of the major voting systems will be available by 2020.

---

[51] The fractional margin calculation for ballot-polling audits differs from ballot-level comparison audits. Here, for a ballot-level comparison audit, we calculate a "fully-diluted margin" by dividing the margin in votes by the total number of ballot sheets, rather than by the number of ballot sheets with a valid vote for one of the candidates being audited. For this hypothetical example we also assume that every ballot is a 3-sheet scanned ballot, further reducing the fractional margin.

[52] See the GitHub archive https://github.com/ron-rivest/consistent_sampler or associated arXiv paper https://arxiv.org/abs/1808.10016

Election officials considering audits are encouraged to contact Verified Voting[53] to see what is available.

Creating a ballot sheet manifest without relying on the voting system can be a challenge

The ballot sheet manifest used in the pilot audit in June 2018 was problematic. Hand counting ballots on several thousand eSlate paper rolls would take too long. For the pilot the manifest for these ballots was created based on the voting system's counts of ballots per precinct polling place, which is permissible only if these counts are eventually corroborated with physical evidence. The manifest for the hand-counted paper ballot sheets was also problematic: the voting system's counts were checked by physically weighing the batches, but this check was imprecise, allowing an error of up to two in the count of ballot sheets in each batch.

While independently creating a manifest directly from the physical set of paper records is desirable, there are other practical ways to create a manifest without relying on the voting system. One option for generating an independent manifest for eSlates and/or scanned ballots would be to derive the manifest from counts of voters who signed in at each vote center or precinct polling place. This would fit in well with the ballot reconciliation audits which are generally included under the wider umbrella of evidence-based elections. If possible the number of voters who checked in but didn't end up casting a ballot ("fleeing voters"), and the number of provisional ballots in the polling place which were eventually rejected, should be subtracted.

Such a manifest may, in practice, not match the count of physical ballots. Having more ballots than voters should be considered evidence of ballot stuffing and be further investigated. However, if the number of voters is more than the number of ballots, the audit can still proceed, using the process described in the Pulling Ballot Sheets section.

Report Conclusions

The goal of the June 2018 audit pilot in Orange County was to collect:
- useful data on the process of conducting a risk-limiting audit
- evidence that the outcomes of the election were true and correct.
The pilot succeeded in both.

The pilot was not expected to be an ideal risk-limiting audit. After an ideal risk-limiting audit, it might be possible to conclude that the chance that an erroneous outcome of the tabulation of either the Auditor-Controller, Assessor or Clerk-Recorder contest would go undetected by the audit is less than 20%.

---

[53] One way is by email: audit@verifiedvoting.org.

As for any tabulation audit, the evidence collected in the pilot audit concerns only the tabulation process and does not purport to address any of the other crucial election processes, such as voter registration and access to the polls.

The 2018 pilot audit in Orange County did show that
- **if the eSlate counts of ballots tabulated per precinct polling place are close enough**
- **if the imprecision of the weighing of the ballot sheet batches was not excessive**
- **if no faxed UOCAVA ballot sheets were selected for examination**

then the chance that an erroneous outcome of the tabulation of either the Auditor-Controller, Assessor or Clerk-Recorder contest would go undetected by the audit is less than 20%.

One of the major conclusions from this pilot is the impact of election technology systems on the conduct of risk-limiting audits. In addition to the baseline requirement that the technology provide voter-verified paper records, requiring technology that creates cast vote records linked to the paper records and supports independent creation of the manifest of ballots -- or ballot sheets -- will make routine risk-limiting audit even more efficient.

Routine auditing of a large variety of contests is an important goal to consider as election agencies upgrade equipment and procedures. In the meantime, pilot auditing does not need to be comprehensive to be valuable.

The authors of this report recommend an incremental approach, with frequent audit pilots allowing election officials discretion, along with funding for the purchase of election technology which supports efficient audits.

Acknowledgements

Appendix: Risk-limiting ballot-polling tabulation audit procedure for Orange County, CA, June 2018

Prepared by Stephanie Singer, Neal McBurnett, Justin Berardino and Neal Kelley

This document specifies all steps for the Orange County Registrar of Voters (OCRV) to conduct a ballot-polling audit.

Here are the steps for conducting the ballot-polling audit

1. Install Free & Fair RLATool and associated infrastructure
   - Will use the installation hosted by Free & Fair.
2. Run election and tabulate results
3. Choose the contests to audit
   a. Registrar of Voters will choose contests to audit.
      - Note: in an ideal audit, no one should have the power to exclude specific contests from auditing. In this pilot there are logistical obstacles due to limitations of the existing voting system. To create a mechanism to control workload, this pilot allows the Registrar to exclude contests.
4. Choose risk limit
   a. Registrar of Voters will choose risk limit.
      - Note: in an ideal audit, the risk limit would be chosen before the election. In this pilot there are logistical obstacles due to limitations of the existing voting system. To create a mechanism to control workload, this pilot allows a late choice of the risk limit.
5. Choose maximum number of ballot sheets to examine at random (Necessary for pilot audit because of software constraints)
   a. At most 1,000 ballot sheets will be examined at random.
      - A "ballot sheet" is a single piece of paper that is part of the ballot. In Colorado these are called "ballot cards", and the RLAtool uses the "card" terminology.
6. Choose number of ballot-interpretation-and-data-entry-teams ("Audit Boards")
   a. There will be four teams.
      - Each real-life OCRV Audit Board corresponds to a Colorado county in RLAtool.
7. Create Ballot-Sheet Manifest file
   a. For scanned paper ballots, independence from the Hart system will be achieved by verifying the number of sheets of paper in each batch. A precision scale will be used during the scanning process to verify the number of sheets in the batch are correct.
      - As the County scans a batch of ballots, we will verify that the total number of sheets in the batch matches a count calculated by a precision scale.

- ■ This will enable us to independently verify that the reports of the number of sheets off the voting system are accurate, and can be used to identify ballot sheets that need to be pulled.
- ■ If the number of ballot sheets in a batch as reported by the scanner is not within 2 of the number indicated by the weight of the batch as measured by the precision scale, steps should be taken to ascertain the true number of ballot sheets in the batch (without relying solely on the scanner or any other part of the voting system). The cause of the error must be investigated and resolved.

  Record the number of ballots in each batch and the storage location of the batch in a file in the Ballot Manifest format described in Appendix A.

- b. For ballots cast on eSlate machines, use the Polling Place Status report from the tally system (using only the eSlate database for input) to create a csv file in the Ballot Manifest format described in Appendix A. Here each polling place is a batch, and the location of the batch should be the location of the paper rolls containing the verified records of voter intent

- c. For each polling place, specify an ordering of the eSlate machines, to define which ballot on which eSlate roll corresponds to a given random selection within the polling place.

- d. Count the number of ballots on each eSlate roll. If desired, the counting can be done after random selections are made (to minimize number of rolls to count).

8. Compare contest_table.txt, ballot manifest, published election results reports, legal election definition, etc. to ensure that everything is consistent.
   - ■ Some can be computerized comparisons
   - ■ Some will be human comparisons (e.g., legal election definition compared with contest_table)

9. Create the MockCVR.csv file from contest_table.txt and selection of contests to be audited. Create the MockBallotManifest.csv file
   - ■ The RLA Tool requires each "county" to upload a recognizable (mock) "CVR file" and (mock) "Ballot Manifest file".
   - ■ The format of MockCVR.csv is specified by the RLAtool to match the export format from Colorado's Dominion system. Sample files are provided)
   - ■ Note that because we are using software (RLAtool) for a situation it does not yet support (ballot polling audits for one county with multiple Audit Boards) the mock CVR and Ballot Manifest files might not have all the features one would expect from files of real cast vote records and real ballot manifests.
   - ■ The number of records in MockCVR.csv does not have to match the number of ballots in the election, nor the number of ballots indicated in MockBallotManifest.csv. There should be at least 1,000 records, enough to ensure that the RLAtool will be able to record information from every ballot sheet chosen in the random phase.

- The (mock) cast vote records in MockCVR (i.e., the lines of zeroes and ones corresponding to choices within contests) do not have to correspond to actual choices by actual voters.
- Vote totals calculated from MockCVR.csv do not have to correspond to actual, or even realistic, vote totals.
- In this ballot-polling application of the RLAtool, the MockCVR.csv file affects:
  - The contests and choices presented to the Audit Boards
  - The total number of ballots that can be entered in the course of the ballot polling audit.
- The total number of cast vote records in MockCVR.csv gives a hard maximum for the number of entries the Audit Boards can make. If there were only 1,000 (mock) cast vote records in MockCVR.csv, then the RLAtool would not support the entry of more than 1,000 ballots by the Audit Boards.
- Format specification of contest_table.txt is in Appendix A (Note that the character set is UTF-8).
  b. Neal McBurnett's parse_hart.py Python script generates a MockCVR.csv from a contest_table.csv. For the MockBallotManifest.csv, we can use the default testing ballot manifest in ColoradoRLA.

10. Create a 20-digit random seed
    a. The Registrar will convene a public meeting ___ days after Election Day and give public notice of the meeting at least seven calendar days in advance. The seed is a number consisting of 20 digits, and each digit will be selected in order by sequential rolls of a 10-sided die. The Registrar will randomly select members of the public who attend the meeting to take turns rolling the die, and designate one or more staff members to take turns rolling the die in the event that no members of the public attend the meeting. [Note: this process is copied and slightly adapted from Colorado's Election Rule 25 (8 CCR 1505-1).]
11. Initialize the election from the RLAtool SoS Dashboard
    a. See Appendix B for exact process. Enter 10% into the "Risk Limit" field in the RLAtool.
       - The value entered in the Risk Limit field, along with the (mock) contest margins calculated from MockCVR.csv in the selected contests, determine how many ballots the RLAtool will present for data entry by Audit Boards.
12. Hash and upload MockBallotManifest.csv and MockCVR files via the RLAtool for each of four counties.
    a. Neal McBurnett or Stephanie Singer will perform the hashing and upload, using Ubuntu sha256sum utility.
13. Enter audit-driving information into the RLAtool.
    a. Log into RLAtool SoS dashboard (if not already logged in)
    b. Navigate to "contests" menu
    c. Select Contests screen

- - Select one contest from each "county" for audit
    - Double-check carefully!
    - Hit "Save"
  d. Enter Random Seed screen
    - Enter 20-digit random seed chosen in public ceremony
    - Double-check carefully!
    - Hit "Save & Next"
  e. Audit definition review
    - Double-check that all data was entered as desired (if not, start over from the <u>Election Initialization step</u>; note that this will require "counties" to upload files again)
    - Hit "Launch Audit"
14. Create a file RandomSequence.csv listing the ballots in the random sequence determined by the random seed and a good pseudo-random number generator
   a. Use <u>Philip Stark's ballotPollTools.htm</u> to generate the RandomSequence.csv based on the Ballot-Sheet Manifest.
   b. The format for RandomSequence.csv is given as an example in <u>Appendix A</u>.
   c. The initial random sequence should be long enough to accommodate the initial round of the ballot-polling audit.
15. Decide how many ballot sheets ("N") the Audit Boards should interpret in the next stage of the audit.
    - Choosing the number is an art, not a science. McBurnett & Singer can provide guidance on this step.
    - We want to balance these considerations:
      - The bigger the number of ballots, the more likely you will be able to stop the audit once all ballots on all lists have been processed
      - Depending on the ballots chosen, there is a significant chance that a significantly smaller number of ballots also might satisfy the risk limit
      - Retrieval of ballots should be as convenient as possible
      - You may want to avoid having some Audit Boards waiting around for the last Audit Board to finish. Note that, depending on procedures for pulling and managing the paper ballots, it might be possible for an Audit Board that finishes its list to take on ballots originally assigned to another Audit Board.
      - The bigger the number of ballots, the longer it will take to process all ballots on all lists. But the more efficiency you'll get from potentially pulling multiple ballots per batch or container.
   b. Run Neal McBurnett's analyze_rounds.py to compute an initial mean sample size, and, for later rounds, to evaluate evidence to date and compute what the risk levels are
   c. OC Registrar of Voters uses results of tool and other considerations into account and picks actual number "N" of ballot sheets to audit for the next round

16. Using RandomSequence.csv, create four lists of ballot selections, one for each Audit Board. Note that if a ballot is duplicated in the random sequence, it should be duplicated in the list.
    a. The four lists will be created by taking four equal subsets of the list in RandomSequence.csv, unless the Registrar of Voters specifies otherwise.
        ■ Together the four lists should cover a contiguous part of the random sequence, of the length N determined in the previous step and starting with the first entry in the random sequence that has not yet been processed.
        ■ There may be duplicates in the Random Sequence; to avoid conflict (between two "counties" needing one piece of paper), lists should be disjoint.
    b. For each of the four lists, create file LocationOrder.csv of the ballots listed in order of location. Note that if a certain ballot appears multiple times on the list, then that ballot must be entered the same number of times into RLATool. So any sheet listed twice (or three times, etc.) in the list taken from RandomSequence.csv should be listed twice (or three times, etc.) in the LocationOrder.csv file.
17. Each Audit Board processes all ballots (for eSlates) and ballot sheets (for scanners) on its LocationOrder list. Any selection listed twice (or three times, etc.) on the LocationOrder list should be entered twice (or three times, etc.) into the RLAtool.
    a. Collect paper ballot sheets on the list
        ■ Give staff the batch-precinct-serial number list from the four LocationOrder.csv files on paper.
    b. Confirm that paper ballot sheets collected exactly match the list
    c. Track which ballots have been entered, and how many times.
    d. Use the Audit Board in the RLAtool to enter human interpretations of ballot marks.
        ■ See Appendix C for the exact process
        ■ After entering an interpretation the Audit Board will have the option to review and either correct or approve what the RLAtool will record. Note that there is no procedure for changing audit board ballot sheet interpretations which have been entered and approved.
18. When all Audit Boards have completed interpretation of ballots on their lists, decide next steps.
    a. Use the analyze_rounds.py program from McBurnett, which employs the BRAVO algorithm from "BRAVO: Ballot-polling Risk-limiting Audits to Verify Outcomes" by Lindeman, Stark and Yates to calculate, for each contest, whether the audit of that contest can conclude because strong enough evidence for the outcome has been obtained. (See Appendix D for a sketch of the process.)
    b. Registrar of Voters will decide which, if any, contests should be moved to a full hand count. (This decision to be made by the Registrar of Voters based on results provided by the RLA.)

19. If each contest has either been confirmed by the evidence of the audit (per the BRAVO algorithm) or moved to a full hand count, stop the audit. Otherwise go to the <u>list creation step</u>.

Protocol Appendix A: Format Specifications

1. Ballot Manifest file for input to Stark's web tool for generating the random sequence must be a csv (comma-separated variable) file with fields:

Field Name	Type	Description
batch_label	char(100)	Should include physical location of batch
batch_size	Integer	The number of ballots in the batch

Sample file contents:

BatchName,NumBallots
Scanner1-Batch4-Bin1,49
1002-Arroyo School,732
Scanner3-Batch3-Bin17,51

2. Contest_table.txt
Character set is UTF-8.

Field Name	Type	Description
Precinct_Name	Char(100)	Name of the precinct
Split_Name	Char(100)	Name of the split precinct
precinct_split Id	integer	Precinct Id as assigned in the election system
Reg_voters	integer	Register voters for the split
Ballots	integer	Ballots cast for the split
Reporting	integer	Non-zero if the precinct is reporting for election Day
Contest_id	Integer	Contest Id as defined in the election system
Contest_title	Char(100)	Contest Title
Choice_id	Integer	Candidate Id as defined in the election system
Candidate_name	Char(100)	Candidate name
Candidate_Type	Char(1)	Candidate Type
Absentee_vote	integer	Votes cast for the candidate absentee

s		
Early_votes	integer	Votes cast for the candidate early
Election_Votes	integer	Votes cast for the candidate on election day

3. RandomSequence.csv
    a. From Stark's ballot polling audit webtool the output has this form:

Field Name	Type	Description
sorted_number	Integer	index of ballot in random sequence sorted into location order,
ballot	Integer	index of ballot in ballot manifest
batch_label	char(100)	Precinct Id as assigned in the election system
which_ballot_in_batch	integer	position of ballot within its physical batch

Sample file contents:
sorted_number, ballot, batch_label, which_ballot_in_batch
1, 253, Scanner1-Batch4-Bin1, 253
2, 380, Scanner1-Batch4-Bin1, 380
3, 460, 1002-Arroyo School, 60
4, 492, 1002-Arroyo School, 92
5, 647, 1002-Arroyo School, 247
6, 670, 1002-Arroyo School, 270
7, 803, 1002-Arroyo School, 270
8, 1132, 1002-Arroyo School, 732
9, 1587, 1002-Arroyo School, 1187
10, 1705, Scanner3-Batch3-Bin17, 105
11, 1896, Scanner3-Batch3-Bin17, 296

Protocol Appendix B: Setting Election Definition in RLAtool

a. Log into RLATool SoS dashboard (if not already logged in)
b. Navigate to "Define Audit" menu
c. Administer an Audit screen

- ■ Election Date -- any
- ■ Election Type -- any
- ■ Public meeting date -- any
- ■ Risk Limit -- enter the value 10%
- ■ Double-check that RLAtool recorded 10%. This is *not* necessarily the actual risk limit for the Ballot Polling Audit.
- ■ Hit "Save". (new screen will come up)
- ■ Double-check that RLATool recorded 10% for the "mock risk limit".
- ■ Hit "Next".

Protocol Appendix C: Entering Ballot Interpretations into RLAtool

- ■ Log into RLAtool as County (if not already logged in)
- ■ Navigate to Audit Board, sign in if necessary (enter any names, choose any parties, click "Sign In" then click "Start Audit" then click "Next")
- ■ Click "Next" at the bottom of the screen.
- ■ Ballot Sheet Verification Screen
    - ● Take the next ballot from the list of ballots to examine
        - a. Note that one ballot can appear multiple times in the random sequence, in which case it will be taken multiple times.
        - b. Ignore the "current ballot" info on the RLATool screen, which in this implementation is not supposed to match the info for the ballot you are entering.
        - c. Interpret the voter's intent from the paper ballot and enter interpretations into tool
        - d. If overvote, do not enter marks for any candidate.
    - ● Click "Review"
    - ● If the review screen shows an entry error, click "Back" and correct the error). Otherwise click "Submit & Next Ballot Sheet".
    - ● If there are more ballots to examine, go to Step 1 of this section. Otherwise stop.

Protocol Appendix D: The BRAVO Stopping Calculation Details

The BRAVO procedure (adapted from "BRAVO: Ballot-polling Risk-limiting Audits to Verify Outcomes" by Lindeman, Stark and Yates):

1. Set $m = 0$ and set $T\_wl = 1$ for each winner w and each loser l . Also set s_wl to be the fraction of valid votes w was reported to have received among valid votes for w and l combined.

2. Draw a ballot uniformly at random with replacement from those cast in the contest and increment m.

3. If the ballot shows a valid vote for a reported winner w, then for each loser l that did not receive a valid vote on that ballot multiply T_wl by 2*s_wl. Repeat for all such w.

4. If the ballot shows a valid vote for a reported loser l, then for each winner w that did not receive a valid vote on that ballot multiply T_wl by 2*(1 − s_wl). Repeat for all such l.

5. For each winner w and loser l such that T_wl ≥ 1/α, declare that there is satisfactory evidence that w got more votes than l (i.e. reject the corresponding null hypothesis). Do not update such T_wl after subsequent draws.

6. If all null hypotheses have been rejected, stop the audit: The reported results stand. Otherwise, if m < 1,000, return to step 2.

7. Perform a full hand count; the results of the hand count replace the reported results.

This method limits the overall risk to α.

Run the rla_export tool to pull the requisite data from corla db; then use BRAVO. Requisite data:

  a. Vote totals for each choice in each contest from election tabulation system

  b. `all_contest_audit_details_by_cvr` output of the rla_export tool, available either as csv or as json.

   i. Note: Count each record in the` all_contest_audit_details_by_cvr` table exactly once, ignoring the "count" column. Any ballot that appeared multiple times in the random sequence will correspond to multiple records in the table. Note that the ballot IDs will also not match the random selections. But check that the total number of ballots audited is right.

2. Calculation outline:

  a. From the election tabulation system vote totals, for each contest, for each winner-loser pair (w,k), calculate s_wk = (number of votes for w)/(number of votes for w + number of votes for k)

  b. For each contest, for each winner-loser pair (w,k), set T_wk =1.

  c. For each line in `all_contest_audit_details_by_cvr` with consensus = "YES", change any T_wk values as indicated by the BRAVO algorithm.

  d. Multiply the final T_wk values by the risk limit. If for any contest the result is greater than or equal to one, that contest has met the risk limit.

Appendix: Protocol for RLA Started Before End of Tabulation

Process for Rigorous Ballot-Polling RLAs started before end of tabulation in Orange County, 2018

This process document amends the full procedure specification to indicate how to begin a ballot-polling RLA before all the ballots are tabulated and reflected in a manifest. This allows a jurisdiction to get a head-start on the audit, if desired.

We will use the term "Ballot Sheets" for both individual pieces of paper (aka "cards") in a scanned ballot, and the entire ballot for VVPATs produced by eSlates.

Note that some details of this process are specific to Orange County, CA's June 2018 RLA. Specifically, in the initial stage of the audit, Orange County created 8 "Tick Sheets" to contain the Early Random Sequence. The Tick Sheet approach is adapted from their 1% audit process. You can think of the Ballot Sheet selections on each Tick Sheet as a "round" of the audit. Tick Sheet #1 contained the first 20 elements of the Early Random Sequence in location order. Tick Sheet #2 contained elements 21-40 of the Early Random Sequence in location order, and so on.

1. Pick a date & time ("First Timepoint") to start first stage of the RLA.
   ○ The "Initial Tabulation" is the most complete tabulation available as of the First Timepoint.
2. Create "Initial Manifest" including ballot sheets tabulated and independently counted (via a precision scale) as of the First Timepoint, plus one entry for each "Anticipated Batch" of Ballot Sheets not included in the Initial Manifest. The size of each Anticipated Batch should be an upper bound on the number of Ballot Sheets that will be tabulated in the category described by the Anticipated Batch.
   ○ In Orange County, there were four Anticipated Batches of Ballot Sheets:
      ■ Early Ballot Sheets
      ■ Valid Provisional Ballot Sheets
      ■ Ballot Sheets tabulated as of the First Timepoint but not included in the Initial Manifest because the independent Ballot Sheet count data was not yet available
      ■ Other (i.e., not early nor provisional) Ballot Sheets not tabulated as of the First Timepoint.
   ○ Call the Ballot Sheets tabulated as of the First Timepoint "Initial-Manifest Ballot Sheets".
   ○ Call entries in the Anticipated Batches the "Anticipated Ballot Sheets". At the end of the entire tabulation, some of these entries will correspond to Ballot Sheets that are actually tabulated; some of these won't correspond to any Ballot Sheets at all, because the size of the batch is an upper bound, not an exact bound.

3. Publish Initial Manifest and Initial Tabulation and their hashes, and tweet the hashes to establish a publication timestamp.
4. Create a random seed ("Initial Seed") as described in the full procedure.
5. Use Initial Tabulation for approximate margins and sample size
6. Choose an initial sample size
   ○ The goal here is to front-load work so that we can finish the audit as quickly as possible after the final tabulation.
   ○ Note that no ballots should be entered into the RLAtool unless the county commits to tracking down and entering all ballots which precede it in the random sequence. Otherwise, it would take additional software work to separate out a valid complete initial sequence ("prefix") from the others. This is because valid RLA stopping rule calculations cannot include ballots which are out-of-order (i.e. they can't be "cherrypicked").
   ○ Use the results of analyze_rounds.py, multiplied by the anticipated average number of audit units per ballot, as input to this decision.
   ○ In Orange County in June 2018, the initial sample size was chosen to be 160.
7. Create "Early Random Sequence" from Initial Seed, Initial Manifest and initial sample size. Note that selection is with replacement, and duplicates need to be tracked from this overall master sequence forward
8. Audit boards find and enter Initial-Manifest Ballot Sheets in Early Random Sequence, (taking no action on Anticipated Ballot Sheets).
9. Complete tabulation of all valid ballots.
10. Create "Late Manifest" including all tabulated Ballot Sheets corresponding to the Anticipated Batches of the Initial Manifest, and only those Ballot Sheets.
    ○ Check that the numbers of Ballot Sheets in each category does not exceed the sizes chosen for the corresponding Anticipated Batches.
11. Publish Late Manifest and results of complete tabulation. Calculate final margins
12. Create a new random seed ("Second Random Seed")
13. Open Philip Stark's web tool. For "Ballots cast in all contests", enter the correct ballot sheet count from just the Late Manifest. Also enter the Late Manifest, and the Second Random Seed. Neal McBurnett will go thru these steps with you remotely to make sure things go smoothly.
14. For Orange County: Go through the 8 Tick Sheets in order and do the following steps for each:
    ○ Some of the entries on the Tick Sheet will not correspond to either Late- or Initial-Manifest Ballot Sheets, but rather to invalid provisionals or nonexistent ballots, because the Anticipated Batch sizes were upper bounds. Ignore these entries in the Early Random Sequence. Specifically, for each Anticipated Batch:
      1. Figure out the number of Ballot Sheets in that category that are were tabulated (e.g., suppose there were 30,000 valid provisional ballot sheets...)
      2. Cross off any column on the Tick Sheet that corresponds to a ballot beyond that final number. (E.g., if there are actually 30,000 provisional ballot sheets, and a column on the Tick

Sheet designates the 30,001st provisional ballot sheet, just cross off that column.)

- ○ Count the number of remaining columns on the Tick Sheet. These are columns that were neither filled in in the first stage of the audit nor crossed off. Call the number of remaining columns "N".
- ○ Use Philip Stark's web tool to find the next "N" columns in the Late Random Sequence. If the tool picks a duplicate of any ballot sheet already audited in this Late Selection, you'll need to pick another single ballot with the tool, since this time we're doing sampling without replacement. (Technically speaking, the Late Selection without replacement is a way of implementing a random permutation of the Late-Manifest Ballot Sheets.)
- ○ Have the audit boards examine these Ballot Sheets and enter their interpretations (aka Audit CVRs, or ACVRs) into the remaining columns on the Tick Sheet. [NB: the type (early, provisional, not on initial manifest, not tabulated) and number printed at the top of the tick sheet column will *not* match the type and number of the new random sequence. Just cross off the preprinted information and hand-enter the new information from the Late Selection.]
- ○ Document these steps by scanning the tick sheet for publication on the web site, and entering summary data from it as a row in a new Tick Sheet Summary spreadsheet, including the tick sheet number and the number of initial, late, and invalid ballot sheets on the Tick Sheet. Include this spreadsheet in a folder with the scans of the original 8 Tick Sheets and publish this folder on the website.
- ○ When any of the original 8 Tick Sheets is completed, enter any new information from that Tick Sheet into the RLA tool.
  - ■ Do *not* re-enter information from the June 11-15 stage of the audit.
  - ■ Do *not* enter anything for the crossed-off columns.
  - ■ *Do* make sure to use the right County (Adams, Alamosa, etc.) login for the Tick Sheet. (AnnaSophia knows which Tick Sheet gets entered into which County).
  - ■ Also for the second Audit Board Member in the RLA Tool, enter, "Late Sheet" as the first name, and the tick sheet number as the last name, so the rla_export tool can reconstruct the ordering of the selections, provide insights on progress during the audit and contrast with other possible approaches.
15. Use final tabulation and results to date to calculate whether the risk limit has been met. If not, proceed with the audit as normal, escalating further as necessary.