Suggestions for CO 2018 election audits

By Mark Lindeman, Ronald L. Rivest, Philip B. Stark, and Neal McBurnett

Overview

This note documents thoughts and suggestions regarding goals and approaches for the CO 2018 election audits, resulting from the 1/4/18 CORLA Representative Group teleconference, and subsequent communication among the present authors. (A formal statistical discussion of some of the issues touched upon here is being prepared, for future use in the detailed specification.)

Most of these suggestions focus on software changes and enhancements. The note is not intended to be comprehensive. As suggested in the 2018 RFI, it seems helpful to distinguish: (1) minimum software and procedural changes needed by June in order to meet what we see as basic goals of the audit; (2) additional changes that we see as highly desirable before November. A separate document titled "SAWG CO RLA Software Recommendations for November Elections" provides further suggestions for November.

Until a better term is found, we herein refer to risk-limiting audits of contests for statewide office as "statewide RLAs," since they may appear on at least some ballots in every county. Most of the changes needed for June are necessary for statewide RLAs. Additional changes are needed to support RLAs of smaller multi-county contests and to efficiently audit contests that do not appear on all ballots in a county. These would be useful for June also, especially since in party primaries, no contest appears on all ballots, but that inefficiency is considered tolerable, especially given the intention to perform RLAs of contests for both major parties.

Introduction: basic audit procedure for June

Probably the simplest way to complete statewide RLAs in June is a stratified audit that combines a ballot-level comparison audit in counties that can conduct them ("CVR counties") with a ballot polling audit in the few – anticipated two – "legacy counties" that cannot. (Our understanding is that the four counties that hand-count their ballots will not be required to audit those counts. Treating those results as authoritative, the RLA calculations can readily be adjusted to take account of those "known good" subtotals.) This stratified approach might turn out to require very little or no additional auditing beyond that required to audit countywide or local contests, depending on the sizes of and margins in the other contests Secretary Williams selects for auditing. (Additional auditing is most likely if a statewide primary contest is much closer than the chosen countywide contest in at least one large county, or the audit finds many discrepancies.) To compute risk measurements and implement escalation rules will require modifications to RLATool, and some software (which could be external to RLATool) will be required to calculate which sampled ballots "count" in calculating the measured statewide risk.[1]

[1] An alternative approach would combine ballot-level comparison where possible with *batch*-level comparison in the legacy counties. This approach would require legacy counties to create batches and obtain vote counts for them, presumably by exporting cumulative vote counts after each batch and then computing the differences. We do not discuss this method further here.

Immediate changes

Procedural change: Establish and use consistent names for multi-jurisdiction contests (and their choices). As previously discussed, auditing statewide and other multi-jurisdiction contests depends on this consistency.

Procedural change: Export audit data. In order to coordinate a statewide audit with minimal changes to existing RLATool modules, the following RLATool data – available from the rla_export tool – must be made available:

1. A list of the ballots that were examined by the audit, in random selection order.
2. For each ballot examined by the audit, and for each contest on the examined ballot (or, at bare minimum, for each contest subject to a risk limit), the audit interpretation of the choice made by the voter. (This can be construed as an "audit CVR" for each audited ballot.)
3. For ballots examined as part of a comparison audit, the contents of the corresponding voting system CVR, and an indication of any discrepancies found on that ballot for each contest on that ballot (1- or 2-vote over- and understatements).
4. Details of any additional information that might be necessary to allow the audit computations to be checked by others.

We strongly believe that all these data should be made **public**, not just to the parties and to election officials. We understand that some modifications may be needed to enforce substantive anonymity, i.e., the unlinkability of a voter's name to the choices on the voter's ballot. We believe that it should be safe to publish CVRs and corresponding audit CVRs for audited ballots containing, at bare minimum, all contests for statewide office and all other contests that are subject to a risk limit (but not containing district style information).

Procedural change: revise the audit data collection from ballot-polling counties to support multiple contests. The current spreadsheet template only works for a single contest. The audit board interpretations should be provided in standard CSV format, with one ballot card per row.

Software change: Modify RLATool to allow external specification of minimum county-level sample sizes for an audit round. The RLATool should be modified to accept external input specifying the minimum number of ballots to be sampled within each county. (Currently, RLATool computes these sample sizes from county-level margins in contests subject to a risk limit.) To coordinate a statewide audit, there needs to be a way to update the minima between rounds during the audit. This modification, along with the audit data export, would allow external software to work with RLATool to perform any calculations needed to complete RLAs of contests for statewide office.

New software module. Here we describe what we see as the minimum "new" software functionalities needed to complete valid statewide RLAs in June. This module could be added to the existing RLATool or could operate external to it. This description does not constitute a formal specification, but should indicate the scope of the work.

- **Aggregate county-level data to produce sampling strata and audit inputs.** The software needs to produce statewide results from county-level results, find the total number of ballots cast in CVR counties and legacy counties (separately, making one "CVR stratum" and one "legacy stratum"), find the total number of reported votes for each candidate in each audited contest in CVR counties and legacy counties (separately), and make other straightforward modifications to the record-keeping to accommodate statewide and stratum-wide tallies. (As mentioned above, statewide tallies must take account of the results in hand-count counties.)

- **Calculate per-round sample sizes for statewide RLAs**: For a contest subject to risk limit, given that risk limit and the initial reported vote totals by county – and, if applicable, the audit results from previous rounds – compute two sample sizes. One sample size is for all CVR counties; the other is for all legacy counties. From these sample sizes, determine sample sizes for each county through simple random sampling.

  (Additional design parameters will influence the sample sizes. For instance, what seems to be the simplest method of combining comparison and ballot-polling audits entails setting a maximum acceptable error for each stratum in each contest subject to a risk limit, such that the total is less than the apparent margin.)

- **Construct stratum-wide samples using county sampling sequences.** This can be done in any of several ways, and is conceptually straightforward. One can think of a stratum-wide sample as having two stages. The first stage is just a simple random sample across all (or, after the first round, all remaining) ballots in the stratum – so larger counties contribute proportionally more ballots. In the second stage, the actual ballots to be audited in each county are found in the county's sample sequence, as previously reported by the RLATool. A county's contribution to the stratum-wide sample may be larger than the sample required to audit county contests, in which case additional ballots are retrieved and audited. Or it may be smaller, in which case some ballots are not included in the statewide RLA calculations for the current round.

- **Compute measured risk for statewide offices.** These risk levels will combine separate calculations for the CVR and legacy counties. (Published data may allow additional, perhaps sharper risk measures to be computed.)

We know that additional changes have been proposed, such as allowing multiple simultaneous audit stations per county. Clearly, all else equal, we would like all improvements to be made as soon as possible.

Changes for November (or possibly beyond)

As we have said before, a proper and efficient implementation of RLAs in Colorado requires some additional functionality:

Revise the RLATool to treat contests as first-class entities. The tool currently regards each county-contest combination as distinct: it has no idea that the U.S. Senate contest in Adams

County is the same as the U.S. Senate contest in Yuma County. Thus, the RLATool cannot appropriately coordinate an audit or even measure risk in such contests. The stratified approach described for June works for contests for statewide office, but extending it to smaller multi-county contests seems far inferior to revising the tool.

Revise the RLATool to import SCORE data as an upper bound on the number of ballots cast of each ballot style (and, thus, per contest). To efficiently audit contests that appear on a fraction (especially a small fraction) of ballots, one needs to be able to find those ballots and disregard others. The CVRs can be helpful in this task, but because the voting system produces the CVRs, they should not be treated as the sole authority on which ballots have which ballot styles. It seems feasible to use SCORE data to provide an upper bound on how many ballots may have been cast in each style (and, therefore, contest). It is then relatively easy to complete a risk-limiting audit that tests, rather than trusts, the CVRs.

Complete the transition to ballot-level comparison. Support for a "legacy stratum" is a short-term expedient which requires more auditing, with less benefit, in these counties.

Continue to expand publication of data in ways consistent with substantive anonymity. Despite extensive data publication, the November 2017 audits in some respects was a "black box" from the perspective of observers. More can and must be done to make the audit observable both in real time and in retrospect. An unobservable audit is not one that can command or warrant public confidence. This will be a topic of ongoing discussion.