Open Source Election Technology Institute Inc.
530 Lytton Avenue, 2nd Floor
Palo Alto, California 94301 USA
+1.650.600.1450
hello@osetinstitute.org

OSET INSTITUTE

TRUST THE VOTE PROJECT

Wednesday, 10.March 2021

Hon. Sen. Mae Flexer & Hon. Rep. Daniel Fox
Co-Chairs
Hon. Sen. Rob Sampson & **Hon. Rep Gale Mastrofrancesco**
Ranking Members
Committee on Government Administration and Elections
Connecticut General Assembly
Legislative Office Building, Room 2200
300 Capitol Avenue #5100
Hartford, CT 06106

RE: HB 6575 AN ACT CONCERNING RISK-LIMITING AUDITS FOR ELECTION RESULTS

May it please the Co-Chairs and Ranking Members—

My name is Philip B Stark of the University of California, Berkeley serving as a Board Advisor to the Open Source Election Technology (OSET) Institute. I have been authorized by its Board of Directors to prepare and submit written testimony on behalf of myself and the Open Source Election Technology (OSET) Institute, Inc.—a 501(c)(3) nonprofit nonpartisan election technology research and development organization headquartered in Palo Alto, CA with over 15 years experience in election system design for verifiability, accuracy, security and transparency.

I, together with Dennis Mema, OSET Election Technology Public Policy Analyst, with review by OSET Chief Legal Officer offer this testimony to the Committee for reference in its information gathering process regarding HB 6575—legislation we fully support. We have made every effort to provide accurate information to the best of my knowledge and experience.

We appreciate the invitation to submit this testimony. We hope this will help inform the Committee's work and that the Institute and I can be of volunteer service in matters of Risk Limiting Audits and related technology, given that I am the inventor of the process.

Respectfully Submitted,

Dr. Philip B. Stark
Associate Dean, Division of Mathematical and Physical Sciences,
Professor of Statistics
University of California Berkeley &
OSET Institute Board Advisor

Dennis Mema
Election Technology
Public Policy Analyst

Christine M. Santoro, Esq.
Chief Legal Officer &
Corporate Secretary

Before the
CONNECTICUT GENERAL ASSEMBLY
COMMITTEE ON GOVERNMENT ADMINISTRATION AND ELECTIONS

In the Matter of	)	PUBLIC HEARING
	)	
HB 6575	)	Wednesday, March 10th, 2021
	)	
AN ACT CONCERNING RISK LIMITING	)	10:00 am EST
	)	
AUDITS FOR ELECTION RESULTS	)	Via ZOOM Video-Conference

PUBLIC TESTIMONY SUBMISSION

THE OSET INSTITUTE'S TESTIMONY BY BOARD ADVISOR DR. PHILIP B. STARK REGARDING HB 6575 AND IMPLEMENTING RISK LIMITING AUDITS

Introduction

May it please the Co-Chairs and Ranking Members, my name is Philip B. Stark and as a Board Advisor to the Open Source Election Technology (OSET) Institute, I have been authorized by its Board of Directors to prepare this written testimony on behalf of myself and the OSET Institute—a nonprofit election technology research organization located in the Silicon Valley with over 15 years of experience in election system design for verifiability, accuracy, security and transparency. My biography is appended to this testimony for your review. Pertinent to this testimony I am the Associate Dean for the Division of Mathematical and Physical Sciences and a Professor of Statistics at the University of California Berkeley, and notably the inventor of Risk Limiting Audits.

We appreciate that the Legislature is contemplating Risk Limiting Audits to improve the verifiability of Connecticut elections. We further understand that HB 6575 is to establish a working group to determine, among other things, "the feasibility of implementing such audits; the different methods used in such audits; any potential equipment necessary to implement one or more of such methods; and the procedures necessary to implement one or more of such methods. We fully support this legislation.

Due to the Institute's subject matter expertise in election technology, and my expertise in Risk Limiting Audits in particular, we offer the following content to assist the Committee, and to accelerate the work of such a working group, assuming passage of the enabling legislation, HB 6575.

Executive Summary

1. The OSET Institute and I fully support the passage of HB 6575 to establish a working group to examine employing risk-limiting audits to determine the accuracy of election results.
2. Elections should be conducted in a way that gives the public convincing evidence that reported election outcomes are correct.
3. Such evidence can be provided by publicly verifiable *risk-limiting audits* (RLAs) of voter-verified paper ballots that a *compliance audit* has demonstrated to be trustworthy.
4. There are many ways to conduct risk-limiting audits, involving different ways of drawing samples of ballots and different demands on the voting system and on auditors.
5. All RLAs require manually inspecting voter-verified paper ballots.
6. Every jurisdiction that uses paper ballots could conduct *ballot-polling* RLAs immediately, without changing their voting equipment; the software requirements are minimal and free software is available.
7. There are more efficient methods for RLAs, in particular, *ballot-level comparison audits*, which have higher demands on the voting system and require more software, but which involve manually inspecting fewer paper ballots.
8. Audit laws should embody a number of principles, including requiring serious checks of the integrity of the paper trail, truly limiting the risk of certifying electoral outcomes that are incorrect, specifying risk limits, specifying how contests subject to RLAs are to be selected, ensuring that the audit cannot be subverted, and providing the public enough information to verify that the audit did not stop prematurely.
9. Voter-verifiable paper ballots are necessary for evidence-based elections, but "paper" alone is not sufficient: the mode of marking paper should not make it harder for voters to confirm that the paper record accurately reflects their preferences. For voters without disabilities, hand-marked paper ballots may be substantially more usable than ballot-marking devices (BMDs) for this purpose, especially if the BMD prints only a summary ballot, rather than a full-face ballot. More usability research should be conducted before such BMD systems are deployed as the primary or only method of casting a vote.
10. A publicly owned (open-source) voting system built from the ground up to be reliable, secure, and auditable would save taxpayers large sums, increase the trustworthiness of elections, make RLAs more efficient, and encourage innovation.

11. Connecticut (and the entire U.S.) could transition to evidence-based elections quickly and economically by:
    - Requiring the legal ballot of record to be voter-verified paper in every jurisdiction.
    - Requiring local election officials to produce *ballot manifests* describing how and where paper ballots (specifically, ballot cards) are stored.
    - Requiring physical inventories of ballots constructed without reliance on the voting system
    - Promulgating regulations that ensure and demonstrate that the voter- verifiable paper ballots were preserved complete and intact.
    - Developing and deploying open-source, easily auditable systems based to the extent possible on commodity hardware, and cultivating a robust, competitive market for support.
    - Immediately requiring ballot-polling RLAs of countywide and statewide contests.
    - Phasing in RLAs of smaller contests as voting equipment is replaced.

1. Evidence-Based Elections[1]

There is no perfect, infallible way to count votes. All methods–including optical scan, touchscreen, and hand counting–are subject to errors, procedural lapses, and deliberate manipulation.

If there is a trustworthy, voter-verified paper trail of voter intent, that paper trail can be used to check (and if necessary, to correct) the *electoral outcomes* of the contests in an election. *Electoral outcome* means the winning candidates or positions,[2] not the exact numerical tally.

The principle of *evidence-based elections* is that local election officials should not only find the true winner(s) of an election, they should also provide the electorate convincing evidence that they did. Generally, that means that the election must be conducted using voter-verified paper ballots; all validly cast ballots must be protected to ensure that they continue to represent the correct outcome; there must be convincing evidence that those ballots were kept inviolate through the audit; and the reported outcomes need to be checked against the paper trail by suitable audits or hand counts.

Evidence = auditability + auditing

---

1 Evidence-based elections were introduced in Stark, P.B., and D.A. Wagner, 2012. Evidence- Based Elections, *IEEE Trans. Security Privacy*, *10*, 33–41, doi.org/10.1109/MSP.2012.62 Preprint.

2 Or, for instance, whether there is a runoff.

2. Risk-Limiting Audits[3]

In most states that audit elections, statutory audits provide no assurance that, if a reported outcome is wrong, the error will be detected, much less corrected. Indeed, in some states supporting a 1% Post-Election Manual Tally (PEMT) local election officials are not required to give greater scrutiny to closer contests, or to take any particular action if the audit discovers tabulation errors. In fact, in some states existing regulations hamper the ability to discover outcome-changing errors by allowing local election officials to exclude provisionally cast ballots and vote-by-mail ballots that arrive after Election Day. Audits of this nature and type waste resources, and do not accomplish anything in particular.[4]

In contrast, a *risk-limiting audit* (RLA) is any post-election procedure that offers the following statistical guarantee:

> *If a full manual tally of the complete voter-verified paper trail would show a different electoral outcome, there is a known, pre-determined minimum chance that the procedure will lead to a full manual tally.*

If the procedure does lead to a full manual tally, the result of that tally replaces the reported result, thereby correcting it. The maximum chance that the procedure will *not* lead to a full manual tally if that tally would show a different outcome is called the *risk limit*. Equivalently, the risk limit is the largest chance that the audit will fail to correct an outcome that is incorrect, where *incorrect* means that a full manual tally of the voter-verified paper trail would find different winner(s). For instance, a RLA with a risk limit of 5% has at least a 95% chance of requiring a full manual tally, if that tally would show a different outcome.

---

[3] Risk-limiting audits have been endorsed by the Presidential Commission on Election Administration, the American Statistical Association, the League of Women Voters, Common Cause, Verified Voting Foundation, and many other organizations concerned with election integrity. They are required by law in Colorado and Rhode Island, and have been tested in California, Ohio, and Denmark. They were developed in 2007; the first publication is Stark, P.B., 2008. Conservative Statistical Post-Election Audits, *Ann. Appl. Statistics*, *2*, 550–581. Reprint. Since then, there have been extensions for other social choice functions (e.g., proportional representation, see Stark, P.B., and V. Teague, 2014. Verifiable European Elections: Risk-limiting Audits for D'Hondt and Its Relatives, *JETS: USENIX Journal of Election Technology and Systems*, *3*, 18–39. https://www.usenix.org/system/files/jets/issues/0301/overview/jets_0301_stark_update_9- 10-15.pdf), for auditing any number of contests simultaneously, for different types of voting equipment, etc. For a general but still somewhat technical introduction, see Stark, P.B., and M. Lindeman, A Gentle Introduction to Risk-Limiting Audits, *IEEE Security and Privacy*, *10*, 42–49, doi:10.1109/MSP.2012.56 Preprint

[4] I have heard many people claim that it checks whether the machines are working correctly. That assertion lacks appropriate nuance. Machines never work perfectly. The question is whether they worked well enough, in a given election, to find the true winner(s). That is the question a risk-limiting audit answers.

There are many methods for conducting risk-limiting audits. For instance, a full hand count is a risk-limiting audit, with a risk limit of zero. But by inspecting randomly selected ballots and using appropriate statistical methods, it is possible to conduct risk-limiting audits much more efficiently—when the electoral outcome is correct.[5]

3. Compliance Audit

A risk-limiting audit of an untrustworthy paper trail, or any audit that purports to ascertain voter intent from an electronic record or from an artifact that the voter did not have the opportunity to check, is "Security Theater." There is little reason to believe that a full manual tally of such records would reveal the true winner(s). It is therefore crucial to base audits on voter-verified paper records; to ensure that those records include every validly cast vote exactly once, and no others (checking the determination of eligibility, in particular); to ensure that those records remain complete and intact from the election through the audit; and to assess the evidence that they are trustworthy. Absent affirmative evidence that the paper trail is a trustworthy record of voter intent—that it accurately reflects the intent of every voter who legitimately cast a ballot in the contests under audit, and no others—the audit might simply confirm the incorrect outcome. The process of assessing the trustworthiness of the paper trail is called a *compliance audit*. Compliance audits should include the following steps, among others:

- Ballot accounting.
  - Check that the number of ballots sent to polling places equals the number returned voted plus the number returned spoiled, plus the number returned un-voted.[6]
  - Check that the number of ballots returned from each polling place does not exceed the number of voters registered at that polling place or the number of pollbook signatures at the polling place.
  - Check that the number of ballots of each style corresponds to the number of ballots of each style reported by the voting system. Ballot counts for this purpose should be based on the physical paper, not on the voting system: the audit needs external touchstones to check the voting system.
- Eligibility.

---

5 When the outcome is incorrect, the audit is intended to have a large probability of requiring a full manual tally, so it generally will not save labor then.

6 For systems that print ballots on demand, check that the paper stock (sheets cast, spoiled, and still blank) adds up to the number of sheets sent to the polling place or vote center.

- Check signature verification on vote-by-mail ballots, especially if signature verification was automated.
- Check the disposition of provisional ballots to ensure that all that were validly cast (and no others) were included in the results.
- Check that each voter received the correct ballot style based on her eligibility. For vote-by-mail ballots, there should be a record of the ballot style mailed to the voter; for in-person voting, this might require recording (e.g., in pollbooks) the ballot style given to the voter. For provisionally cast ballots, this might be more complicated.

- Physical chain of custody.
  - Record seal numbers whenever a batch of ballots is sealed.
  - Check physical seals for signs of tampering whenever a batch is unsealed.[7]
  - Use numbered seals that are hard to forge or bypass; check seal numbers against the numbers recorded when the boxes or bags were sealed; and log the result.
  - Review custody logs. Check that at least two staff accompanied the ballots whenever ballots were not locked securely and under surveillance.
  - Review surveillance video of secure ballot storage facility to ensure there was no unauthorized access to ballots.
- Due diligence regarding processes, equipment, etc.
  - Review voting equipment event logs.
  - Review any complaints made by voters or anomalies or problems noted by poll workers.

Some of these steps are formally or informally part of the canvass procedure in some jurisdictions. Ideally, the Secretary of State would require these steps (and others) to be conducted in a way that is publicly verifiable, and require jurisdictions to report the results. Protocols around physical seals and physical chain of custody are uneven at best. Before the election, there should also be scrutiny of voter registration databases and change-logs for those databases. Pre-election "logic and accuracy testing" should also include checking the usability of each ballot style by ordinary voters. We are not aware of any jurisdiction that currently checks the usability of ballots, despite high profile, high-consequence ballot usability problems (e.g., "butterfly ballots" in Florida).

---

[7] Good procedure is to photograph each seal after it is applied, check each seal against its photograph before breaking the seal (perhaps taking a second photograph to document the state of the seal), and record any discrepancies.

Compliance audits should be part of any recount, not just a precursor to risk- limiting audits. Absent a compliance audit, there might not be much reason for the public to trust that a recount will find the true winner(s).

4. Methods for Conducting Efficient Risk-Limiting Audits

The basic strategy behind current methods for risk-limiting audits is to start by acknowledging that the reported electoral outcome might be incorrect, then to examine more and more randomly selected ballots until either (a) the evidence is convincing that a full manual tally would confirm the reported outcome, or (b) there has been a full manual tally. There is more than one way to do this. Two basic approaches are *ballot-polling audits* and *comparison audits*. Both can be conducted by randomly selecting either groups of ballots (*batch-level audits*) or individual ballots (*ballot-level audits*).[8]

Ballot-polling audits are like exit polls, but instead of asking voters how they voted, they manually examine randomly selected voter-verified ballots.[9] If a sufficiently large random sample of ballots shows a sufficiently large margin in favor of the reported winner, that is evidence that the reported winner really won.[10] Ballot-polling audits have the advantage that they require very little of the voting system: just the reported winners, and access to the ballots. They do require local election officials to organize the paper trail well enough to draw a random sample of ballots.

Comparison audits compare how the voting system tallied groups of ballots to how humans tally the same physical group of ballots. A group might be, for instance, all ballots tallied in a given precinct, or by a given machine. That yields a *batch-level comparison audit*. The most efficient comparison audits use groups consisting of individual ballots: *ballot-level comparison audits*. To conduct a ballot-level comparison audit, the voting system must report how it interpreted individual ballots. Such interpretations are called *cast-vote records*. The cast-vote record for a ballot lists the voting system's interpretation of voter intent for each contest on the ballot. Some states (e.g., California) require new

---

8 Ballot-level audits tend to require examining fewer ballots in all than audits based on larger batches. Roughly speaking, the number of batches one needs to examine to confirm a contest with a given margin of victory at a given risk limit is about the same, regardless of the batch size. Hence, to attain a given risk limit, an audit that uses batches the size of precincts (say, 500 ballots per batch on average) requires examining about 500 times as many ballots as an audit that uses batches consisting of a single ballot (i.e., a ballot-level audit).

9 Unlike voters, ballots have to reply, and have to reply truthfully, so ballot-polling audits give strong statistical evidence while exit polls generally suffer from large biases.

10 How to quantify the strength of the evidence depends on how the sample is drawn, among other things.

voting systems to be amenable to ballot-level comparison audits, and several vendors now make systems that report cast-vote records that can be linked to the corresponding physical ballot.

One method for conducting a ballot-level comparison audit with a 5% risk limit requires manually inspecting approximately 7/(diluted margin) ballots, unless the audit finds errors in the cast-vote records. Let's consider California for example. The *diluted margin* is the margin of victory in votes, divided by the total number of ballot cards[11] in the population from which the sample is drawn (which must include all ballots cast in the contest, and may include others). In the 2018 California gubernatorial primary, Newsom and Cox advanced to the general election. The margin of Cox over the runner-up, Villaraigosa, was 618,215 votes out of 7,060,646 ballots cast, including under-votes. The diluted margin is thus 618,215/7,060,646 = 8.75%. A ballot-level comparison audit with a risk limit of 5% would have required inspecting approximately 7/0.0875 = 80 ballots selected at random from the entire state (assuming the audit did not find any errors). That is a trivial amount of work.

Note that ballot-polling audits and comparison audits can be combined in a single audit to take advantage of equipment capabilities (e.g., ballot-polling for precinct count optical scan (PCOS) and ballot-level comparison for Vote-by-mail and central count optical scan (CCOS)).[12]

Most ways of conducting RLAs require a *ballot manifest* describing how ballots are stored, for instance, "There are 913 boxes of ballots, numbered 1 through 913. Box 1 contains 301 ballots. Box 2 contains 199 ballots. . . . " It is reasonable to require local election officials to construct ballot manifests routinely: if an official cannot keep track of how much paper there is and where it is, they are not doing their job. Some counties might not currently organize their paper flow in a way that makes constructing ballot manifests possible; they would probably need to revise their procedures to conduct RLAs efficiently. Ballot manifests should be constructed without relying on the voting system to count the paper; otherwise, we are relying on the voting system to check itself. (Moreover, common sources of error in elections are to scan the same box of ballots twice, or to omit a box. Relying on the voting system to construct a manifest would miss such errors.)[13]

---

11 About "ballot cards:" In many elections in California, a "ballot" consists of two or more "ballot cards" that contain different contests. Sorting the physical ballot cards into homogeneous groups can greatly reduce the number of cards that must be inspected at random to yield a given number of cards that contain a particular contest—and increase the diluted margin.

12 See: Ottoboni et al. (2018) https://arxiv.org/abs/1809.04235

13 However, ballot manifests can be augmented by data from the voting system to facilitate audits, provided the audit is designed to take into account the possibility that the voting system data is incorrect. For instance, there are ways to

5. Resources Required for Risk-Limiting Audits

Ballot-polling RLAs require a ballot manifest and the reported results; the hardware and software requirements are minimal; open-source software already exists for all the computations.[14] Batch-level comparison audits using precincts as batches do not save effort compared to ballot-polling RLAs for typical margins and precinct sizes; therefore, we do not recommend them.

Ballot-level comparison audits require voting systems that can report cast-vote records for individual ballots in a way that the corresponding physical ballot can be retrieved, and vice versa. Ballot-level comparison audits also require exporting those CVRs and "committing" to them in a verifiable way. This involves more software, but Colorado (for example) commissioned open-source software that contains much of the needed functionality.[15]

The OSET Institute is developing next-generation open source RLA software based on foundational work of myself and including several capabilities unavailable elsewhere. Voting Works has an open source version of RLA software as well. Modifying any of these for Connecticut's requirements should not be expensive.[16]

Methods for combining ballot polling in some jurisdictions with ballot-level comparisons in others to produce a RLA of cross-jurisdictional contests have been developed (and software has been written to implement them)[17], but these methods have not yet been used in a real election.

RLA methods exist for plurality contests, majority contests, super-majority contests, vote-for-n contests, and proportional representation.[18] Note: SHANGRLA software supports RLA for instant-runoff voting.

---

combine cast-vote records with ballot manifests to make it easier to sample ballots that contain specific contests and still ensure that the procedure is a RLA. See, e.g., https://github.com/pbstark/SHANGRLA

14 See: e.g., https://www.stat.berkeley.edu/~stark/Vote/ballotPollTools.htm

15 Microsoft's open source SDK ElectionGuard also supports this; I believe it's been recently incorporated into Arlo an open source RLA software platform from Voting Works. This was piloted by Inyo County in November 2020.

16 Based on our tracking of the capabilities, updates, and improvements of these systems to date, we would expect that any additional modifications for Connecticut's purposes would cost on the order of $100K-$200K, possibly less by the time implementation decisions are made. Other states might be willing to share the cost of generally useful enhancements. The underlying calculations are well worked out (see, e.g., https://github.com/pbstark/SHANGRLA or https://www.stat.berkeley.edu/~stark/Vote/auditTools.htm); the main issues concern user interfaces for state and local election officials, and some changes to the data structures that the existing software currently employs.

17 Lindeman, M., N. McBurnett, K. Ottoboni, and P.B. Stark, 2018. Next Steps for the Colorado Risk-Limiting Audit (CORLA) Program, https://github.com/pbstark/CORLA18/blob/old-version1/CO18.pdf

18 For examples, see, e.g., Stark, P.B., 2009. Efficient post-election audits of multiple contests: 2009 California tests. 2009 Conference on Empirical Legal Studies. (Preprint), and Stark, P.B., and V. Teague, 2014. Verifiable European Elections:

Probably the most difficult aspect of auditing is to coordinate the actions of different jurisdictions, for contests that cross jurisdiction boundaries. A pilot risk-limiting audit of a statewide contest in Connecticut would be very instructive.

In my experience, it takes about 2 minutes to retrieve a particular randomly selected ballot and transcribe the votes for two or three contests.[19] Additional contests take on the order of 10 seconds each per ballot. The cost of conducting RLAs seems to be very small compared to the overall cost of holding an election. In Colorado for example, some local election officials report that RLAs are easier than the statutory audits that RLAs replaced, even though the previous audits were limited to tallying at most 500 ballots. There are many ways Connecticut can reduce the effort required to conduct RLAs, including printing serial numbers on ballots after they have been made anonymous (so that voters' preferences cannot be linked to their identities).

Some vendors are promoting systems that make digital images of ballots, claiming that the images make performing RLAs easier, because fewer (or no) paper ballots need to be inspected. **That is simply not true**, **as a matter of statistics**: if a risk-limiting audit relies on images of ballots, it must check that the error in making the images from the voter-verified paper ballots plus the error the system made interpreting those images to make cast-vote records is not large enough to cause the electoral outcome to be wrong. It is a mathematical fact that this requires examining at least as many physical ballots as an audit that just compares cast-vote records to a human reading of the voter-verified ballots, without relying on the digital images.[20]

6. Principles for a State's Audit Legislation

A risk-limiting audit law should satisfy a number of principles.

1. It should require rigorous physical custody of ballots, and compliance audits, as discussed above. A RLA that relies on an untrustworthy paper record accomplishes little.

Risk-limiting Audits for D'Hondt and Its Relatives, *JETS: USENIX Journal of Election Technology and Systems*, *3*, 18–39. https://www.usenix.org/system/files/jets/issues/0301/overview/jets_0301_stark_update_9- 10-15.pdf).

19 The process is much faster if serial numbers are printed on the ballots (after the voted ballot has been dissociated from the voter's identity).

20 Moreover, (i) there are demonstrations of ways that scanners inadvertently alter images in ways that would change the appearance of voter intent, including erasing votes; and (ii) one common source of large errors in election results is failing to scan a batch of ballots, or scanning the same batch twice. "Picking errors" (where a scanner picks up more than one card) and paper jams can also lead to ballots being omitted or scanned more than once. Expecting digital images to accurately reflect voter intent from every validly cast ballot, exactly once, is wishful thinking, even in the absence of subversive digital activity (hacking). Of course, hacking the scanners or the image processing software is within the technical ability of many undergraduate computer science/engineering students.

2. It should **require genuine RLAs**: the procedures and calculations should ensure that whenever an outcome is incorrect, the audit has the requisite chance of leading to a full hand count.[21] That in turn entails a number of things:
    a. **The audit must ascertain voter intent manually, directly from the human-readable marks on the paper ballots** the voters had the opportunity to verify. It is not adequate to rely on digital images of ballots, paper printed from an electronic record, barcodes, or other artifacts that are not verifiable by the voter and/or are not tamper evident; nor is it adequate to re-tabulate the votes electronically, either from images of the ballots or from the original paper. Digital images, re-printed ballots, and other computer data are not reliable records of voter intent: they can be incomplete, fabricated, or altered accidentally or maliciously, by software errors (bugs), procedural lapses, or intentional compromise (hacking). The bill should prohibit relying on such things for the determination of voter intent. Making this prohibition explicit is important because, as mentioned above, voting system vendors are marketing technology that purports to facilitate RLAs by allowing auditors to examine digital images of ballot instead of paper ballots. But relying on those images as an accurate representation of voter intent would in fact undermine RLAs. Relying on an electronic record created by the voting system to accurately reflect voter intent amounts to asking the same doctor for a second opinion (or asking a defendant whether s/he is guilty).
    b. **The audit must take all validly cast ballots into account.** If ballots are omitted from consideration, for instance, vote-by-mail ballots that did not arrive by election night or provisionally cast ballots, the audit cannot be a genuine risk-limiting audit. The simplest way to do this is to start the audit when the canvass is complete but before the results are certified. Local election officials may want to start the audit sooner, for instance, before all provisionally cast ballots have been resolved. There are a number of ways this can be done and still yield a true RLA. For example, the audit can treat ballots that have not been tallied when the audit starts as if they had the votes that cast the most doubt on the outcome based on the votes already tallied. (In a plurality contest, that would mean treating ballots that had not been adjudicated and tallied as if each ballot showed a valid vote for every losing candidate or position.) Another approach involves stratified sampling, drawing independent random

[21] We do not believe any statute should dictate methods or calculations, only principles. That makes it possible to use improved methods as they are developed and/or as voting systems are replaced.

samples from different collections of ballots after those collections are tabulated, and combining the results using statistically appropriate formulae. While these approaches enable election officials to start the audit as soon as election night, the risk calculations are more complicated, as is the "escalation" process of examining more ballots if the initial sample turns out to be insufficient to confirm the outcome of every contest. More generally, there are tradeoffs involved in choosing among RLA methods, both in logistical complexity for election officials and in the ability for the public to verify that the audit was performed correctly. But basing the audit on vote totals that omit provisionally cast ballots or vote-by-mail ballots that were not tabulated as of some date *cannot* yield a true RLA.

c. **The audit must have the ability to correct incorrect outcomes.** This might mean that the audit must take place before results are certified, or that the audit can revise already-certified results.

3. **The risk limit(s) should be in statute.** Allowing the Secretary of State or local election official to choose the risk limits may create a real or apparent conflict of interest.

4. **The statute should specify how the contests to be audited are selected.**
   a. If not every contest will be audited in every election, the selection of contests to audit should involve a random element to ensure that every contest has some chance of being selected, to ensure that a malicious opponent would not be able to predict that any particular race will not be audited.

   b. Every contest not audited with a RLA should be audited using a *risk-measuring audit* instead.[22]

   c. **The statute must apply to cross-jurisdictional contests, including statewide contests.** Because the point of an RLA is to ensure that reported contest outcomes are correct, every county involved in a particular contest must examine ballots in such a way that the overall cross-jurisdictional procedure is an RLA of that contest. Operationally, auditing

[22] Risk-measuring audits are related to risk-limiting audits, but they do not have a pre-specified minimum chance of requiring a full manual tabulation when that tabulation would show a different result. In statistical terminology, a risk-measuring audit reports a p-value for the compound hypothesis that a full count would yield a different electoral outcome, based on the audit data. Equivalently, they report the smallest value for which a risk-limiting audit conducted using that value as its risk limit would have stopped without examining more ballots.

cross-jurisdictional contests—whether by an actual RLA or a partial RLA—will entail a number of things:

i. Contest-level results (not merely jurisdictional-level results) must be known before the audit can conclude.

ii. Audits of cross-jurisdictional contests need to be coordinated in some way, so that each county knows when its portion of the audit can stop.

iii. The risk calculations need to match the way the sample is drawn. The approach that involves examining the fewest ballots (when the contest outcome is correct) is to sample directly from all ballots cast in the contest. That requires the Secretary of State or another entity to tell each jurisdiction how many ballots it needs to draw from each cross-jurisdictional contest, in light of the margin and what the audit reveals as it progresses. A rigorous audit can also be based on a stratified sample, which would de-couple the sampling in different jurisdictions. However, stratification generally requires inspecting more ballots to attain the same risk limit, and the risk calculations are more complicated.

5. **The audit sample must not be predictable before the audit starts**.
   a. The "seed" for selecting the sample must be sufficiently random (e.g., involve rolling 10-sided dice 20 times), with public participation.[23]
   b. The sample from any collection of ballots should not be selected before election officials have "committed" to the tally of those ballots. For example, nobody should be able to know whether "precinct 207" will be audited until the election official has published the tally for "precinct 207".[24]
6. **The public must be able to verify that the RLA did not stop prematurely**, not merely "observe" the RLA. Among other things, this requires election officials to:
   a. Disclose the algorithms used to select the sample, to calculate the risk, and to determine when the audit can stop.
   b. Provide public opportunity to observe the selection of the "seed" for the random sample.
   c. Provide adequate evidence that the paper trail of cast ballots is complete and intact (evidence generated in part by the *compliance audit*).

---

[23] Colorado's public ceremony is a good model. See https://youtu.be/ysG4pFFmQ-E

[24] There are examples (notably, in Cuyahoga County, OH, https://www.wired.com/2008/03/the-mysterious/) where election officials altered the tallies in precincts selected for recount after the sample was selected, to ensure that the inspection would not find any discrepancies.

d. Provide public opportunity to verify that the correct ballots were inspected during the audit.
e. Provide public opportunity to observe the voters' marks on the ballots that were inspected by the audit.[25] For some ways of conducting RLAs, the public needs additional information. For instance, with "ballot-level comparison audits," the public also needs to see the cast-vote record for each audited ballot and proof that the full set of cast-vote records yields the reported contest results.

7. Four Recommendations for Connecticut

Connecticut already requires paper ballots; thus, all that is required to conduct evidence-based elections is better procedures for safeguarding ballots, compliance audits, and risk-limiting audits. Therefore, I offer two recommendations for implementing RLAs, and two related recommendations regarding voting systems more generally to support efforts for more verifiable elections and support risk-limiting audits.

Recommendation 1.

Pass legislation requiring the Secretary of State to develop regulations for the secure, verifiable custody of voted ballots (including protocols for using seals, transporting ballots, and storing ballots), and regulations for "compliance audits" to ensure that eligibility was determined correctly and that the security of the physical ballots was maintained.

RLAs of statewide contests are reasonably economical using ballot polling, which requires little software or other infrastructure and no change to voting equipment. To conduct such RLAs would require local election jurisdictions to report final, but uncertified vote totals to the Secretary of State earlier, and would require coordinating the audits of different jurisdictions. It would make the results of those contests trustworthy, and would be a stepping-stone towards auditing all contests to a risk limit.

Recommendation 2.

Pass legislation requiring risk-limiting audits for all statewide contests and risk-measuring audits for all other contests based on the ballots selected for the risk-limiting audits. I would recommend using a risk limit no larger than 5 percent for those contests. I would expect that auditing most such contests will require examining far fewer ballots than a 1% post election manual tally (PEMT). At some point, RLAs of

[25] It is important to have published rules governing how human marks on ballots are to be interpreted in audits and recounts. For instance, if a voter makes a write-in vote for a candidate who is also listed on the ballot, is that a valid vote? If a voter marks a vote for a listed candidate and also writes in that candidate's name, is that a valid vote? If a voter marks a vote for a candidate, crosses through the mark, and marks a vote for a second candidate, is that a valid vote for the second candidate? If a voter makes a stray mark on the ballot that is distinctive enough to identify the ballot, is the ballot valid?

all contests should become mandatory, but the risk limit for smaller contests might not be the same as that for larger contests.

As Connecticut replaces its voting systems, there should be a regulatory requirement that new systems support ballot-level comparison audits. As jurisdictions acquire such systems, they can move from ballot-polling audits to ballot-level comparison audits. But note the cost of such systems at this writing may remain unnecessarily high, largely as a result of the current approach to certification by the US Election Assistance Commission (EAC) and potentially state laws.

Recommendation 3.

The OSET Institute and I believe that our nation could (and should) develop a publicly owned, open-source[26] voting system designed from the start to be accessible, secure, reliable, and auditable, and to run on commodity "off-the-shelf" hardware (modified for certain security requirements to ensure hardware integrity and trusted start-up) to the extent possible.[27] We estimate that this would cost roughly $10- $15 million to complete the development and testing of the system based on technology research and development already underway or nearing completion around the country. We further estimate $10-$20 million would be required to train local election officials to use the system and to catalyze a competitive market for commercial support.[28] Existing voting system vendors could compete to package, sell, lease, and support the systems. But taxpayers would benefit immediately: current annual maintenance costs greatly exceed the cost of developing, acquiring, and maintaining a system of this sort.

Accordingly, we recommend that the state of Connecticut pass legislation requiring Connecticut, alone or in collaboration with other states with financial assistance from the federal government, to develop an

---

26 Subject to an OSI approved open source (public) license that addresses any and all government I.T. procurement issues.

27 Most of the hardware a voting system needs can be general-purpose office equipment (high-speed scanners and commodity CPUs). Being able to buy or replace components at retail outlets (i.e., "commodity off-the-shelf" or "COTS") rather than paying premium prices for vendor-branded (OEM), special-purpose "voting" hardware would immediately save jurisdictions significant sums. To be sure, there remain requirements for deployment configuration of COTS hardware to incorporate the necessary firmware and potential "swappable" components for hardware and system security; however such still would produce significant cost-savings.

28 We can imagine using tax incentives, small business loans, and other programs to encourage entrepreneurship. These businesses would perform the role that RedHat, Oracle, IBM, VMWare, and others provide in the enterprise computing operating system software world of Linux—although an important difference is ballot casting and counting systems need to operate on local, dedicated, air-gapped hardware, not in the cloud.

accessible, auditable, secure, public technology (open source) based voting system built upon, to the extent possible, "COTS" hardware.[29]

Again, there are several such projects underway in the U.S., with varying degrees of progress, success, and shortcomings—all of which can be leveraged with much learning. However, this raises a final issue, accessibility.

Every voter should have the opportunity to vote using a method that is usable, given his or her abilities. More than one kind of usability matters: voters must be able to record their preferences accurately and easily, and to verify accurately and easily that the paper record—which state law should specify is the official ballot—reflects those preferences. Voters with dexterity or visual disabilities may need assistive technology to cast a vote. There remains appetite for using ballot-marking devices (BMDs) for all voters, and in particular, devices that do *not* print a full-face ballot but instead print "selections-only" or "summary" ballots, possibly with a barcode or QR code to assist the voting system in transcribing voter intent. Election officials like BMDs for a variety of reasons:

- No ambiguous voter marks, which occur from time to time with hand- marked paper ballots;
- Less paper to store;
- No need to pre-print ballots in general, nor anticipate how many ballots of a given style, or in a given language, will be needed;
- Risk mitigation with respect to lawsuits on behalf of voters with disabilities.

Recent recounts have shown that the number of truly ambiguous voter marks is negligible, so I discount the first concern.

However, my personal experience is that hand-marked paper is more usable by voters *without* disabilities, not only for recording intent but also for verifying that the paper ballot correctly portrays that intent. We cannot force voters to check that the paper ballot accurately reflects their preferences, but we should not field voting systems that make it unnecessarily hard for voters to check. The extra verification step of checking whether the paper accurately reflects one's choices is especially important for ballot-marking devices (compared to hand-marked ballots) because the technology might alter the vote as a result of flawed calibration, software errors, or even hacking.

---

[29] This would be accelerated if the US Election Assistance Commission adopted standards for interoperability of voting systems, and guidelines for certifying modules of systems rather than only complete voting systems—fortunately work that is slowly proceeding within the EAC and with a new Executive Director expressing an aggressive agenda for operational improvements and bolstered resources, this may increasingly become a priority for the agency.

Experiments and observations of voters have shown that relatively few voters notice errors in BMD printout, or even check.[30] Moreover, even if voters notice errors, there is no way they can prove to a poll worker that a machine malfunctioned, and there is no way to know how many votes were mis-recorded. Votes cast on BMDs should be treated as if they are not trustworthy.[31]

Recommendation 4.

Do not certify voting equipment that requires all in-person voters to use BMDs. Every voter should have the opportunity to cast a vote by hand marking a full-face paper ballot, whether the voter casts a ballot in person or by mail.

Respectfully Submitted,

Philip B. Stark
Strategic Board Advisor
OSET Institute
Palo Alto, CA.

&

Associate Dean, Division of
Mathematical and Physical Sciences,
Professor of Statistics
University of California Berkeley

Dennis Mema
Election Technology Public Policy Analyst
OSET Institute
Milford, CT

Tuesday, 09th March, 2021 6:50 AM PST

---

30 Bernhard, M., A. McDonald, H. Meng, J. Hwa, N. Bajaj, K. Chang, and J.A. Halderman, 2020. Can Voters Detect Malicious Manipulation of Ballot Marking Devices? IEEE Symposium on Security and Privacy, https://ieeexplore.ieee.org/document/9152705. DeMillo, R., R. Kadel, and M. Marks, 2018. What Voters are Asked to Verify Affects Ballot Verification: A Quantitative Analysis of Voters' Memories of Their Ballots http://dx.doi.org/10.2139/ssrn.3292208. Kortum, P., M.D. Byrne, and J. Whitmore, 2020. Voter Verification of BMD Ballots Is a Two-Part Question: Can They? Mostly, They Can. Do They? Mostly, They Don't. https://arxiv.org/abs/2003.04997. Stark, P.B., and R. Xie, 2020. Testing Cannot Tell Whether Ballot-Marking Devices Alter Election Outcomes, https://arxiv.org/abs/1908.08144

31 Appel, A.W., R. DeMillo, and P.B. Stark, 2020. Ballot-marking devices cannot assure the will of the voters. *Election Law Journal,* 19:3. https://doi.org/10.1089/elj.2019.0619. Appel, A.W. and P.B. Stark, 2020. Evidence-Based Elections: Create a Meaningful Paper Trail, Then Audit, *Georgetown Law Technology Review*, 4, 523–541 https://georgetownlawtechreview.org/wp-co ntent/uploads/2020/07/4.2-p523-541-Appel-Stark.pdf.

Appendix

About the Witness Dr. Philip B. Stark

Philip B. Stark is Professor of Statistics and Associate Dean of Mathematical and Physical Sciences at the University of California, Berkeley. His research centers on inference problems and uncertainty quantification, with applications including astrophysics, causal inference, climate modeling, cosmology, earthquake forecasts and seismic hazard, elections, endangered species, food webs, the geomagnetic field, geriatric hearing loss, information retrieval, Internet content filters, legislation and litigation, risk assessment, seismic structure of Sun and Earth, spectrum estimation, and urban foraging. He was a Presidential Young Investigator and won the UC Berkeley Chancellor's Award for Research in the Public Interest for developing methods for auditing elections, which have been incorporated into laws in California, Colorado, Oregon, Nevada, Rhode Island, Texas, Virginia, and Washington. He is a fellow of the American Statistical Association and the Institute of Physics. He currently serves on the Board of Advisors of the U.S. Election Assistance Commission, and serves on the Advisory Board to the nonprofit OSET Institute, Inc. See: https://www.stat.berkeley.edu/~stark and more specifically: https://www.stat.berkeley.edu/~stark/Vote/index.htm.