DEBRA BOWEN | SECRETARY OF STATE | STATE OF CALIFORNIA
1500 11th Street, 6th floor | Sacramento, CA 95814 | Tel (916) 653-7244 | Fax (916) 653-4620 | www.sos.ca.gov

March 1, 2011

Mr. Gregory Schmidt
Secretary of the Senate
State Capitol
Sacramento, CA 95814

Mr. E. Dotson Wilson
Chief Clerk of the Assembly
State Capitol
Sacramento, CA 95814

VIA EMAIL AND HAND DELIVERY

Dear Mr. Schmidt and Mr. Wilson:

Attached please find my office's report of the Post-Election Risk-Limiting Audit Pilot Program, which is required by AB 2023 (Saldaña), Chapter 122, Statutes of 2010. If you have any questions about this report, please feel free to contact me at (916) 653-7244.

Sincerely,

Debra Bowen

Debra Bowen
Secretary of State

DB:elg:pc:op

Attachment

cc: Ms. Diane Boyer-Vine
Legislative Counsel
The Honorable Lou Correa, Chair
Senate Elections and Constitutional Amendments Committee
The Honorable Paul Fong, Chair
Assembly Elections and Redistricting Committee

DEBRA BOWEN | SECRETARY OF STATE | STATE OF CALIFORNIA
1500 11th Street, 6th Floor | Sacramento, CA 95814 | Tel (916) 653-7244 | Fax (916) 653-4620 | www.sos.ca.gov

AB 2023 (Saldaña), Chapter 122, Statutes of 2010
Post-Election Risk-Limiting Audit Pilot Program
March 1, 2012, Report to the Legislature

Summary

The Secretary of State Post-Election Risk-Limiting Audit Pilot Program is designed to test the effectiveness and efficiency of risk-limiting audits. Audits were conducted in eight counties last year following elections held between March and November 2011. All eight audits successfully confirmed the official election results by reviewing a small number of ballots (i.e., a few dozen to a few hundred ballots) cast in each contest that was audited. By contrast, California law requires elections officials to hand tally 100% of the ballots from 1% of all precincts after each election, an exercise that typically requires hand counting thousands of ballots.

Despite the high number of ballots hand tallied for the 1% manual tally, the pilot project team's analysis showed this statutorily-mandated manual tally to be ineffective and inefficient at confirming election results. The post-election risk-limiting audits were able to confirm with 90% confidence that election results were correct after hand counting very few randomly selected ballots. By contrast, the 1% manual tally requiring 100% of ballots in 1% of the precincts to be hand counted in the same elections gave very little statistical proof that the election outcomes were correctly calculated by the voting system.

The U.S. Election Assistance Commission (EAC) awarded the Secretary of State (SOS) a $230,000 grant to help fund the pilot program. Counties that participate in the program are eligible to receive up to $5,000 to cover staffing and other costs associated with conducting post-election audits.

Twenty-one counties have volunteered to participate in the program, and one county, Orange County, volunteered to do two audits: one in 2011 and one in 2012. Therefore, the SOS plans to complete 22 audits over the two-year period.

The $230,000 two-year grant from the federal EAC helps fund:

1) Audits of election results following live elections in 20 California counties;

2) Detailed analyses of the efficacy of risk-limiting audits and recommendations on modifications needed to make current voting systems auditable; and

3) The creation of auditing tools for elections officials. The pilot program team has developed draft audit rules for selecting the initial sample size and for determining when enough ballots have been audited, methods for ballot-level audits, and user-friendly web-based tools and procedures for conducting and

reporting on risk-limiting audits. (A preliminary version is available at statistics.berkeley.edu/~stark/Vote/auditTools.htm)

Risk-limiting audits are based on modern statistical principles. The number of ballots initially reviewed in a risk-limiting audit varies based on the margin of victory. The audit can escalate – potentially to a full hand count of every ballot cast – depending on the observed differences between the hand tally and the voting system tally. Risk-limiting audits are most efficient when conducted at the "ballot level," meaning individual ballots from the entire voting jurisdiction are subject to the random draw and the audit. This contrasts with California's statutorily-mandated 1% manual tally, where precincts are subject to the random draw from across the entire election jurisdiction, rather than the ballots themselves. Put another way, risk-limiting audits generally involve hand counting fewer ballots overall, but those ballots are likely to come from across the entire voting jurisdiction, whereas the 1% manual tally generally involves far more ballots but only from specific areas of the voting jurisdiction.

Now that a variety of methods for risk-limiting audits have been tested, the SOS believes efficient and effective election auditing requires auditing at the ballot level. Therefore, the team will develop standards, procedures and tools for such audits as part of this project. This will:

1) Help California and other states develop new, more robust and informative election auditing laws,

2) Inform the design of next generation voting systems,

3) Provide election auditing best practices and procedures that can be used in California and other states using a wide variety of voting systems; and

4) Build public confidence that if there are errors in election results, those errors will be caught and corrected.

The Post-Election Risk-Limiting Audit Pilot Program will be completed in 2012. The Secretary of State will share with the Legislature the final report to the EAC on the project.

Background

Why audit?

State and federal voting system testing and certification help ensure voting systems used in the U.S. can mark and tally ballots accurately and securely, while protecting voter privacy. But front-end regulation and testing isn't enough. How can the public know whether voting systems actually got the job done right, unless the election results are audited after the election?

History has shown that election fraud is not theoretical. Computer experts have demonstrated that voting systems can be hacked. But even setting aside the chance of voter fraud or tampering, no voting system – no machine – can operate to perfection. Neither can humans. Voting machines misinterpret ballots, voters mismark ballots. Errors happen, and auditing determines whether those errors matter – in other words, whether a full hand count of the ballots would result in a different person or side winning an election.

What is a "risk-limiting" audit?

Risk-limiting audits can determine with precision how much hand counting is necessary to confirm election results with a high confidence. The pilot program audits provide 90% confidence that the audit will require a full hand count if the voting system tally was so wrong that the incorrect winner was called. (Confirming exact vote totals for each candidate or side of a measure would always require a full hand count.) The size of the initial sample of ballots depends on the margin of victory in the contest: the narrower the margin, the larger the initial sample. For a higher confidence level – for example, 99% confidence that any errors that led to the wrong winner being declared would be caught – more ballots would be audited. The audit can stop once the sample has shown with high statistical confidence that a full hand count would show the same winners that the voting system reported.

There are two kinds of risk-limiting audits: ballot-polling audits and comparison audits. A ballot-polling audit is analogous to an opinion poll. It examines randomly selected ballots until the human eye interpretation of the votes on those ballots gives sufficiently high statistical confidence that a full hand count would agree with the machine results. A ballot-polling audit does not require detailed information from the voting system.

In contrast, a comparison audit compares a human interpretation of the votes on randomly selected ballots to the voting system's interpretation of the votes on those ballots. A comparison audit continues until there is sufficiently high statistical confidence that a full hand count would agree with the machine results, despite any errors the audit has found. An "error" is any difference between the voting system's record of the votes on a ballot and a human eye interpretation of the votes on the ballot. Ballot-level comparison audits require very detailed information from the voting system: The system must be able to report how it interpreted each physical ballot. The benefit of a comparison audit over a ballot-polling audit is that it generally requires examining fewer ballots to confirm that the machine outcome is in fact correct. Among the 15 contests audited in the eight pilot audits conducted during 2011, one was a ballot-polling audit, 13 were ballot-level comparison audits, and one was a comparison audit based on a combination of machine-level results and ballot-level results.

Risk-limiting audits can lead to a full hand count of ballots to confirm whether the machine tally of ballots correctly determined the election winner. Until the sample provides sufficiently strong evidence that a full hand count would show the same outcome that the machine count reported, the audit escalates, and additional ballots are selected for hand counting. Depending on the true and reported margins and – in the case of comparison audits – the errors the audit finds, the audit can lead to a full hand count of ballots to confirm (or overturn) the machine results.

Ballot-level risk-limiting audits sample individual ballots spread across a jurisdiction – a method that is dramatically more effective and efficient than hand counting a far larger number of ballots selected by drawing entire precincts. Sampling individual ballots "stirs the pot" more effectively than sampling entire precincts, giving a larger chance of uncovering any problem that might exist, even though the sample is smaller.

California's 1% manual tally law dates back to the 1960s and requires elections officials to hand count the votes for all contests on all ballots from 1% of all precincts in each jurisdiction. In a regular election year, counties must count tens of thousands of ballots as part of the required 1% manual tally, yet doing so gives little or no statistical evidence that the voting system found the right winners. With risk-limiting audits, individual contests or groups of contests on the same ballot can be audited, and the winners confirmed, by looking at relatively few individual ballots. For the simplified version of the risk-limiting audit model, called the "super-simple" audit, which has been used for some of the audits in this pilot project, the size of the initial sample is the number 4.8 divided by the margin of victory. If elections officials expect to see errors, then a slightly larger sample is drawn:

Margin	**Equation**	**Initial Sample Size (in ballots)**
50%	4.8/.5	10
40%	4.8/.4	12
30%	4.8/.3	16
20%	4.8/.2	24
10%	4.8/.1	48
5%	4.8/.05	96
2%	4.8/.02	240
1%	4.8/.01	480
.5%	4.8/.005	960

Professor Philip B. Stark (Stark) from the University of California, Berkeley, is the lead researcher on the pilot program and has been a pioneer in developing risk-limiting election audit methods. His methodology has since been endorsed by national organizations, including Common Cause, the League of Women Voters, the American Statistical Association, Verified Voting, and other groups involved in election integrity.

Overview of the Program

For the pilot program, the Secretary of State:

1) Gathered a mix of 21 urban and rural counties, including at least one county representing each voting system in use in California, to participate;

2) Entered into a contract with the University of California (UC) to allow Stark to serve as lead researcher for the pilot program;

3) Convened an advisory panel and established a webpage for the pilot program;

4) Developed and tested draft web-based tools, instructions and methods for conducting risk-limiting audits; and

5) Conducted audits in eight counties during 2011.

Counties

The following counties have volunteered to participate in the pilot program, which will continue through 2012, thanks to EAC grant funding.

County	Election Date	Audit Date
Alameda	Nov. 8, 2011	Dec. 5, 2011
Alpine	TBD	TBD
Colusa	Nov. 6, 2012	TBD
El Dorado	Nov. 6, 2012	TBD
Humboldt	Nov. 8, 2011	Dec. 16, 2011
Madera	TBD	TBD
Marin	TBD	TBD
Merced	Nov. 8, 2011	Dec. 12, 2011
Monterey	May 3, 2011	May 6, 2011
Napa	Jun. 5, 2012	TBD
Orange	Mar. 8, 2011	Mar. 14, 2011
Sacramento	TBD	TBD
San Diego	TBD	TBD
San Francisco	TBD	TBD
San Luis Obispo	Aug. 30, 2011	Sept. 12, 2011
Santa Cruz	TBD	TBD
Stanislaus	Nov. 8, 2011	Dec. 2, 2011
Sutter	Nov. 6, 2012	TBD
Ventura	Nov. 8, 2011	Nov. 29, 2011
Yolo	TBD	TBD
Yuba	TBD	TBD

Two of the counties, Orange and San Luis Obispo, participated in pilot audits prior to the EAC grant award, so the expenses related to those counties' audits were absorbed by the counties. The remaining audits are funded through the EAC grant.

Advisory Panel

At the outset of the pilot program, the Secretary of State established an advisory panel of the following experts, advocates, and community activists in the field of election auditing and reform:

Susannah Goodman
Director, Common Cause National Campaign for Election Reform

Joseph Lorenzo Hall
Postdoctoral Research Fellow, New York University Department of Media, Culture and Communication

Mark Halvorson
Director and Founder, Citizens for Election Integrity Minnesota

Dean Logan
Registrar-Recorder/County Clerk, Los Angeles County

Hovav Shacham
Assistant Professor, University of California, San Diego, Department of Electrical Engineering and Computer Science

Pamela Smith
President, Verified Voting

Conducting Ballot-Level Risk-Limiting Audits Using a Parallel Scan and Tally

For most election audits, the results of a hand tally are compared to the results recorded by the voting system. For California's 1% manual tally, elections officials hand tally 100% of the ballots in 1% of the precincts and compare those hand tally totals to the machine-tallied totals generated by the voting system for those precincts.

For risk-limiting audits to be efficient, they must be conducted at the individual ballot level, not at the precinct level. A ballot-level ballot-polling audit makes no demands on the vote tabulation system, but requires examining more ballots than a ballot-level comparison audit. A ballot-level comparison audit compares the result tallied by the voting system for a given ballot to a hand tally of the same ballot. To conduct a risk-limiting comparison audit at the ballot level, two things are necessary: 1) the voting system must report a cast vote record (CVR) for each ballot. A CVR shows how the votes on a given ballot were actually tallied by the voting system; and 2) elections officials must be able to match a CVR to the corresponding physical ballot, which requires either marking ballots as they are tallied or keeping ballots in the same order in which they are tallied.

The pilot project team conducted a series of conference calls with voting system vendors to determine the capabilities of existing voting systems. Through these calls and discussions with participating counties, the team determined that none of the voting systems in use in California is capable of exporting CVRs that can be associated with corresponding physical ballots.

For this reason, the team conducted most of the audits for this pilot program by means of a parallel scan and tally of the votes. A parallel scan and tally is a second tally of the ballots using commercial off-the-shelf (COTS) scanners and open-source tally software, which was developed during the spring and summer of 2011 for use in the pilot program. Unlike commercial, federally certified vote-tabulation systems, the open-source software does report a CVR for every ballot.

County elections officials scanned the ballots using a COTS scanner and either marked the ballots or kept the ballots in order to permit each physical ballot to be paired with its CVR. This method allowed auditing the interpretation of individual ballots rather than auditing vote subtotals for entire precincts. Making individual ballots auditable – i.e., creating auditable "batches" of one ballot each – brings incredible efficiency, as described above. The hand counting workload for a ballot-level audit can be smaller than the workload of a precinct-level audit by a factor of 1,000 or more. Since the parallel tally for each audit showed the same results (winners and losers) as the official voting system, the audit was able to confirm the official results *transitively* (i.e., If A and B agree and B is correct, then A is also correct).

<u>Conducting "Ballot-Polling" Audits</u>

Stark developed a new risk-limiting audit model for the pilot program, called a "Ballot-Polling" audit. The advantage of the Ballot-Polling audit model is that elections officials need only the overall election results – not precinct subtotals or individual ballot results (CVRs). The Ballot-Polling audit model is useful for contests with wide margins, but becomes inordinately time consuming if the margin of victory is very small. The hand-count workload for ballot-polling audits grows rapidly as the margin shrinks. Ballot-level comparison audits, which check individual ballots against the voting system's results for the same ballots, are more efficient, but such audits require knowing how the voting system interpreted every ballot.

To perform a ballot-polling audit, physical ballots are selected at random and interpreted by the human eye. This selection continues until the sample gives sufficiently high confidence that the winners according to a full hand count would be the same as the winners according to the original machine count.

Ballot-polling audits make their own statistical assessment of who won directly from a random sample of ballots. This kind of auditing was developed and tested for the first time during the pilot program in the Monterey County audit (see below). Ballot-polling audits may be an excellent way to efficiently confirm large statewide contests or contests that cross two or more counties, if the margins of victory in the contests audited are relatively large or if the contest is so large that the counting burden in each jurisdiction is manageable. Stark is continuing to develop better methods for ballot-polling audits.

Web-Based Tools and Instructions for Elections Officials

The pilot program team, led by Stark, has honed the audit models and developed a set of web-based tools (http://statistics.berkeley.edu/~stark/Vote/auditTools.htm) and draft instructions (attached) designed for elections officials to use to conduct risk-limiting audits. The tools explain how the audits work and show the math that the tools implement, so elections officials and the public can understand risk-limiting audits. These resources are continually being refined and improved as a part of the pilot program.

Audits Conducted in 2011

The pilot program team conducted successful risk-limiting audits in eight counties during 2011. For each audit, the team worked with participating counties and voting system vendors to plan the audits. Stark traveled to all but one of the counties to provide on-site assistance to jurisdictions carrying out the audits, including performing all computations and helping with the random draw and hand count as required.

In some counties, the team successfully conducted simultaneous audits of several contests at one time. The simultaneous audits proved very efficient where the contests audited overlapped completely (or almost completely) in jurisdiction. The team found that unless there are large margins in the contests to be audited, conducting audits of each contest separately was more efficient if the jurisdictions among the contests did not overlap. Stark has developed the web tool in a way that allows elections officials to enter contest data to determine whether it is more efficient to audit contests simultaneously or separately.

1. Orange County: March 14, 2011

The first pilot audit followed an election in Orange County, California. The election was March 8, 2011, and the audit took place on March 14, 2011. The contest audited was a special election for San Clemente Measure A, Playa del Norte Commercial Development Project. There were 17,823 ballots cast, with 42.8% voting Yes and 57.2% voting No. Orange County uses Hart BallotNow v. 3.3.11 and the Hart eSlate v. 4.2.13 for polling place voting.

This audit was conducted as follows:

1) Ballots cast on eSlate direct recording electronic (DRE) voting machines were audited at the DRE level by randomly selecting eSlate DRE machines. Auditors hand counted the voter-verified paper trails (VVPATs) from the selected DREs and matched the totals against the DRE generated totals.

2) Ballots cast on paper (i.e., vote-by-mail ballots and polling place ballots cast using paper instead of a DRE) were audited at the ballot level, because the Hart system was able to generate a CVR for each paper ballot and county elections officials kept the paper ballots in the order they were fed into the Hart system.

The initial sample size was:

1) 12 randomly selected Hart eSlate machines for a total of 446 ballots;

2) 21 individual paper ballots.

Overall, 467 of the 17,823 ballots cast were manually reviewed and tallied for this audit. No errors were found, meaning the hand tally of these ballots matched the machine tally of these same ballots exactly.

The hand counting burden for this audit was relatively high. The 467 ballots hand tallied represented about 2.5% of all ballots cast. This was because part of the audit had to be conducted at the DRE machine level, not the ballot level. The Hart eSlate DREs did not have the capability to produce CVRs that could be associated with each VVPAT, so entire DRE machines had to be selected and all ballots cast on the DREs hand counted.

If all of the votes had been cast on paper, the entire audit could have been conducted at the ballot level, which would have required the hand counting of only roughly 33 ballots, about one-tenth of one percent (0.10%) of all ballots cast.

2. Monterey County: May 6, 2011

The second pilot audit followed an election in Monterey County, California. The election was May 3, 2011, and the audit took place on May 6, 2011. The contest audited was a special all-mail election for Monterey Peninsula Water Management District Director, Division 1. Monterey uses the Sequoia Optech 400-C/WinETP v. 1.12.4 voting system with the Sequoia AVC Edge Model II v. 5.0.24 for accessibility in polling places. Two candidates, Brenda Lewis and Thomas M. Mancini, were on the ballot, along with write-in candidates. There were 2,111 ballots cast in all. The reported totals were 1,353 votes for Lewis, 742 for Mancini, and 13 for various write-in candidates. The remaining 3 ballots were recorded as undervotes or overvotes, and as a result, those ballots were voided in the official count. According to the voting system results, Lewis received 64% of the valid votes, while Mancini received 35% of valid votes.

The audit was a "ballot-polling" audit (see description under "Background" above) which relied only on comparing the margin of victory in the overall election results to the margin of victory in a hand tally of a sample of randomly selected ballots.

Confirming the outcome with 90% confidence required examining 89 ballots selected at random. The ballot-polling audit was designed to ensure that if Lewis had at least 64% of the vote, there was at most a 1% chance that the audit would lead to a full hand count (i.e., 99% confidence level). The audit took about 90 minutes, including the time Stark spent explaining the audit procedure to public observers. Public observers helped roll the dice used to select ballots at random and had an opportunity during the audit to confirm that their interpretation of ballots agreed with the auditors' hand tally.

3. San Luis Obispo County: September 12, 2011

The third pilot audit followed a special election in San Luis Obispo County, California. The election was August 30, 2011, and the audit took place on September 12, 2011. Both contests on the ballot, City of San Luis Obispo Measures A-11 and B-11, were audited. The county uses Premier AccuVote-OS v. 2.0.12 with AutoMARKs for accessibility.

This was the first simultaneous risk-limiting audit of two contests. The method Stark used was the "super-simple" model (described above under "Background"), because it uses a relatively simple, easy-to-understand mathematical formula* to determine the initial sample size and confirm election outcomes. The audit involved a random sample of just 16 ballots, and was finished in one hour, confirming the winner of each measure. There were 10,689 ballots cast in the election, and the narrower of the two margins of victory in the contests was 45%.

Stark and San Luis Obispo County elections officials performed the audit in front of 10 public observers. The public was able to see, hear, and compare the manual tally results for each ballot audited against the CVR for the ballot. All 16 ballots matched the votes tallied from the CVRs, so no escalation was required.

The audit was extremely efficient because Stark conducted a ballot-level audit – i.e., selected individual ballots for the sample – instead of whole precincts. The audit confirmed the winner with 90% confidence. Public notice and observation are built into the audit process. The model calculates the size of the initial audit sample based on the margin of victory.** The smaller the margin, the larger the initial sample must be.

*The drawback to the "super-simple" model is that the formula is slightly less efficient (i.e. leads to a slightly larger sample sizes) than more complex and precise models.

**The margin used for risk-limiting audits is the "diluted margin of victory," which is the margin in votes divided by the number of <u>ballots cast</u> in the contest(s) audited. This is slightly different from the usual "margin of victory," which is the number of votes between the winner and loser divided by the number of <u>valid votes</u> cast. The diluted margin takes into account the fact that the voting system might have misinterpreted a valid vote as an undervote or overvote (an undervote or overvote corresponds to a "ballot cast" by a voter, but is not a "valid vote").

The audit stops after the initial sample if auditors find zero or only statistically insignificant differences between the voting system interpretation and a human eye interpretation of the votes on the ballots. The number of differences permissible depends on the margin and the size of the initial sample. As discussed above, risk-limiting audits can lead to a full hand count, if enough statistically significant differences are found.

4. Ventura County: November 29, 2011

The fourth pilot audit followed an election in Ventura County, California. The election was November 8, 2011, and the audit took place on November 29, 2011. The contest audited was the City of San Buenaventura City Council, for which there were three at-large seats to be filled. Ventura uses the Sequoia Optech 400-C/WinETP v. 1.12.4 voting system, the Sequoia AVC Edge Model II v. 5.0.24 for accessibility, and the Sequoia OptechInsight APX K2.10 HPX K1.42 in polling places.

This was the first multi-winner contest to be audited using the super-simple method. The audit was successful: The election outcome was confirmed by looking at just 90 individual ballots.

Stark developed a draft set of web-based auditing tools and tested the web tools for the first time in the Ventura audit. There were 11 candidates in this vote-for-three contest. Official results showed the winners to be Cheryl Heitmann with 7,090 votes, Carl E. Morehouse with 6,793 votes, and Christy Weir with 6,515 votes. The runner up was Kenneth M. Cozzens, with 5,564 votes. There were 17,376 ballots cast in all.

To prepare for the audit, Ventura County staff scanned all of the paper ballots* cast in the election to produce digital images. The digital images were processed using ballot tally software developed for the pilot program. The software created a CVR for each ballot and tallied the votes reflected by those CVRs. Ventura County staff kept the ballots in the physical order in which they were scanned so the CVRs could be associated with the paper ballots they represented. The ballots were organized into batches of a maximum of 50 before scanning, to make it easier to find individual ballots.

The initial sample size was 90 individual ballots. The 90 ballots were retrieved and compared to the CVRs. All ballots matched their CVRs exactly, so the audit stopped and the election outcomes were confirmed with 90% confidence (10% risk limit).

*Three ballots could not be located. Since the ballots could have contained votes for the loser, auditors treated the ballots as such, slightly narrowing the margin of victory calculation for the audit, which affects the initial sample size calculation. This ensured that the initial sample size took into account the fact that the missing ballots might have been cast for the loser.

5. Stanislaus County: December 2, 2011

Stanislaus County conducted a risk-limiting audit of City of Oakdale Measure O, in which 3,152 ballots were cast. To prepare for the audit, Stanislaus County staff rented a scanner for a day and scanned all of the paper ballots cast in the election to produce digital images. (One ballot could not be located for scanning; it was treated as a “no” vote by the audit, to ensure that the audit was conservative.) Stanislaus County staff kept the ballots in the physical order in which they were scanned so the CVRs could be associated with the paper ballots they represented. The ballots were organized into batches for scanning to make it easier to find individual ballots. The digital images were processed using software developed for the pilot program. The software created a CVR for each ballot and tallied the votes on the CVRs. According to the software, there were 1,728 “yes” votes and 1,391 “no” votes, a margin of 336 with the missing ballot treated as a “no.” This corresponds to a diluted margin of 336/3152 = 10.6%.

The web tools at http://statistics.berkeley.edu/~stark/Vote/auditTools.htm were used to determine an initial sample size for an audit at 10% risk limit, which turned out to be 49 ballots. A seed for the random number generator was selected by drawing film canisters containing numbered slips of paper at random from an opaque bag. The web tools were then used to select the ballots to audit. The human eye interpretation of all 49 ballots matched the CVRs for those ballots, so the audit stopped. It took approximately 1 hour and 5 minutes to conduct the audit.

The statutory 1% audit required a hand tally of all the ballots cast in one of the five precincts that contained the contest. The precincts ranged in size from 452 ballots cast to 792 ballots cast. The average number of ballots – the expected number of ballots the 1% audit would require tallying in this contest – was 630 ballots. Even though the 1% audit examined far more than the 49 ballots the risk-limiting audit examined, the statutory 1% manual tally could have had a chance as large as 80% of not finding a single error even if the machine-count winner had been wrong. In contrast, the risk-limiting audit had a 90% chance of requiring a full hand count if the machine-count winner had been wrong. Again, this shows the power and efficiency of risk-limiting audits compared to the current statutory audit.

6. Alameda County: December 5, 2011

To prepare for the audit, Alameda County staff used a small county scanner and scanned all of the paper ballots cast in the election to produce digital images. Before scanning the ballots, county staff stamped each ballot with an identification number to make it easier to associate CVRs with the physical ballots. The digital images were processed using ballot tally software developed for the pilot program. The software created a CVR for each ballot and tallied the votes on those CVRs. Alameda County staff kept the ballots in the physical order in which they were scanned so the CVRs could be associated with the paper ballots they represented. The ballots were organized into batches for scanning to make it easier to find individual ballots.

Four City of Alameda contests were audited simultaneously: City Council (vote for 3 of 5) and three measures. All votes were cast on paper ballots; 1,374 ballots were cast in all. The software developed for the pilot found one extra vote for Bukowski for City Council (409 versus 408) and one extra "no" vote for measure F (841 versus 840). The tools at http://statistics.berkeley.edu/~stark/Vote/auditTools.htm were used to determine an initial sample size for an audit at 10% risk limit: 17 individual ballots to be selected at random from the 1,374. Numbered ping-pong balls were drawn at random from a bingo-like tumbler by county staff to generate a seed for the random number generator in the web tool. The 17 ballots were retrieved and compared to the CVRs. All 17 ballots matched their CVRs, so the audit stopped without escalation. Two members of the public observed the audit, which took approximately 25 minutes.

7. Merced County: December 12, 2011

Two City of Merced contests were audited simultaneously, Mayor and City Councilmember (vote for 3 of 8). A total of 7,321 ballots were cast in these contests. The reported winner in the mayoral contest was Stan Thurston, with 2,231 votes; the runner-up was Bill Blake with 2,037 votes. The three reported winners of the City Council contest were Noah Lor (3,736 votes), Mark "Tony" Dossetti (3,669 votes) and Mike Murphy (3,375 votes); runner-up was Richard L. Cervantes (2,416 votes). The diluted margin for the two contests was (2231 - 2037)/7321 = 2.6%, the smallest diluted margin among contests audited under the pilot so far.

To prepare for the audit, Merced County staff used an office scanner they owned to scan all of the paper ballots cast in the election to produce digital images. The digital images were processed using software developed for the pilot program. The software created a CVR for each ballot and tallied the votes on those CVRs. Merced County staff kept the ballots in the physical order in which they were scanned so the CVRs could be associated with the paper ballots they represented. The ballots were organized into batches for scanning to make it easier to find individual ballots.

The web tools at http://statistics.berkeley.edu/~stark/Vote/auditTools.htm were used to determine an initial sample size for an audit at 10% risk limit (which turned out to be 198 ballots), to draw the random sample, and to locate the selected ballots within bundles of stored ballots. The human eye interpretation of all 198 ballots matched the CVRs for those ballots, so the audit stopped without escalation. It took about 3 hours and 15 minutes to conduct the audit.

8. Humboldt County: December 16, 2011

Humboldt County was the first in the program to conduct risk-limiting audits of election results without on-site help from the pilot program team. The county used the draft instructions and web tools developed for the pilot program and conducted a risk-limiting audit of three contests on the ballot.

Humboldt County works with the Humboldt Transparency Project after each election to confirm election results by scanning ballots and creating a parallel tally using Transparency Project software (TEVS). Humboldt County conducted risk-limiting audits of three contests using the Transparency Project CVRs and results for the contests:

Resort Improvement District #1 (elect 3):
(6 candidates)

Total Ballots Cast	193
Ballots Examined for 1% Manual Tally	72
Ballots Examined for Risk-Limiting Audit	52

Eureka City Schools Trustee Area 4 (elect 1):
(2 candidates)

Total Ballots Cast	5,455
Ballots Examined for 1% Manual Tally	15
Ballots Examined for Audit	34

Ferndale Unified School District (elect 2):
(3 candidates)

Total Ballots Cast	640
Ballots Examined for 1% Manual Tally	89
Ballots Examined for Audit	57

Even though Humboldt County examined 176 ballots for the 1% manual tally of the three contests above, the statutory 1% manual tally left at least a 50% chance of not finding a single error, even if the machine-count found a wrong winner.

In contrast, the risk-limiting audit involved reviewing fewer ballots – 143 ballots – and guaranteed a 90% chance of catching and correcting errors that could have caused the voting system to find a wrong winner. Since no errors were found in the initial sample for each contest, no escalation was needed to confirm results.

As with the prior audits in other counties, the Humboldt County audit showed the power and efficiency of risk-limiting audits compared to the flat 1% manual tally currently required by law.

<u>Cost-Efficiency Analysis: Risk-Limiting Audits Compared to the 1% Manual Tally</u>

The time it took to conduct the audits was minimal – a few minutes to a few hours – compared to the time it takes to conduct the 1% manual tally. However, counties participating in the pilot audits typically spent significantly more money on the risk-limiting audit than they did for their 1% manual tally.

There are two reasons for this:

1) Counties spent a considerable amount of time scanning the ballots in preparation for the audits. Each county used a standard office scanner or scanner/copier, rather than a high-speed scanner. For the 2012 audits, the team plans to help counties rent high-speed scanners to minimize the time spent scanning ballots in preparation for the audits; and

2) The audits conducted in 2011 were of small local elections. During 2012, larger audits are planned following the June and November elections to show the efficiencies that can be created with risk-limiting audits of larger elections.

Participating counties submitted spreadsheets, detailing the risk-limiting audit costs compared to the cost of the 1% manual tally for the election (see attached).

Final Report

The Post-Election Risk-Limiting Audit Pilot Program will be completed at the end of 2012. The Secretary of State will share with the Legislature the final report to the EAC on the project.

ALAMEDA COUNTY
Calculation of Conducting Risk-Limiting Audit vs. 1% Tally
for the November 8, 2011 UDEL Election

Risk-Limiting Audit:

		Hours/Mins	Rate	Total
Three Board Members:	Chun Lin	15.50	$ 36.13	$ 560.02
	Van To	15.50	$ 25.01	$ 387.66
	Janet Yabut	16.00	$ 49.91	$ 798.56
Find Particular Ballots:	David Pink	21.75	$ 35.21	$ 765.82
	Andrew Seto	21.75	$ 68.60	$ 1,492.05
Survey and Research:	Lolita Francisco	20.00	$ 58.99	$ 1,179.80
Totals		**110.50**		**$ 5,183.90**

vs.

1% Tally:

		Hours/Mins	Rate	Total
Nine Board Members:	Jan Blythe	1.00	$ 51.55	$ 51.55
	Jacqueline Lam	6.00	$ 54.31	$ 325.86
	Gemma Nimedez-Arbas	3.00	$ 53.14	$ 159.42
	Lauren Perez	7.50	$ 45.82	$ 343.65
	Hester Sun	7.50	$ 50.73	$ 380.48
	Van To	7.00	$ 25.01	$ 175.07
	Zenaida Valerio	7.00	$ 25.01	$ 175.07
	Philip Cable	7.00	$ 25.01	$ 175.07
	Leo Fernandez	4.50	$ 45.50	$ 204.75
Paper Ballots/Tally Reports:	David Pink	7.50	$ 35.21	$ 264.08
Totals		**58.00**		**$ 2,254.99**

HUMBOLDT COUNTY
Calculation of Conducting Risk-Limiting Audit vs. 1% Tally
Consolidated District Election

Risk-Limiting Audit:							1% Tally:						
		Hours/Mins	**Rate**		**Total**	**vs.**			**Hours/Mins**	**Rate**		**Total**	
Four Board Members:	Amanda Windsor	4.00	$	9.22	$	36.88		Four Board Members:	Florence Sheldon	5.00	$	9.22	$ 46.10
	Nancy Nieboer	2.00	$	9.22	$	18.44			Alice Gay	5.00	$	9.22	$ 46.10
	Kelly Sanders	11.00	$	47.83	$	526.13			Slater Smith	5.00	$	9.22	$ 46.10
	Judi Hedgpeth	8.00	$	34.83	$	278.64			Fern Miller	5.00	$	9.22	$ 46.10
	Mitch Tractenberg	3.00	$	35.00	$	105.00			Virginia Prince	8.00	$	24.94	$ 199.52
	Teddy Wilwerding	2.00	$	29.86	$	59.72							$ -
	Carolyn Crnich	12.00	$	78.30	$	939.60							$ -
													$ -
Totals		**42.00**			**$**	**1,964.41**		**Totals**		**28.00**			**$ 383.92**

MERCED COUNTY
Calculation of Conducting Risk-Limiting Audit vs. 1% Tally
for the City of Merced Election

Risk-Limiting Audit:

		Hours/Mins	Rate	Total
Board Members:	Stacey Cotter	4.00	$ 83.54	$ 334.16
	Shawnesti Machado	4.00	$ 43.85	$ 175.40
Find Particular Ballots:	Alicia Ponder	4.00	$ 48.39	$ 193.56
Scanning:	Shawnesti Machado	71.00	$ 43.85	$ 3,113.35
Batching:	Marina Ortega	9.50	$ 46.48	
Totals		**92.50**		**$ 3,816.47**

vs.

1% Tally:

		Hours/Mins	Rate	Total
Board Members:	Betty Hale	11.00	$ 16.50	$ 181.50
	Bev Raggio	14.00	$ 16.50	$ 231.00
	Alicia Ponder	1.00	$ 48.39	$ 48.39
	Cyndi Helton	14.75	$ 15.98	$ 235.71
	Tammy Lyons	5.00	$ 35.83	$ 179.15
	Lea Hernandez	4.00	$ 46.44	$ 185.76
				$ -
Tally Reports:	Shawnesti Machado	1.00	$ 43.85	$ 43.85
Totals		**50.75**		**$ 1,105.36**

MONTEREY COUNTY
Calculation of Conducting Risk-Limiting Audit vs. 1% Tally
for the County of Monterey Election, May 2011

Risk-Limiting Audit:					**vs.**	**1% Tally:**				
		Hours/Mins	**Rate**	**Total**				**Hours/Mins**	**Rate**	**Total**
Four Board Members:	Election Worker #1	1.00	$ 35.67	$ 35.67		Four Board Members:	Election Worker #1	1.00	$ 35.67	$ 35.67
	Election Worker #2	1.00	$ 25.02	$ 25.02			Election Worker #2	1.00	$ 25.02	$ 25.02
	Election Worker #3	1.00	$ 26.22	$ 26.22			Election Worker #3	1.00	$ 26.22	$ 26.22
	Election Worker #4	1.00	$ 25.02	$ 25.02			Election Worker #4	1.00	$ 25.02	$ 25.02
Totals		**4.00**		**$ 111.93**		**Totals**		**4.00**		**$ 111.93**

NOTE: This was a vote by mail only election. There was only one contest in this election with seven precincts participating. Therefore the 1% tally was one precinct and approximately 400 ballot verified. Given the size of the election it is difficult to see the benefits of the Risk-Limiting Audit, but the process of the Risk-Limiting Audit would most likely save more time and resources during a larger and more complex election.

ORANGE COUNTY
Calculation of Conducting Risk-Limiting Audit vs. 1% Tally
for the San Clemente Election

Risk-Limiting Audit:

		Hours/Mins	Rate	Total
Four Board Members:	Delicia Hsu	1.75	$ 30.80	$ 53.90
	Jay Koo	1.75	$ 30.80	$ 53.90
	Marcia Nielsen	1.75	$ 29.52	$ 51.66
	Jessica Castaneda	1.75	$ 21.06	$ 36.86
Find Particular Ballots:	Christian Tran	1.00	$ 36.82	$ 36.82
	Justin Berardino	4.00	$ 47.76	$ 191.04
Survey and Research:	Brent Adams	3.00	$ 19.87	$ 59.61
Totals		**15.00**		**$ 483.79**

vs.

1% Tally:

		Hours/Mins	Rate	Total
Four Board Members:	Delicia Hsu	1.00	$ 30.80	$ 30.80
	Jay Koo	1.00	$ 30.80	$ 30.80
	Jessica Castenada	1.00	$ 21.06	$ 21.06
	Kay Cotton	1.00	$ 52.22	$ 52.22
Paper Ballots:	Christian Tran	1.00	$ 36.82	$ 36.82
Tally Reports:	Justin Berardino	1.00	$ 47.76	$ 47.76
Eslate Roll:	Brent Adams	0.50	$ 19.87	$ 9.94
	Derek Moore	1.00	$ 28.28	$ 28.28
Totals		**7.50**		**$ 257.68**

SAN LUIS OBISPO COUNTY
Calculation of Conducting Risk-Limiting Audit vs. 1% Tally
for the City of San Luis Obispo Election (all vote-by-mail)
10,389 ballots- 2 measures

Risk Limiting			**vs.**	**1% manual Tally**		
		Hours/Mins				**Hours/Mins**
Four Board Members:	L. Zohns	1.75		Four Board Members:	C. Martin	1.00
	V. Wallen	1.75			C. Fontan	1.00
	E. Cano	1.75			L. Zohns	1.00
	T. Gong	1.75			V. Wallen	1.00
Find Particular Ballots:	J. Martinez	2.00		Sort Paper Ballots	J. Martinez	4.00
				Tally Reports:	T. Gong	1.50
Survey and Research:	J. Rodewald	2.50				
	T. Gong	2.50				
	J. Martinez	2.50				
Scanning Ballots	T. Gong	20.00				
Totals		**36.50**		**Totals**		**9.50**

STANISLAUS COUNTY
Calculation of Conducting Risk-Limiting Audit vs. 1% Tally
for the Oakdale Measure "O" Election

Risk-Limiting Audit:						**1% Tally:**				
		Hours/Mins	***Rate**	**Total**	**vs.**			**Hours/Mins**	***Rate**	**Total**
Three Board Members:	Registrar of Voters	1.25	$ 61.50	$ 76.88		Four Board Members:	Election Staff	0.50	$ 22.01	$ 11.01
	Election Management	1.25	$ 29.70	$ 37.13			Election Staff	0.50	$ 11.88	$ 5.94
	Prof. Stark	1.25		$ -			Election Staff	0.50	$ 11.88	$ 5.94
				$ -			Election Staff	0.50	$ 11.88	$ 5.94
Find Particular Ballots:	Election I.T.	1.00	$ 40.38	$ 40.38		1% Draw Preparation:	Election Staff	1.05	$ 22.01	$ 23.11
	Election Management	0.25	$ 29.44	$ 7.36		Paper Ballots:	Election Staff	0.25	$ 22.01	$ 5.50
						Tally Reports:	Election I.T.	0.08	$ 40.38	$ 3.37
Ballot Scanning	Election Staff	4.00	$ 16.17	$ 64.68						
	Election Supervisor	3.75	$ 23.15	$ 86.81						
	Scanner Rental Cost		$ 402.66	$ 402.66						
Totals		**12.75**		**$ 715.89**		**Totals**		**3.38**		**$ 60.80**

*Rate includes wages only

Ventura County

PEOPLESOFT PAYROLL REPORTS FOR 2011

Payrol Costs for Scanning Process

Name	Hours/Min	Rate	Direct Charges
Chavez,Daniel	39.50	36.36	1,436.27
Lopez,Gonzalo	40.00	27.10	1,084.06
Chavez,Daniel	20.00	38.23	764.63
Lopez,Gonzalo	18.00	30.72	552.97
	117.50	$	3,837.92

Scan Care Kit for Scanner:		550.00
Scanner Maintenance:		700.00
Auto Rental (GSA Charges):		113.86
Total Cost:	$	5,201.78
(Overage) Amount Excluded		(201.78)
Maximum Grant Amount:	$	5,000.00

1% Hand Tally

	Name	Hours	Rate $	Total Charges
Board Members	Lopez,Gonzalo	4	$26.13	104.52
	Philo,Gloria Louise	4	$25.57	102.28
	Reeder,Marilyn Christine	4	$12.93	51.72
	Cornejo,Robert	4	$13.70	54.80
				-
	Chavez, Daniel	4	$35.10	140.40
Set-up	Philo,Gloria Louise	4	$25.57	102.28
		24		556.00

DEBRA BOWEN | SECRETARY OF STATE | STATE OF CALIFORNIA
1500 11th Street, 6th Floor | Sacramento, CA 95814 | Tel (916) 653-7244 | Fax (916) 653-4620 | www.sos.ca.gov

DRAFT
Post-Election Risk-Limiting Audit Pilot Program 2011-2012
Step-by-Step Instructions for Conducting Risk-Limiting Audits

Step 1: Provide Public Notice and Educate the Public

The elections official must provide a five-day public notice of the date, time, and place of the post-election risk-limiting audit. The form and method of public notice may be similar to the notice provided for the 1% manual tally, as required by Elections Code (EC) section 15360. Elections officials should share with the public the "Audit Tools" website and the draft Step-by-Step Instructions created for the pilot program, which describe the audit model and the statistical calculations behind the model.

Step 2: Secure the Audit Trail

For purposes of risk-limiting audits, the county's official canvass procedures, adopted and implemented under EC section 15003, must include the following:

1. Vote-by-mail (VBM) ballot security procedures, as specified by EC sections 3019 and 15101 including:

   a. Secure returned VBM ballots: Until 29 days prior to the election, the elections official must keep all returned VBM ballots in a secure room[1]. Beginning 29 days prior to the election, all VBM ballots awaiting processing must be held in a secure room. The elections official must keep track of the number of ballots received. (best practices)

   b. Verify VBM envelopes: Beginning 29 days prior to the election, the elections official may begin processing returned VBM envelopes.

      i. Verify signature on outside of envelope by comparing it to the signature on the voter registration form; and

      ii. Update voter's history file.

   c. Remove VBM ballots from envelopes: Beginning 7 days prior to the election, the elections official may begin processing VBM ballots.

      i. Remove ballots from signature-verified envelopes and prepare the ballots for machine tally; and

[1] A "secure room" must be locked. Video monitoring of the room is recommended. A two-person rule for access to the room is recommended, similar to the two-person rule below for ballot transport of ballots to the elections office.

ii. Duplicate damaged ballots.[2]

d. Scan VBM ballots: Beginning 7 days prior to the election, the elections official may machine tally the votes on VBM ballots.

   i. VBM ballots that have been processed and prepared for the machine tally must be batched and brought to a secure room to await the machine tally; (best practice)

   ii. As with the tallying of other ballots, quality control procedures, such as attaching a batch report to each batch of scanned VBM ballots, must be established and followed; (best practice)

   iii. VBM ballots that have been machine tallied must be returned to a secure room; and (best practice)

   iv. Results of any VBM ballot tabulation may not be accessed, examined, or released prior to the close of the polls on Election Day. EC section 15101.

2. Polling place ballot security procedures, as specified in EC sections 15201 and 15202, including the following:

   a. Seal[3] the container used to transport voted ballots;

   b. Ensure that the precinct number is designated on the ballot container;

   c. Certify, sign, and seal polling place materials packages and envelopes; and

   d. Deliver the ballot container and packages to the elections official. At least two precinct board members must travel together to make the delivery. The ballot container and packages must remain in the exclusive possession of the two precinct board members until delivered.

3. Canvass security procedures, including ballot accounting and reconciliation tasks, as specified in EC sections 15302 et seq. and 15370, including the following:

   a. Inspect all materials and supplies returned by poll workers;

---

[2] If a duplicate ballot is selected for the post-election audit, both the original ballot and the duplicate ballot must be viewed to ensure the duplicate reflects the voter's intent. Therefore, the originals of duplicate ballots should be stored in a manner that allows for easy retrieval.

[3] The elections official must have a protocol for affixing, checking, removing and replacing seals. The protocol must include steps that must be taken if a broken seal is discovered.

b. Reconcile[4] the number of signatures on the roster with the number of ballots recorded on the ballot statement;

c. Reconcile the number of ballots counted, spoiled, canceled, or invalidated with the number of ballots recorded by the voting system;

d. Implement chain-of-custody procedures for voted ballots; (best practice)

e. Seal ballots after counting is complete;

f. Implement quality control measures, such as verifying that precinct subtotals sum to overall totals tabulated by the voting system; and (best practice)

g. Report final canvass results to the public[5], the county's governing board, and the Secretary of State.

4. Public observation procedures:

   a. Vote-by-Mail: The processing and counting of vote-by-mail ballots, both prior to and after the election, shall be conducted as required by EC section 15104, as follows:

      i. Elections officials must make VBM processing and counting open to the public;

      ii. Elections officials must provide 2-day public notice of the dates, times, and places where VBM ballots will be processed and counted;

      iii. Elections officials must allow observers to challenge the manner in which VBM ballots are handled, processed and counted. The elections official must document and share with the public all challenges;

      iv. Elections officials must allow observers "sufficiently close access" to enable them to observe the VBM ballot return envelopes and signatures and challenge procedures, including signature verification, duplication of ballots, securing of VBM ballots to prevent tampering; and

      v. Observers may not touch or handle the ballots.

---

[4] The elections official must have a protocol to address instances where the number of voted ballots does not match the number of signatures or the voting system's record of the number of ballots counted.

[5] Publishing results on the county website is recommended.

b. Polling Place: The precinct board must conduct polling place security procedures in the presence of the public, as required by EC section 15201.

c. Canvass: All proceedings at the central county place, or places, must be open to the view of the public. Observers may not touch any ballot container. Access to the area where electronic data processing equipment is operated may be restricted. EC section 15204.

Step 3: Finish and Publish the Official Canvass Results

If the voting system can produce ballot-level results (i.e., cast vote records), publish ballot-level results in addition to precinct- and contest-level results.

Step 4: Conduct a Transparency Scan and Tally of the Ballots

Skip this step if the county has a voting system that can produce ballot-level results, i.e. a cast vote record for each ballot that can be associated with the corresponding physical ballot.

To be efficient, risk-limiting audits must be conducted at the ballot level – not the precinct level. California's 1% manual tally law (EC section 15360) requires elections officials to select and hand tally all of the ballots from 1% of all precincts. In contrast, efficient risk-limiting audits involve selecting and interpreting individual ballots to verify overall election outcomes.

Since current vote tabulation systems cannot produce a record of the votes tallied from each ballot (called a "cast vote record") that can be associated with each ballot, the elections official conducts a "transparency scan" of all ballots in the contest(s) to be audited.

Software has been developed (which will be licensed as free, open-source software in the future) for use in the California Secretary of State's post-election risk-limiting audit pilot program that can be configured to create cast vote records from ballots, given a detailed description of each ballot style in the election.

Transparency scan and tally procedures:

1. Public Notice: Provide a five-day public notice of the date, time, and place of the transparency scan and tally.

2. Public Observation: The transparency scan and tally must be open to the view of the public.

3. Public Education: The elections official should speak to public observers to explain the audit process, the goal of the audit (confirm winners), and the need for a transparency scan due to current voting system constraints.

4. Ballot security: Ballots should be held in a secure location before and after the scanning process. Observers may not touch or handle ballots.

5. Configure transparency software: In order to configure the transparency software used in the pilot, the auditor will need: 1) a sample ballot; 2) a pdf copy of each ballot style created by the voting system, with camera-ready art for ballot printing – i.e., auditors need a copy of the files provided to your ballot printer; and 3) test scans of blank ballots and marked ballots, using the scanner(s) that will be used in the audit. The test scans should be either 200 dpi grayscale or color, with nearly lossless compression, such as nearly lossless .tiff files. Auditors will use the files to calibrate the transparency software, so that the software can interpret the marks on ballot images as votes.

6. Use a commercial off-the-shelf (COTS) scanner/copier/fax machine to scan the ballots cast in the contests to be audited: The scanner must be capable of producing 600dpi .jpg color or 200dpi lossless tiff color images of each ballot scanned. Black and white images are not sufficient, but grayscale may be sufficient, depending on whether elections staff have “enhanced” voter marks using a highlighter.

7. Scan ballots in batches. Create one digital image file for each ballot scanned, either in 600dpi .jpg color format or in 200dpi lossless tiff color format.

8. Keep ballots in the order in which they were scanned and create labeled stacks of scanned ballots, placing markers every 50 ballots in each batch. This allows individual ballots selected during the audit to be retrieved efficiently. Create a ballot “manifest” that lists how scanned ballots are organized (labeling of stacks, listing the number of ballots in each stack). Conduct quality control measures to ensure that each stack of ballots contains the number of ballots listed in the manifest.

9. Tally the votes on the ballot images and create a cast vote record (CVR) for each ballot image. Auditors will use the transparency software created for the pilot program to accomplish this task. The software interprets voters' marks from each ballot image and associates that interpretation (CVR) with the corresponding physical ballot. Auditors will do a human eye interpretation of ballots with marks flagged as questionable, for example ballots with light marks or marks outside of voting targets.

10. Check the transparency tally results against the county’s voting system results. If the transparency tally and the official results both show the

same winner(s) for each contest, the risk-limiting audit may proceed. (The risk-limiting audit confirms winners, not exact vote totals.) If the results do not show the same winner(s), then the elections official should make sure all ballots were included in both tallies – the voting system tally and the transparency tally – and undertake any other quality control measures deemed necessary. If the winners found by the transparency scan still differ from the official voting system results for any contest, then all the ballots for that contest must be counted by hand.

11. Explain to public observers that since both systems show the same winners, an audit of either system confirms ***transitively*** that both systems show the right winners or it corrects both outcomes.

12. Commit to the transparency tally results: There are several ways to preserve a copy of the final transparency tally to prove to the public that the results are untouched and unchanged before, during, and after the audit. The preferred method is to publish CVRs on a website using a digital signature key. Alternative methods are: burn a DVD copy of the CVRs and upload the DVD to a laptop with no Internet connection; use cryptographic methods to secure a copy of the CVRs; or, for example, CD copies of the CVRs may be provided to election observers to use in observing the audit, which will compare individual physical ballots with the corresponding CVR. The elections official should commit to the transparency tally results in the presence of the public.

Step 5: Determine the Size of the Initial Audit Sample

For the "super-simple" audit model, the initial sample size depends on the "diluted margin," which is the margin of victory[6] in votes divided by the number of ballots cast. This differs slightly from usual ways of calculating the margin, because it divides by the number of ballots rather than the number of valid votes. Dividing by the number of ballots takes into account that the vote tabulation system might confuse an undervote or overvote for a valid vote, or vice versa. The super-simple method starts with an initial sample of individual ballots and either stops or escalates, depending on the number and nature of errors found in the initial sample.

Go to the Post-Election Audit Tools Website and follow the instructions to enter contest data. The elections official may develop software tools based on the audit model. For purposes of the pilot program, the risk limit setting should be left at 10%: http://statistics.berkeley.edu/~stark/Vote/auditTools.htm

---

6 "Margin of victory" is the number of votes between the winning candidate with the fewest votes and the losing candidate with the most votes. For a contest involving a measure, it is the number of votes between the "Yes" and threshold for passage (e.g., 50% + 1 or 66 2/3% +1). The margin of victory for contests that include more than one jurisdiction is the overall margin of victory in the contest, not the margin of victory within the jurisdiction.

Illustration: For *zero* expected over/understatements, the math is simple. The number 4.8 divided by the diluted margin provides the sample size:

Diluted Margin	**Equation**	**Initial Sample Size (in ballots)**
50%	4.8/.5	10
40%	4.8/.4	12
30%	4.8/.3	16
20%	4.8/.2	24
10%	4.8/.1	48
5%	4.8/.05	96
2%	4.8/.02	240
1%	4.8/.01	480
.5%	4.8/.005	960

Step 6: Randomly Select Individual Ballots for the Sample

The elections official may be able to use the same random selection method used for the 1% manual tally. Alternatively, the elections official may use the following method:

1. Ask public observers to roll eight 10-sided dice to generate a "seed" number for a public-source pseudo-random number generator.

2. Enter the seed on the “Random Sampling” section of the Audit Tools website: http://statistics.berkeley.edu/~stark/Vote/auditTools.htm

3. Enter the number of ballots in the contest(s) to be audited.

4. Click “draw sample” to select ballots.

5. Each random number corresponds to one ballot. For example, if there are 1,000 ballots total in 5 equal stacks, then the ballots in Stack 1 can be ballots 1 through 200; Stack 2 can be ballots 201 through 400; Stack 2 can be ballots 401 through 600, etc. If ballot number 341 is randomly selected, then the elections official can pull that ballot by counting into Stack 2. As noted above, the instructions recommend placing a marker every 50 ballots in the stacks to facilitate retrieving individual ballots.

Step 7: Compare Physical Ballots in the Sample with Corresponding CVRs

Because risk-limiting audits involve comparing individual ballots to the voting system results for each ballot, the audits do not involve a "tally" or count in the traditional way the 1% manual tally is conducted. That is, the 1% manual tally ballot counters are assembled to tally entire precincts of ballots and compare the hand tally totals to the totals produced by the voting system for the same precinct. For risk-limiting audits, a human eye interpretation of each ballot is compared to the CVR for that ballot as recorded by the voting system, so ballots are not "tallied" or counted up and totaled.

Compare each ballot in the sample with its CVR as follows:

1. Retrieve the ballots chosen for the sample. Those designated to retrieve ballots should not have access to the CVRs for the ballots they retrieve in order to ensure the integrity of the audit.

2. Retrieve the CVR for each ballot (created in the transparency tally) and determine whether the CVR matches a human eye interpretation of the votes on the corresponding ballot. Existing methods used for the 1% manual tally may be used to the extent applicable, including the rules for determining voter intent.

3. Ensure public observers have the opportunity to compare the CVR and the corresponding ballot to verify whether the CVR matches a human eye interpretation of the voter's intent.

4. Document and share with the public any differences found between the hand tally and the software tally of the votes on each ballot.

5. Document and share with the public any instances, in which one or more public observers disagreed with the hand tally interpretation of a ballot or with the audit procedures.

6. Establish procedures to handle observer challenges to the audit or audit procedures. The public must be allowed to observe, verify and point out procedural problems in all phases of the audit but without interfering with the process.

Step 8: Escalate If Necessary

Depending on the number and type of overstatements and understatements found in the initial sample, the audit may need to be expanded to look at more ballots. To determine how many more ballots should be hand tallied, if any, assuming a similar rate of over/under statements:

1. Go to the Post-Election Audit Tools Website and follow the instructions to determine whether escalation is necessary.

2. If escalation is necessary, the elections official should explain to the public that the audit may lead to a full hand count if significant differences persist.

Step 9: Finish and Publish Results

1. Complete the audit by releasing the results to the public and sending a brief report to the Secretary of State, which includes:

   a. The time it took to conduct the audit, with a breakdown of the time needed to scan ballots compared to the time needed to conduct the audit itself (random selection and manual tally);

   b. The cost of the audit, with a breakdown of both parts: ballot scan and audit; and

   c. The cost of the 1% manual tally for the same election.

Tools for Risk-Limiting Election Audits

This page implements some of the methods described in A Gentle Introduction to Risk-Limiting Audits (AGI), by Lindeman and Stark.

To hide or show everything but the tools, click this link.

A *risk-limiting audit* is a procedure that is guaranteed to have a large chance of progressing to a full hand count of the votes if the electoral outcome is wrong. The outcome according to the hand count then replaces the outcome being audited. The *risk limit* is the maximum chance that the audit will not progress to a full hand count if the electoral outcome is incorrect, no matter why it is incorrect—whether because of voter error, bugs, pollworker error, or deliberate fraud—provided the audit trail is complete and accurate.

There are many methods for conducting risk-limiting audits. This page performs calculations for a particularly simple method described in AGI. The method is a type of *comparison audit*: It involves comparing the interpretation of ballots according to the voting system (the cast vote record or CVR) to a human interpretation of the same ballot. Differences between the two interpretations are noted. Determining whether the audit can stop depends on the number and nature of those differences, the number of ballots examined so far, the risk limit, and the diluted margin. The smaller the risk limit or the diluted margin, the larger the number of ballots that must be audited, all else equal.

The difference can be neutral, an *understatement*, or an *overstatement*, depending on the effect of changing the voting system interpretation of the ballot to match the hand interpretation: Consider the *pairwise margin* between each winner and each loser in a contest. For instance, a city council election might involve voting for three candidates from a pool of ten, to fill three seats on the council. Then each of the three winners can be paired with each of the seven losers, giving twenty-one pairwise margins in that contest. If changing the interpretation of a ballot according to the voting system to make it match the human interpretation of the ballot would widen *every* pairwise margin in *every* contest under audit, that ballot has an *understatement*. Understatements do not call the outcome into question. If changing the interpretation according to the voting system to match the human interpretation would narrow *any* pairwise margin in *any* contest under audit, the ballot has an *overstatement*. If enough ballots have overstatements, the outcome could be wrong.

The sample size calculations for this method depend on the risk limit as well as the *diluted margin*, which is the margin in votes divided by the number of ballots cast in any of the contests being audited together, including undervotes and overvotes. Undervotes and overvotes are included because they might have been intended as votes for candidates, misinterpreted by the voting system as undervotes or overvotes. Because the

Efficient risk-limiting audits generally count votes by hand until there is convincing evidence that the outcome according to a full hand count would agree with the outcome under audit. If convincing evidence is not forthcoming, the audit progresses to a full hand count, which is used to correct the outcome under audit if the two disagree.

The tools on this page help perform the following steps:

Visualizing the required sample size

The ultimate sample size required to confirm the outcome depends on the diluted margin and the number of errors (both understatements and overrstatements) found in the sample, as well as the risk limit. The following graph plots the sample size as a function of the number of 1-vote overstatement errors the audit finds, for diluted margins of 0.5%, 1%, 5%, 10%, and 20%, all at risk limit 10%. It also plots the expected final sample size as a function of the diluted margin, for various rates of observed 1-vote overstatements. The plot assumes that there are no 2-vote overstatements and no understatements.

Expected sample size as a function of the diluted margin, 10% risk limit

☐ Plot final sample size as a function of observed 1-vote overstatements

Initial sample size

The initial sample size tool lets you enter the particulars of the contest(s) to be audited as a group: the total ballots cast across all the contests combined, and the vote totals for each candidate in each contest. The form helps you anticipate the number of randomly selected ballots that will need to be compared to their CVRs to attain a given limit on the risk, under assumptions about the rates of differences anticipated. It is completely legitimate to sample one at a time and check whether enough have been sampled using the "stopping sample size tool," (later in this page) but this form can help auditors anticipate how much work will be required and retrieve ballots more efficiently, by reducing the number of times a given batch of ballots is opened.

Enter the total number of ballots cast in all contests to be audited. Add candidates and contests as necessary until the results from all contests have been entered. Enter the desired risk limit and the expected rates of one- and two-vote differences. Select whether to round up the expected number of differences of each type. Finally, click "calculate" to find the starting sample size.

Initial sample size

Contest information

Ballots cast in all contests: [ ] Smallest margin (votes): undefined. Diluted margin: undefined.

Contest 1. Contest name:

[ ]

Vote for no more than 1

Reported votes:

Candidate 1 Name: [ ] Votes: [ ]
Candidate 2 Name: [ ] Votes: [ ]

[Add candidate to contest 1] [Remove last candidate from contest 1]

[Add contest] [Remove last contest]

Audit parameters

Risk limit: 10%

Expected rates of differences (as decimal numbers):

Overstatements. 1-vote: 0.001 2-vote: 0.0001
Understatements. 1-vote: 0.001 2-vote: 0.0001

Starting size

☑ Round up 1-vote differences. ☐ Round up 2-vote differences. [Calculate size] ...

By default, this form assumes that rate of one-vote understatements and overstatements is one in a thousand (0.001) and that the rate of two-vote understatements and overstatements is one in ten thousand (0.0001). These values are conservative, in my experience, but the choice is up to the user. The larger these rates are assumed to be, the larger the initial sample size will be. Taking a larger initial sample can avoid needing to expand the sample later, depending on the rate of errors the audit actually finds. Avoiding "escalation" can make the audit less complicated.

Considerations for deciding which contests to audit together

The number of ballots the audit must examine before stopping depends on the smallest diluted margin among the contests to be audited together (as well as the risk limit, the errors the audit finds, and so on). All else equal, the larger the diluted margin is, the smaller the sample size needs to be.

Because the diluted margin is the smallest margin in votes divided by the total number of ballots cast in all the contests under audit, auditing small contests together with large contests can be inefficient: Dividing the vote margin in small contests by the number of ballots cast in large contests can make the diluted margin very small, which makes the required sample very large.

Generally, if two contests overlap substantially—for instance, if both are jurisdiction-wide contests—it is more economical to audit them together: Fewer ballots will need to be inspected in all. Conversely, if two contests do not overlap at all, it is more efficient to audit them separately.

Auditing small contests together with (overlapping) large contests generally is not efficient unless the vote margin in the small contests is a substantial fraction of the vote margin in the large contests. That is, auditing small contests that have large percentage margins together with large contests that have small percentage margins can be efficient, but auditing small contests together with large contests that have comparable vote margins generally is not efficient, because it makes the diluted margin of the combination much smaller.

The tool above can be used to explore whether it makes sense to audit a collection of contests together by checking whether the required starting sample size when they are audited together is greater than the sum of the required starting sample sizes when they are audited separately. (If you experiment with different groupings of contests, be sure to change the entry for "Ballots cast in all contests" to reflect only the contests that are to be audited together.)

Hide technical notes.

Random sampling

The next tool helps generate a pseudo-random sample of ballots. To start, input a random seed with at least 20 digits (generated by rolling a 10-sided die, for instance), the number of ballots from which you want a sample, and the number of ballots you want in the sample. Further below, there is a form to help find the individual, randomly selected ballots among the batches in which ballots are stored.

Pseudo-Random Sample of Ballots

Seed: Input seed here

Number of ballots: 1000

Current sample number: 0

Draw this many ballots: 1 draw sample reset

Ballots selected: ☑ show sequence numbers ☐ show hash values

Ballots selected, sorted:

Ballots selected, sorted, duplicates removed:

Repeated ballots:

Hide technical notes.

Find ballots using a ballot manifest

Generally, ballots will be stored in batches, for instance, separated by precinct and mode of voting. To make it easier to find individual ballots, it helps to have a *ballot manifest* that describes how the ballots are stored. For instance, we might have 1,000 ballots stored as follows:

Batch label	ballots
Polling place precinct 1	130
Vote by mail precinct 1	172
Polling place precinct 2	112
Vote by mail precinct 2	201
Polling place precinct 3	197
Vote by mail precinct 3	188

If ballot 500 is selected for audit, which ballot is that? If we take the listing of batches in the order given by the manifest, and we require that within each batch, the ballots are in an order that does not change during the audit, then the 500th ballot is the 86th ballot among the vote by mail ballots for precinct 2: The first three batches have a total of 130+172+112 = 414 ballots. The first ballot in the fourth batch is ballot 415. Ballot 500 is the 86th ballot in the fourth batch.

The ballot look-up tool transforms a list of ballot numbers and a ballot manifest into a list of ballots in each batch. Batch labels should not contain commas. Use a comma to separate each batch label from the number of ballots in that batch (or the range of ballot numbers or the set of ballot identifiers—see below). The manifest should have one line per batch and no empty lines.

For instance, to input the ballot manifest above, you would enter:

```
Polling place precinct 1, 130
Vote by mail precinct 1, 172
Polling place precinct 2, 112
Vote by mail precinct 2, 201
Polling place precinct 3, 197
Vote by mail precinct 3, 188
```

Some jurisdictions number the ballots cast in an election. If all the ballots in an election are numbered sequentially, the numbers on the ballots that contain a particular contest might not be sequential. For instance, an election might cover precincts 1, 2, and 3, but only voters in precincts 1 and 3 are eligible to vote in the contests to be audited with the current sample. In the previous example, suppose that the jurisdiction had stamped numbers on all the ballots, sequentially, so that the ballots from the polling place in precinct 1 were numbered 1 to 130, the vote by mail ballots from precinct 1 were numbered 131 to 302, the ballots from the polling place in precinct 2 were numbered 303 to 414, and so on, as summarized in the following table:

Batch label	ballot range
Polling place precinct 1	1 to 130
Vote by mail precinct 1	131 to 302
Polling place precinct 2	303 to 414
Vote by mail precinct 2	415 to 615
Polling place precinct 3	616 to 812
Vote by mail precinct 3	813 to 994

Provisional ballots for precinct 1	996, 998, 1000
Provisional ballots for precinct 2	997
Provisional ballots for precinct 3	995, 999

Since the ballots already have numbers on them, it makes sense to look them up using those numbers. If we were auditing a collection of contests that included only precincts 1 and 3, the ballots subject to audit would be the 686 ballots labeled 1 to 130, 131 to 302, 616 to 812, and 813 to 994, and 995, 996, 998, and 1000. In this case, the ballot manifest would include *only* the six batches that comprise precincts 1 and 3, not all eight batches; there are only 686 ballots in these batches. Each line in the manifest would consist of a batch label and a range of ballot numbers, where the range is denoted by a colon, or of a batch label and a set of ballot identifiers in parentheses, separated by spaces. Ballot ranges cannot have gaps: There can be no missing numbers within the range for any single batch. (If there is in fact a gap, input the numbers as a set of identifiers, rather than as a range.) Again, separate the label from the range or set of ballot numbers by a comma. The label must not contain any commas, and the range of ballot numbers or set of identifiers must not contain commas. In this example, we would enter the ballot manifest as follows:

```
Polling place precinct 1, 1:130
Vote by mail precinct 1, 131:302
Polling place precinct 3, 616:812
Vote by mail precinct 3, 813:994
Provisional precinct 1, (996 998 1000)
Provisional precinct 3, (995 999)
```

The total number of ballots in the manifest must equal the number cast in the contests that are to be audited together using the sample (686 in this example).

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.

Ballots to look up (separated by commas):

look up ballots

Sorted lookup table:

Should more ballots be audited?

The stopping sample size tool determines whether enough ballots have been examined for the audit to stop, and if not, estimates how many more ballots will need to be audited. The answer depends on the risk limit, the margin, and the differences between the cast vote records and the manual inspection of the ballots in the sample.

Differences matter according to how they affect the *pairwise margin* between some winner and some loser in some contest. Suppose we are auditing a mayoral contest with four candidates, a city council contest that allows voting for up to three of ten candidates, and a simple measure that involves voting either "yes" or "no." The mayoral contest has three pairwise margins: The winner can be paired with each of the three losers. The city council contest has twenty-one pairwise margins: each of the three winners can be paired with each of the seven losers. The measure has but one pairwise margin, since it has only one winner and one loser. In all, there are 3+21+1 = 25 pairwise margins among the three contests being audited.

If there is any difference between the cast vote record and the human interpretation of a ballot, that ballot as a whole may have an understatement of one or two votes, or an overstatement of one or two votes. No matter how many contests on the ballot have differences and no matter how many candidates in those contests have differences, the ballot as a whole has an understatement of one or two votes, or an overstatement of one or two votes, or neither an understatement nor an overstatement. (Of course, the sample might contain many ballots in each of these categories.)

If changing the interpretation of the ballot according to the voting system to make it match the human interpretation of the ballot would widen *every* pairwise margin in *every* contest under audit, that ballot has an *understatement*. If it would widen *every* pairwise margin in *every* contest by two votes, the ballot has a two-vote understatement; otherwise it has a one-vote understatement. If the ballot does not contain every contest under audit, it cannot have an understatement. Since there is an understatement only if changing the machine interpretation of the ballot to match the hand interpretation would increase *every* pairwise margin, understatements are quite rare. Understatements do not call the outcome into question, so they do not increase the sample size required to confirm the outcome.

If changing the interpretation of the ballot according to the voting system to match the human interpretation of the ballot would narrow *any* pairwise margin in *any* contest under audit, that ballot has an *overstatement*. If changing the interpretation of the ballot according to the voting system to match the human interpretation of the ballot would narrow *any* pairwise margin in *any* contest under audit by two votes, that ballot has a two-vote overstatement. No matter how many margins would be narrowed by one or two votes, the overstatement on a ballot is at most two votes, because only the maximum overstatement enters the calculations. If enough ballots have overstatements, the outcome could be wrong, so overstatements increase the sample size required to confirm the outcome.

As an example, suppose that we are auditing five contests simultaneously. Tables 1 and 2 below show two hypothetical CVRs and manual interpretations of the same ballots.

Table 1: Hypothetical CVR and hand interpretation of a ballot that contains four of five contests under audit. Overall, the ballot has an overstatement of 2 votes, because that is the largest overstatement of any margin in any of the contests.

	contest 1	**contest 2**	**contest 3**	**contest 4**	**contest 5**
CVR	undervote	winner	loser	not on ballot	winner
Hand	loser	loser	winner	loser	not on ballot
discrepancy	1 over	2 over	2 under**	1 over	1 over

**Contest 3 has an understatement of 2 votes *only* if the contest has only two candidates. If there are two or more losers in the contest (and only one winner), this contest has an understatement of only one vote, because only one pairwise margin was understated by two votes; the others were overstated by one vote. Similarly, if there are two or more winners in the contest and only one loser, this contest has an understatement of only one vote. If there are at least two winners and at least two losers, there is no understatement in this contest, because at least one pairwise margin was not affected at all by the discrepancy. Regardless, the *ballot* has an overstatement of 2 votes, because the ballot has an overstatement of 2 votes in contest 2.

Table 2: Hypothetical CVR and hand interpretation of a ballot that contains four of five contests under audit. Overall, the ballot has an overstatement of 1 vote, because that is the largest overstatement of any margin in any of the contests.

	contest 1	**contest 2**	**contest 3**	**contest 4**	**contest 5**

CVR	winner	winner	undervote	not on ballot	winner
Hand	overvote	undervote	loser	loser	not on ballot
discrepancy	1 over	1 over	1 over	1 over	1 over

To determine whether the audit can stop, enter the number of ballots in the sample with overstatements or understatements of one or two votes, then click "Calculate." If the sample size is not large enough to confirm the outcome based on the number of differences of each type observed, the value of "If no more discrepancies are observed" will be larger than the current sample size, and the value of "Estimated additional ballots if difference rate stays the same" will be greater than zero. That value is the estimated number of additional ballots that will need to be audited to confirm the outcome at the desired risk limit, assuming that the rate of one and two-vote understatements and overstatements does not change as the sample expands.

Stopping sample size and escalation

Ballots audited so far: 0

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

Estimated stopping size

Calculate Audit incomplete

If no more differences are observed: …
If differences continue at the same rate: … .
Estimated additional ballots if difference rate stays the same: …

If the contest being audited has more than two candidates or positions, the calculation above can be very conservative if overstatements do not affect the margin between the winner with the fewest votes and the loser with the most votes. The formula above can be modified to take that into account.

Hide technical notes.

P.B. Stark, statistics.berkeley.edu/~stark. http://statistics.berkeley.edu/~stark/Java/Html/auditTools.htm
Last modified 17 February 2012.

A Gentle Introduction to Risk-limiting Audits

Mark Lindeman
Philip B. Stark
Department of Statistics, University of California, Berkeley

January 6, 2012

Abstract

Risk-limiting audits provide statistical assurance that election outcomes are correct by hand counting portions of the audit trail—paper ballots or voter-verifiable paper records. We sketch two types of risk-limiting audits, *ballot-polling audits* and *comparison audits*, and give example computations. Tools to perform the computations are available at `statistics.berkeley.edu/~stark/Vote/auditTools.htm`.

1 What is a risk-limiting audit?

A risk-limiting audit is a method to ensure that at the end of the canvass, the hardware, software, and procedures used to tally votes found the real winners. Risk-limiting audits do not guarantee that the electoral outcome is right, but they have a large chance of correcting the outcome if it is wrong. They require examining portions of an *audit trail* of (generally paper) records that voters had the opportunity to verify accurately recorded their selections before they cast their votes.

Risk-limiting audits address limitations and vulnerabilities of voting technology, including difficulties ascertaining voter intent algorithmically, configuration and programming errors, and malicious subversion. Computer software cannot be guaranteed to be perfect or secure, so voting systems should be *software-independent*: An undetected change or error in voting system software should be incapable of causing an undetectable change or error in an election outcome

[Rivest and Wack, 2006, Rivest, 2008]. An audit trail provides software independence; a risk-limiting audit leverages software independence by checking the audit trail strategically.

Systems that do not produce voter-verifiable paper records, such as paperless touchscreen voting systems, cannot be audited this way. Records of cast votes printed after the fact do not confer software independence because voters had no chance to verify them.

A risk-limiting audit is an "intelligent" incremental recount that stops when the audit provides sufficiently strong evidence that a full hand count would confirm the original (voting system) outcome. As long as the audit does not yield sufficiently strong evidence, counting continues, possibly to a full hand count (as part of the audit or a separate recount). "Sufficiently strong" is quantified by the *risk limit*, the largest chance that the audit will stop short of a full hand count when the original outcome is in fact wrong, no matter why it is wrong, including "random" errors, voter error, bugs, equipment failures, or deliberate fraud. A full hand count reveals the correct outcome.

Smaller risk limits entail stronger evidence that the outcome is correct: All else equal, the audit counts more ballots if the risk limit is 1% than if it is 10%. Smaller margins—smaller differences between the vote shares of the winners and the losers—require more evidence to attain a given risk limit, because there is less room for error: All else equal, the audit counts more ballots if the margin is 1% than if it is 10%.

The risk limit is *not* the chance that the outcome (after auditing) is wrong. A risk-limiting audit replaces the original outcome if and only if it leads to a full count that disagrees with the original outcome. Hence, a risk-limiting audit cannot harm correct outcomes. But if the original outcome is wrong, there is a chance the audit will not correct it. The risk limit is the *largest* such chance. If the risk limit is 10% and the outcome is wrong, there is at most a 10% chance (and typically much less) that the audit will not correct the outcome—at least a 90% chance (and typically much more) that the audit will correct the outcome.

1.1 The audit trail

Risk-limiting audits involve counting votes in portions of the audit trail by hand. The best audit trail is voter-marked paper ballots. Voter-verifiable paper records (VVPRs) generated by touchscreen voting machines are not as good. They are less directly connected to voters' selections, since they are generated by hardware and software, not by the voter. Printers can jam or run out of paper. VVPRs can be fragile and cumbersome to audit. (As noted above, paperless touchscreen voting machines do not provide a suitable audit trail.) Below, we call entries in the audit trail "ballots," even though they might be VVPRs.

Like a recount, a risk-limiting audit assumes there is a "correct" interpretation of each ballot. Rules for interpreting ballots must be established before the audit starts.

1.2 Ballot-level audits

States that mandate hand counting as part of audits generally require counting the votes in selected *clusters* of ballots. For instance, under California law, each county counts the votes in 1% of precincts; each cluster comprises the ballots cast in one precinct.

The smaller the clusters, the less counting a risk-limiting audit requires—assuming the outcome is correct. (If the outcome is wrong, the audit has a large chance of counting all the votes, regardless of the size of the clusters.) Auditing a random sample of 100 individual ballots can be almost as informative as auditing a random sample of 100 entire precincts! Hand counting is minimized when clusters consist of one ballot each, yielding "ballot-level" audits or "single-ballot" audits. See Stark [2010a] for more discussion.

Ballot-level audits save work, but finding individual ballots among millions stored in numerous boxes or bags ("batches") is challenging. It requires knowing the number of ballots in each batch (a *manifest*, discussed below), how to locate each batch, and how to identify each ballot within each batch uniquely. Labeling individual ballots helps, but is prohibited in some jurisdictions. Ballot-level auditing elevates privacy concerns. The most efficient ballot-level audits, comparison audits (explained below), require knowing how the voting system interpreted individual ballots—which no federally certified vote tabulation system reports.

If the voting system does not report its interpretation of each ballot, one can audit using an unofficial system that does. *Transitive auditing* checks the unofficial system, rather than the system of record. If the two systems show different outcomes, all votes should be counted by hand. If the systems show the same outcome, a risk-limiting audit of the unofficial system checks the outcome of the system of record: Either both are right or both are wrong. If both are wrong, the risk-limiting audit has a large chance of requiring a full hand count. See, e.g., Calandrino et al. [2007], Benaloh et al. [2011].

2 Before the audit starts

Because a risk-limiting audit relies upon the audit trail, preserving the audit trail complete and intact is crucial. If a jurisdiction's procedures for curating the audit trail are adequate in principle, ensuring compliance with those procedures can provide strong evidence that the audit trail is reliable. This *compliance audit* should assess the integrity of the audit trail, determining whether all records were secure against loss, spoilage, and tampering. A compliance audit may be subsumed by a comprehensive post-election audit or canvass.

To sample ballots efficiently requires a *ballot manifest* that describes in detail how the ballots are organized and stored. For instance, the jurisdiction might keep cast ballots in 350 batches, labeled 1 to 350. The manifest might say "There are 71,026 ballots in 350 batches: Batch 1 has 227 ballots; batch 2 has 903 ballots; . . . ; and batch 350 has 114 ballots." If the jurisdiction numbers its ballots, the manifest might say, "Batch 1 contains ballots 1–227; batch 2 contains ballots 228–1,130; . . . ; and batch 350 contains ballots 70,913–71,026."

Auditors should verify that the number of ballots according to the manifest matches the total according to the election results. It is good practice to count the ballots in the batches containing the ballots selected for audit to check whether the manifest is accurate. If the manifest is inaccurate, the risk limit may not be correct.

3 Two kinds of simple risk-limiting audits

We present simple examples of two kinds of risk-limiting audits: *ballot-polling audits* and *comparison audits*. (Johnson [2004] makes an analogous distinction, but does not address risk-limiting audits per se.) "Simple" means that the calculations are easy, even with a pencil and paper, so observers can check the auditors' work. Tools that perform these calculations are available at `statistics.berkeley.edu/~stark/Vote/auditTools.htm`, the "auditTools page."

This section addresses risk-limiting audits of a simple single-winner contest. Section 5 discusses auditing more than one contest at once, contests with more than one winner, contests that require a super-majority, and ranked-choice voting.

3.1 Ballot-polling audits

Ballot-polling audits examine a random sample of ballots. When the vote shares in the sample give sufficiently strong evidence that the reported winner really won, the audit stops.

Ballot-polling audits require knowing who reportedly won, but no other data from the vote tabulation system. They are best when the vote tabulation system cannot export vote counts for individual ballots or clusters of ballots or when it is impractical to retrieve the ballots that correspond to such counts. Ballot-polling audits generally require far more counting than ballot-level comparison audits, described below. For a margin half as big, the expected number of ballots to be counted would approximately double in a comparison audit, but approximately quadruple in a ballot-polling audit. For a margin a third as large, the workload of a comparison audit would approximately triple, but the workload of a ballot-polling audit would increase by about a factor of nine.

The following ballot-polling audit, which relies on Wald's sequential probability ratio test [Wald, 1945], has risk limit 10%: There is at least a 90% chance it will require a full hand count if the reported winner actually lost. It assumes that the winner's reported share s of valid votes is greater than 50%: a majority rather than a mere plurality. With small changes, it applies to contests that require a super-majority. Similar procedures can be constructed for plurality winners.

1. Let s be the winner's share of the valid votes according to the vote tabulation system; this procedure requires $s > 50\%$. Let t be a positive "tolerance" small enough that when t is subtracted from the winner's reported vote share s, the difference is still greater than 50%. (Increasing t reduces the chance of a full hand count if the voting system outcome is correct, but increases the expected number of ballots to be counted during the audit.) Set $T = 1$.

2. Select a ballot at random from the ballots cast in the contest.

3. If the ballot is an undervote, overvote, or an invalid ballot, return to step 2.

4. If the ballot shows a valid vote for the reported winner, multiply T by

$$(s-t)/50\%.$$

5. If the ballot shows a valid vote for anyone else, multiply T by

$$(1-(s-t))/50\%.$$

6. If $T > 9.9$, stop the audit. The reported outcome stands.

7. If $T < 0.011$, stop the audit and perform a full hand count to determine who won. Otherwise, return to step 2.

If the reported winner's true share of the vote is at least $s - t$, there is at most a 1% chance that this procedure will lead to a full hand count; that chance can be adjusted by changing the comparison in step 7. To reduce the risk limit, we would compare T to a larger number in step 6.

As a numerical example, suppose one candidate reportedly received $s = 60\%$ of the valid votes. Set $t = 1\%$. If the reported winner in fact got at least $s - t = 59\%$ of the vote, there is at most a 1% chance that the audit will lead to a (pointless) full hand count. Note that $1 - (s - t) = 1 - 59\% = 41\%$. To audit, we repeat steps 2–7, drawing ballots at random (see section 4) and updating T until either $T > 9.9$ or $T < 0.011$.

The number of ballots we end up auditing depends on the winner's actual vote share and on which ballots happen to be selected. If the first 14 ballots drawn all show votes for the reported winner,

$$T = (59\%/50\%) \times (59\%/50\%) \times \cdots \times (59\%/50\%) = (59\%/50\%)^{14} = 10.15,$$

and the audit stops.

If the winner's true vote share is 60%, the audit is expected to examine 120 ballots; for a 55% share, 480; and for a 52% share, 3,860: The expected workload grows quickly as the margin shrinks.

When the outcome is correct, the number of ballots the audit examines depends only weakly on the number of ballots cast, so the percentage of ballots examined in large contests can be quite small. For example, in the 2008 presidential election, 13.7 million ballots were cast in California; Barack Obama was reported to have received 61.1% of the vote. A ballot-polling audit could confirm that Obama won California at 10% risk by auditing roughly 97 ballots in all—seven ten-thousandths of one percent of the ballots cast—if Obama really received over 61% of the votes.

The expected auditing workload for individual counties is proportional to the percentage of ballots cast in the county. Almost 25% of the ballots were cast in Los Angeles county, the largest of California's 58 counties. Over 75% of the ballots were cast in the largest 12 counties. The smallest 14 counties together account for less than 1% of the ballots. So, about 24 of the 97 ballots (25%) would be from Los Angeles; 73 (75%) from the largest 12 counties, including Los Angeles; and perhaps one ballot (1%) total from the smallest 14 counties.

If the winner's share were 52% rather than 61.1%, the expected number of ballots to examine would be 3,860, less than three hundredths of one percent of the ballots cast. Of those, Los Angeles would have expected to examine about 946, the largest 12 counties about 2,922 total, and the smallest 14 counties about 35 total. Since ballot-polling audits do not require data from the vote tabulation system, they are an immediate practical option for auditing large contests. Indeed, *all* statewide contests could be confirmed with a single ballot-polling audit expected to examine 3,860 ballots if the winners' smallest vote share was 52%. Comparison audits, described next, generally involve examining fewer ballots, but require much more from the vote tabulation system.

3.2 Comparison audits

Comparison audits check outcomes by comparing hand counts to voting system counts for clusters of ballots. In ballot-level comparison audits, each cluster is one ballot. Comparison audits can be thought of as having two phases. The first checks whether the reported subtotals for every cluster of ballots sum to the contest totals for every candidate. If they do not, the reported results are inconsistent; the audit cannot proceed. The second phase spot-checks the voting system subtotals against hand counts for randomly selected clusters, to assess whether the subtotals are sufficiently accurate to determine who won. If not, the audit has a large chance of requiring a full hand count.

This section is based on the "super-simple" ballot-level risk-limiting comparison audit [Stark, 2010b]. It presumes we know how the vote tabulation system (or, for transitive audits, an unofficial system) interpreted every ballot. The audit compares a manual interpretation of ballots selected at random to the system's interpretation of those ballots, continuing until there is strong evidence that the outcome is correct or until all ballots have been examined.

Suppose the manual interpretation disagrees with the voting system interpretation. If changing the voting system interpretation to match the manual interpretation would increase the margin between the winner and every loser, the ballot has an "understatement." If the voting system interpretation of a ballot records an overvote but the manual interpretation shows a vote for the winner, the ballot has an understatement. Understatements do not call the outcome into question, because correcting them widens the margin.

If changing the voting system interpretation to match the manual interpretation would decrease the margin between the winner and any loser, the ballot has an "overstatement" equal to the maximum number of votes by which any margin would decrease. If the voting system interpretation of a ballot records an undervote but the manual interpretation finds a vote for one of the losers, the ballot has an overstatement of one vote: The voting system interpretation overstated the margin by one vote. If the voting system interpretation of a ballot recorded a vote for the winner but the manual interpretation finds an overvote, that ballot has an overstatement of one vote.

If the voting system interprets a ballot as a vote for the winner while a manual interpretation finds a vote for one of the losers, that ballot has an overstatement of *two* votes. For voter-marked paper ballots, occasional one-vote misstatements are expected, owing to the vagaries of how voters mark their ballots: From time to time the system will interpret a light mark as an undervote or a hesitation mark as an overvote. But two-vote overstatements should be quite rare: A properly functioning voting system should not award a vote for one candidate to a different candidate.

We now present a simple rule for a risk-limiting comparison audit with risk limit 10%. The rule depends on the "diluted margin" m, the margin of victory in votes divided by the number of ballots cast. Dividing by the number of ballots, rather than by the number of valid votes, allows for the possibility that the vote tabulation system confused an undervote or overvote for a valid vote, or vice versa. Let n be the number of ballots in the audit sample. Let u_1 and o_1 be the number of 1-vote understatements and overstatements among those n ballots, respectively; similarly, let u_2 and o_2 be the number of 2-vote understatements and overstatements. The audit can stop when

$$n \geq \frac{4.8 + 1.4(o_1 + 5o_2 - 0.6u_1 - 4.4u_2)}{m}. \quad (1)$$

This follows from equation [9] of Stark [2010b] with risk limit $\alpha = 10\%$ and $\gamma = 1.03905$, by the same conservative approximation used to derive equation [17] there, with a bit of rounding.

Overstatements increase the required sample size and understatements decrease it, but not by equal amounts. We have more confidence in the outcome if the sample shows no misstatements than if it shows large but equal numbers of understatements and overstatements. In expression [1] a 1-vote understatement offsets 60% of a 1-vote overstatement and a 2-vote understatement offsets 85% of a 2-vote overstatement.

If the diluted margin is 10%, each 1-vote overstatement increases the required sample size by $1.4/10\% = 14$ ballots and each 1-vote understatement decreases the required sample size by $1.4 \times 0.6/10\% = 8.4$ ballots. Each 2-vote overstatement increases the required sample size by $1.4 \times 5/10\% = 70$ ballots and each 2-vote understatement decreases the required sample size by $1.4 \times 4.4/10\% = 61.6$ ballots. For a margin of 5%, these numbers double; for a margin of 2%, they quintuple.

With this method, the auditor can check one ballot at a time against its voting system interpretation sequentially or a larger number in parallel. Moreover, the auditor can decide at any point to abort the audit and finish a full hand count. The risk limit will be 10% provided the audit continues either until condition [1] is satisfied or until all ballots have been counted by hand and the hand-count outcome replaces the reported outcome.

Numerical examples might help. Suppose that 10,000 ballots were cast in a particular contest. According to the vote tabulation system, the reported winner received 4,000 votes and the runner-up received 3,500 votes. Then the diluted margin is $m = (4000 - 3500)/10000 = 5\%$. We consider sampling ballots incrementally and sampling in stages.

3.2.1 Sampling incrementally

In an incremental audit, the auditor draws a ballot and compares a manual interpretation to the voting system interpretation before drawing the next ballot. If there is one 1-vote understatement and no other misstatements among the first 80 ballots examined, $u_1 = 1$ and o_1, u_2, and o_2 are all zero and the audit can stop, because

$$80 \geq \frac{4.8 - 1.4 \times 0.6 \times 1}{5\%}. \tag{2}$$

If there are no overstatements or understatements among the first 96 ballots examined, u_1, o_1, u_2, and o_2 are all zero and the audit can stop, because

$$96 \geq 4.8/5\%. \tag{3}$$

3.2.2 Sampling in stages

To simplify logistics, an auditor might draw many ballots at once, then compare each to its voting system interpretation. If the audit needs to continue, the auditor would draw another set of ballots and compare them to their voting system interpretations. Each set of draws and comparisons is a *stage*.

If the auditor expects errors at some rate, she can select the first-stage sample size so that the audit stops there if her expectation proves correct or pessimistic. Suppose she expects one 1-vote overstatement and one 1-vote understatement per

	0 understatements					1 1-vote understatement				
	# 1-vote overstatements					# 1-vote overstatements				
margin	0	1	2	3	4	0	1	2	3	4
0.2%	2400	3100	3800	4500	5200	1980	2680	3380	4080	4780
0.5%	960	1240	1520	1800	2080	792	1072	1352	1632	1912
1%	480	620	760	900	1040	396	536	676	816	956
2%	240	310	380	450	520	198	268	338	408	478
5%	96	124	152	180	208	80	108	136	164	192
10%	48	62	76	90	104	40	54	68	82	96
20%	24	31	38	45	52	20	27	34	41	48

Table 1: Exemplar sample sizes for ballot-level comparison audits with various margins and misstatements among the sampled ballots, 10% risk limit.

thousand ballots (0.001 per ballot), and expects 2-vote misstatements to be negligibly rare. For a contest with a diluted margin *m* of at least 5%, an initial sample of $4.8/m$ ballots (rounded up) is 96 ballots or fewer. If overstatements are as infrequent as expected, there are unlikely to be any among the first 96 ballots: The audit will stop at the first stage. An initial sample of $6.2/m$ (124 ballots or fewer if the margin is at least 5%) allows the audit to stop at the first stage if it shows one 1-vote overstatement.

If the sample is sorted before checking ballot interpretations, all ballots drawn in the stage should be checked before calculating whether to stop: The first n ballots in a sorted random sample of N ballots are not a random sample of n ballots.

Table 1 gives stopping sample sizes for various margins and numbers of overstatements and understatements, for 10% risk. It can help select the first-stage sample size for different expected rates of error.

4 Random selection

Risk-limiting audits rely on random sampling. (Random samples can be augmented with "targeted" samples chosen by other means; see, e.g., Stark [2009a].) If the sample is not drawn appropriately, the risk limit will be wrong. The risk-limiting methods described above rely on drawing a random sample of ballots with replacement. This is like putting all the ballots into an enormous mixer, stirring them thoroughly, and drawing a ballot without looking. The ballot is returned to the mixer, the ballots are mixed again, and another ballot is drawn (possibly the same ballot), until the audit stops.

Public confidence requires that observers can verify the selection is fair—that all ballots are equally likely to be selected in each draw. This speaks against a number of common methods for selecting samples, including "arbitrary" selection by the election officials; drawing slips of paper, where there is no way to confirm that each ballot is represented by exactly one slip and that the slips have been adequately mixed; using a proprietary software program such as Excel; or using any source of putative randomness that cannot readily be checked.

Trustworthy methods of generating random numbers often have two features: (1) a physical source of randomness, such as dice rolls; (2) inputs from multiple parties, so that even if some parties collude, any non-colluding party could foil an attempt to rig the sample. It can be efficient, effective, and transparent to use a simple mechanical method—such as rolling dice [Cordero et al., 2006]—to generate a "seed" for a well-designed *pseudo-random number generator* (PRNG). PRNGs can generate arbitrarily many "pseudo-random" numbers from a single seed. The numbers are not truly random because a PRNG always produces the same sequence of numbers from a given seed, but they have many of the desirable properties of random sequences. And any observer who knows the seed and the PRNG can check the output. For good PRNGs, small changes in the seed yield quite different sequences, so starting with a random seed makes it effectively impossible for anyone to render the audit less effective by anticipating which ballots will be examined.

The auditTools page (described in section 3) provides a good PRNG suggested by Ronald L. Rivest. It relies on the SHA-256 cryptographic hash function, which is in the public domain and has been implemented in many computer languages. That allows observers to confirm that the sequence of pseudo-random numbers is correct, given the seed.

A ballot manifest can be used to translate random (or pseudo-random) numbers into the particular ballots comprising the sample. Before the audit, we use the manifest to assign a unique number to each ballot, if the ballots are not already marked with unique identifiers. Suppose that the manifest lists 822 ballots in three batches, numbered 1 through 3; the batches contain, respectively, 230, 312, and 280 ballots. Then we can number the 230 ballots in batch 1 ballots 1 through 230; the 312 ballots in batch 2 ballots 231 through 542; and the 280 ballots in batch 3 ballots 543 through 822. Ballot 254 is the 24th ballot in batch 2. We assume that the ballots are stored in some order that remains unchanged during the audit, so that "the 24th ballot in batch 2" uniquely identifies a particular ballot.

To draw the audit sample, we generate random numbers between 1 and 822, and retrieve the corresponding ballot. If 254 is generated, we audit the 24th ballot in batch 2.

5 More complicated situations

We have discussed only contests where the candidate with the most votes wins. The methods can be extended to audit contests that require a supermajority, contests with more than one winner, cross-jurisdictional contests, ranked-choice voting; and to audit a collection of contests simultaneously with a single sample.

Contests with more than one winner and collections of contests can be audited with a comparison audit based on the *maximum relative overstatement of pairwise margins* (MRO) [Stark, 2008b, 2009b], defined as follows. A *pairwise margin* is the margin between any reported winner and loser in a given contest. An overstatement of a pairwise margin, divided by that margin, is the *relative overstatement* of the pairwise margin. A one-vote overstatement of a wide margin casts less doubt on the outcome than a one-vote overstatement of a narrow margin; relative overstatements take this into account. The MRO is the maximum relative overstatement on each audited ballot. The arithmetic can be simplified by treating all overstatements as if they affected the smallest diluted margin. This is conservative, but if overstatements are rare, the workload remains manageable. That is the heart of the "super-simple" simultaneous audit method [Stark, 2010b].

For simultaneous audits of multiple contests, the diluted margin is the smallest margin in votes divided by the total number of ballots on which at least one of the contests appears. If a contest appears on only a small fraction of ballots, it may take less work to audit it separately, so that its diluted margin considers only the ballots that contain the contest.

Auditing contests that cross jurisdictional boundaries is straightforward if all the results are available before the audit starts, and the sample can be drawn from all ballots as a pool. If the jurisdictions draw samples independently, the computations are complicated [Stark, 2008a, Higgins et al., 2011]. Auditing instant-runoff or ranked-choice (IRV/RCV) contests is a topic of research: Even computing the "margin of victory" is difficult [Magrino et al., 2011, Cary, 2011].

6 A practical example: Merced County, California

The methods described above have been used to audit live elections in California, including the November 2011 election in Merced County. That audit, authorized by California's 2010 law AB 2023 and funded by a grant from the Election Assistance Commission, used a single sample to check the outcomes of two City of Merced contests: the mayoral contest, and the (vote-for-three) councilmember contest. In the mayoral contest, which had five candidates, the voting system reported that Stan Thurston received 2,231 votes, and runner-up Bill Blake received 2,037—a margin of 194 votes, or 2.79% of valid votes cast. In the councilmember contest, the margin of decision (between the third-place and fourth-place candidates) was wider, 959 votes.

Because Merced's voting system cannot report its interpretation of individual ballots, a transitive audit was conducted: The 7,120 cast ballots were digitally scanned. A ballot manifest was prepared. Kai Wang, Ph.D. student at the University of California, San Diego, interpreted the images using software he wrote, spot-checking "difficult" cases by hand. His vote totals were slightly higher than the official totals, but gave the same winners. The margin he found for the mayoral contest was 192 votes, a diluted margin m of about 2.70%. Before the audit started, the unofficial interpretations were posted to a website to ensure that those interpretations could not be changed during the audit.

The initial sample was large enough to confirm the original results at 10% risk limit if it revealed few overstatements. The minimum sample size if there were no misstatements would be $4.8/m = 178$. The initial sample size was chosen on the assumption that the rates of one-vote overstatements and understatements would be 0.001, rounded up to the nearest whole number, and that the rates of two-vote overstatements and understatements would be negligible. That led the auditors to anticipate one 1-vote overstatement and one 1-vote understatement in the sample. Expression [1] with $o_1 = 1$ and $u_1 = 1$ yields $n \geq (4.8 + 1.4 \times (1 - 0.6 \times 1))/0.027$, so $n \geq 198.5$. Expression [1] rounds to the nearest tenth but the auditTools page does not; the initial sample was 198 ballots. (To allow for a one-vote overstatement without any compensating one-vote understatement, the initial sample size would be 230 instead: When $o_1 = 1$ and $u_1 = o_1 = o_2 = 0$, $n \geq (4.8 + 1.4 \times 1)/0.027$, giving an initial sample size $n \geq 229.6$.)

Each person present at the audit contributed two digits to a seed. Using the PRNG on the auditTools page and that seed, the auditors generated 198 numbers between 1 and 7,120, the number of ballots. They retrieved each of the corresponding ballots using the manifest and the lookup tool on the auditTools page.

Their manual interpretation of each ballot matched Kai Wang's interpretation, so the audit stopped, transitively confirming the official winners of both contests at 10% risk limit by looking at 198 ballots.

7 Discussion

Risk-limiting audits guarantee that if the vote tabulation system found the wrong winner, there is a large chance of a full hand count to correct the results. Providing this guarantee requires a voting system that produces a voter-verifiable paper record—an audit trail—and requires the local election official to ensure that the audit trail remains complete and accurate. Risk-limiting audits examine portions of the audit trail by hand until there is sufficiently strong evidence that a full hand count would confirm the reported result, or until there has been a full hand count.

There are two general types of risk-limiting audits: *ballot-polling audits* and *comparison audits*. For a ballot-polling audit, the sample size depends on the true vote shares among valid votes; for a comparison audit, it depends on the reported margins and reported invalid votes and undervotes, and on the number and nature of errors in the original tally. For both types of audit, sample sizes depend on the luck of the draw—the particular ballots that happen to be in the sample—but only weakly on the size of the contest. Both types are most efficient when the audit checks individual ballots, *ballot-level auditing*.

Ballot-polling audits require almost nothing but the audit trail and a list of reported winners. In contrast, ballot-level comparison audits require detailed information from the vote tabulation system—its interpretation of each ballot. However, ballot-polling comparison audits examine fewer ballots than ballot-polling audits when the margin is small and the outcome is correct: The number grows like the reciprocal of the margin, versus the square of the reciprocal for ballot-polling audits. At 10% risk limit, assuming the vote tabulation system is perfectly accurate, the ballot-polling method we presented would be expected to examine 120 ballots if the winner's share is 60%, 480 if it is 55%, or 3,860 if it is 52%, versus 24, 48, and 120 for the comparison audit method we present.

Unfortunately, current commercial vote tabulation systems do not report their interpretation of each ballot, so ballot-level comparison audits sometimes rely on unofficial systems, giving *transitive audits*. Ballot-polling audits may be immediately practical for large contests, because they require so little of the vote tabulation system, and the counting burden typically is spread across many jurisdictions.

These types of audit require random samples. The samples must be drawn properly, in a way that precludes manipulation, and ideally in a way that the public can verify is proper. Using a high-quality public pseudo-random number generator with a "seed" generated at random by audit participants satisfies these requirements.

While the mathematics that underlie risk-limiting audits might be daunting, the calculations required to conduct the audit can be extremely simple: arithmetic that could easily be done with pencil and paper or a four-function calculator. Simplicity improves transparency and can increase public confidence by allowing anyone interested to check the calculations.

8 Acknowledgments

We are grateful to Jennie Bretschneider, Ronald L. Rivest, and Barbara Simons for helpful suggestions.

References

Benaloh, J., Jones, D., Lazarus, E., Lindeman, M., and Stark, P. (2011). SOBA: Secrecy-preserving observable ballot-level audits.

Calandrino, J., Halderman, J., and Felten, E. (2007). Machine-assisted election auditing. In *Proceedings of the 2007 USENIX/ACCURATE Electronic Voting Technology Workshop (EVT 07)*. USENIX.

Cary, D. (2011). Estimating the margin of victory for instant-runoff voting. In *Proceedings of the 2011 Electronic Voting Technology Workshop / Workshop on Trustworthy Elections (EVT/WOTE '11)*. USENIX.

Cordero, A., Wagner, D., and Dill, D. (2006). The role of dice in election audits – extended abstract. In *IAVoSS Workshop On Trustworthy Elections (WOTE 2006)*.

Higgins, M., Rivest, R., and Stark, P. (2011). A sharper treatment of stratification in risk-limiting post-election audits. *Working draft*.

Johnson, K. (2004). Election certification by statistical audit of voter-verified paper ballots. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=640943. Retrieved 6 March 2011.

Magrino, T., Rivest, R., Shen, E., and Wagner, D. (2011). Computing the margin of victory in IRV elections. In *Proceedings of the 2011 Electronic Voting Technology Workshop / Workshop on Trustworthy Elections (EVT/WOTE '11)*. USENIX.

Rivest, R. (2008). On the notion of 'software independence' in voting systems. *Phil. Trans. R. Soc. A*, 366(1881):3759–3767.

Rivest, R. and Wack, J. (2006). On the notion of "software independence" in voting systems (draft version of july 28, 2006). Technical report, Information Technology Laboratory, National Institute of Standards and Technology. `http://vote.nist.gov/SI-in-voting.pdf` Retrieved April 20, 2011.

Stark, P. (2008a). Conservative statistical post-election audits. *Ann. Appl. Stat.*, 2:550–581.

Stark, P. (2008b). A sharper discrepancy measure for post-election audits. *Ann. Appl. Stat.*, 2:982–985.

Stark, P. (2009a). CAST: Canvass audits by sampling and testing. *IEEE Transactions on Information Forensics and Security, Special Issue on Electronic Voting*, 4:708–717.

Stark, P. (2009b). Efficient post-election audits of multiple contests: 2009 California tests. http://ssrn.com/abstract=1443314. 2009 Conference on Empirical Legal Studies.

Stark, P. (2010a). Risk-limiting vote-tabulation audits: The importance of cluster size. *Chance*, 23(3):9–12.

Stark, P. (2010b). Super-simple simultaneous single-ballot risk-limiting audits. In *Proceedings of the 2010 Electronic Voting Technology Workshop / Workshop on Trustworthy Elections (EVT/WOTE '10)*. USENIX. `http://www.usenix.org/events/evtwote10/tech/full_papers/Stark.pdf`. Retrieved April 20, 2011.

Wald, A. (1945). Sequential tests of statistical hypotheses. *Ann. Math. Stat.*, 16:117–186.

Assembly Bill No. 2023

CHAPTER 122

An act to amend Sections 15620, 15621, 16401, and 16421 of, and to add Chapter 8.5 (commencing with Section 15560) to Division 15 of, the Elections Code, relating to elections.

[Approved by Governor July 19, 2010. Filed with Secretary of State July 19, 2010.]

LEGISLATIVE COUNSEL'S DIGEST

AB 2023, Saldana. Election results.

Existing law requires that, after an election, each county conduct an official canvass of the ballots cast in the election and report the final results to the relevant governing board and the Secretary of State.

This bill would authorize the Secretary of State to establish a postcanvass risk-limiting audit pilot program in 5 or more voluntarily participating counties for the purpose of verifying the accuracy of election results. Under the program, a participating county would conduct an audit of one or more contests in each election after the tabulation of the unofficial final results, as defined, or after completion of the official canvass for the election. The bill would require that the audit be conducted in public view and by manual tally, and would further require the Secretary of State to report to the Legislature on or before March 1, 2012, on the effectiveness and efficiency of these audits.

The people of the State of California do enact as follows:

SECTION 1. Chapter 8.5 (commencing with Section 15560) is added to Division 15 of the Elections Code, to read:

CHAPTER 8.5. POSTCANVASS RISK-LIMITING AUDIT PILOT PROGRAM

15560. (a) The Secretary of State is authorized to establish a postcanvass risk-limiting audit pilot program in five or more counties to improve the accuracy of, and public confidence in, election results. The Secretary of State is encouraged to include urban and rural counties; counties from northern, central, and southern California; and counties with various different voting systems.

(b) The pilot program described in subdivision (a) shall be conducted as follows:

(1) During the year 2011, each county that chooses to participate in the pilot program shall conduct a postcanvass risk-limiting audit of one or more contests after each election in that county.

(2) An elections official conducting an audit pursuant to this section shall do all of the following:

(A) Provide at least a five-day public notice of the time and place of the random selection of the audit units to be manually tallied and of the time and place of the audit.

(B) Make available to the public a report of the vote tabulating device results for the contest, including the results for each audit unit in the contest, prior to the random selection of audit units to be manually tallied and prior to the commencement of the audit.

(C) Conduct the audit upon tabulation of the unofficial final results or upon completion of the official canvass for the election.

(D) Conduct the audit in public view by hand without the use of electronic scanning equipment using the tally procedures established by Section 15360 for conducting a manual tally.

(3) On or before March 1, 2012, the Secretary of State shall report to the Legislature on the effectiveness and efficiency of postcanvass risk-limiting audits conducted pursuant to this section. The report shall include an analysis of the efficiency of postcanvass risk-limiting audits, including the costs of performing the audits, as compared to the 1-percent manual tallies conducted in the same election pursuant to Section 15360.

(c) An audit shall not be conducted pursuant to this section with respect to a state or multijurisdictional contest unless all of the counties involved in the contest choose to participate in the pilot program authorized by this section.

(d) For purposes of this section, the following terms have the following meanings:

(1) "Audit unit" means a precinct, a set of ballots, or a single ballot. A precinct, a set of ballots, or a single ballot may be used as an audit unit for purposes of this section only if all of the following conditions are satisfied:

(A) The relevant vote tabulating device is able to produce a report of the votes cast in the precinct, set of ballots, or single ballot.

(B) The elections official is able to match the report described in subparagraph (A) with the ballots corresponding to the report for purposes of conducting an audit pursuant to this section.

(C) Each ballot is assigned to not more than one audit unit.

(2) "Contest" means an election for an office or for a measure. "Contest" shall not include either of the following:

(A) An election for a political party central committee, as provided in Division 7 (commencing with Section 7000).

(B) An advisory election, as provided in Section 9603.

(3) "Risk-limiting audit" means a manual tally employing a statistical method that ensures a large, predetermined minimum chance of requiring a full manual tally whenever a full manual tally would show an electoral outcome that differs from the outcome reported by the vote tabulating device

for the audited contest. A risk-limiting audit shall begin with a hand tally of the votes in one or more audit units and shall continue to hand tally votes in additional audit units until there is strong statistical evidence that the electoral outcome is correct. In the event that counting additional audit units does not provide strong statistical evidence that the electoral outcome is correct, the audit shall continue until there has been a full manual tally to determine the correct electoral outcome of the audited contest.

(4) "Unofficial final results" means election results tabulated pursuant to an official canvass conducted pursuant to Chapter 4 (commencing with Section 15300) but not yet reported to the governing board or the Secretary of State pursuant to subdivision (h) of Section 15302.

SEC. 2. Section 15620 of the Elections Code is amended to read:

15620. Following completion of the official canvass and again following completion of any postcanvass risk-limiting audit conducted pursuant to Section 15560, any voter may, within five days thereafter, file with the elections official responsible for conducting an election in the county wherein the recount is sought a written request for a recount of the votes cast for candidates for any office, for slates of presidential electors, or for or against any measure, provided the office, slate, or measure is not voted on statewide. The request shall specify on behalf of which candidate, slate of electors, or position on a measure (affirmative or negative) it is filed.

If an election is conducted in more than one county, the request for the recount may be filed by any voter within five days, beginning on the 29th day after the election, with the elections official of, and the recount may be conducted within, any or all of the affected counties.

For the purposes of this section, "completion of the canvass" shall be presumed to be that time when the elections official signs the certified statement of the results of the election except that, in the case of a city election, if a city council canvasses the returns itself and does not order the elections official to conduct the canvass, "completion of the canvass" shall be presumed to be that time when the governing body declares the persons elected or the measures approved or defeated.

SEC. 3. Section 15621 of the Elections Code is amended to read:

15621. Following completion of the official canvass any voter may, within five days beginning on the 29th day after a statewide election, file with the Secretary of State a written request for a recount of the votes cast for candidates for any statewide office or for or against any measure voted on statewide. Additionally, any voter may file with the Secretary of State a written request for a recount of the votes cast for candidates for any statewide office or for or against any measure voted on statewide within five days following completion of any postcanvass risk-limiting audit conducted pursuant to Section 15560. A request filed pursuant to this section shall specify in which county or counties the recount is sought and shall specify on behalf of which candidate, slate of electors, or position on a measure (affirmative or negative) it is filed.

The Secretary of State shall forthwith send by registered mail one copy of the request to the elections official of each county in which a recount of the votes is sought.

All the other provisions of this article shall apply to recounts conducted under this section.

SEC. 4. Section 16401 of the Elections Code is amended to read:

16401. The contestant shall verify the statement of contest, as provided by Section 446 of the Code of Civil Procedure, and shall file it within the following times after either the declaration of the result of the election or the declaration of the results of any postcanvass risk-limiting audit conducted pursuant to Section 15560 by the body canvassing the returns thereof:

(a) In cases other than cases of a tie, where the contest is brought on any of the grounds mentioned in subdivision (c) of Section 16100, six months.

(b) In all cases of tie, 20 days.

(c) In cases involving presidential electors, 10 days.

(d) In all other cases, 30 days.

SEC. 5. Section 16421 of the Elections Code is amended to read:

16421. The affidavit shall be filed in the office of the clerk of the superior court having jurisdiction within five days after either the completion of the official canvass or the completion of any postcanvass risk-limiting audit conducted pursuant to Section 15560 by the county last making the declaration. In the case of an office for which candidates are certified for the ballot by the Secretary of State, or in the case of a statewide ballot measure, the superior court having jurisdiction shall be the Superior Court for the County of Sacramento.

O