Post-Election Audit Standards Working Group

Report

Evaluation of Audit Sampling Models
and
Options for Strengthening
California's Manual Count

July 27, 2007

Post-Election Audit Standards Working Group

The Working Group was comprised of the following experts in the fields of computer science, financial auditing, statistical analysis, election reform advocacy, and city and county government.

- o David Jefferson (Chairperson), Center for Applied Scientific Computing, Lawrence Livermore National Laboratory

- o Elaine Ginnold, Registrar of Voters, Marin County

- o Kathleen Midstokke, City Clerk, City of Downey

- o Kim Alexander, President, California Voter Foundation

- o Philip Stark, Professor of Statistics, University of California, Berkeley

- o Amy Lehmkuhl, Certified Public Accountant, Ueltzen & Company, LLP

Table of Contents

Executive Summary

Auditing is one of the most important tools for verifying election results and checking the accuracy and reliability of voting systems. If audits are effective, then the public can have confidence in the outcome of elections even if the voting systems used are imperfect, because the audit can detect and be used as the basis to help correct human and voting system errors.[1]

California law requires a manual count of 1% of precincts in every election. This report evaluates new models for determining the appropriate size of the sample for the manual count, examines "escalation" (i.e., expansion) of audits when discrepancies are found between the manual and machine counts, and provides options for strengthening California's manual count.

The following are important rules of thumb to keep in mind when considering the manual count:

- o The larger the random sample, the more likely it is that human and voting system errors will be detected.

- o Close races require larger random samples to determine whether errors could overturn election results.

- o Races involving a small number of precincts require a larger percentage random sample to determine whether errors could overturn election results.

The Working Group has reached a consensus that the most effective way to conduct post-election audits is to take a **risk-based approach**. The sampling model that works best for this approach is the **adjustable sample model**, where the size of the initial random sample depends on a number of factors, including the apparent margin of victory, the number of precincts, the number of ballots cast in each precinct, and a desired confidence level (e.g., 99%) that the winner of the election has been called correctly.

Unfortunately, the statistical models that exist for determining the size of the random sample do not apply to California, because they require the random sample to be selected from the state as a whole. In California, cities, counties, and special districts draw independent random samples. These models do not currently take into account overvotes and undervotes, nor do they cover multi-winner elections. Moreover, the approaches that exist give confidence in the outcome of the election only if no discrepancies are found. Still lacking is a method that can determine whether to trust the outcome of the election when small discrepancies are found.

The Working Group recommends that the Secretary of State develop a comprehensive approach to verifying election outcomes. Such an approach would involve a hand count of a minimum percentage of precincts for all races, and more precincts for close races and races involving only a small number of precincts. The approach would specify an initial sample size and how the sample should be expanded if discrepancies are found. Ultimately the approach must be able to determine whether the outcome of each race is in doubt, given the discrepancies found by the hand count.

---

[1] In the context of voting system auditability, experts describe this concept as "software independence." A software-independent voting system can be audited using a paper ballot or voter verified paper audit trail (VVPAT).

Because developing and testing such an approach for California will take a significant amount of time, this report suggests several **short-term options** for improving the existing manual count process and increasing the chances of detecting errors:

- o **Increase the manual count sample size for close races.** While this involves counting more precincts than the 1% required by law, it need only involve a manual count of the specific races on the ballot that have a narrow margin of victory – not the entire ballot in those additional precincts.

- o **Increase the manual count sample size in jurisdictions that use high-risk voting systems.**

- o **Increase the manual count sample size in small counties.** The Secretary might require small counties – where current law requires only one precinct to be manually counted – to audit at least two precincts.

- o **Consider "targeted audits" where candidates select additional precincts for the manual count in close races.** While not a statistical approach, this may increase the subjective confidence of candidates and the public and can be used *for specific races*.

- o **Escalate the manual count to investigate discrepancies** that indicate a particular race may not have been called correctly *or* that a voting system malfunctioned.

- o **Select precincts using a transparent method for generating random samples, such as rolling 10-sided dice.** Methods that rely on computers are not transparent.

- o **Establish uniform procedures for conducting the manual count**, including procedures for chain of custody; requiring tracking of undervotes, overvotes, and cancelled Direct Record Electronic (DRE) ballots; blind counting; public reporting of discrepancies; and any escalation measures taken.

- o **Increase transparency** of the manual count and post-election process by posting the dates the manual count will take place and posting discrepancy reports. Jurisdictions can increase transparency by improving public reporting and setting up election advisory committees, as for example Alameda, Humboldt, and Marin counties have done.

- o **Educate voters who use DREs that it is crucial to check that their Voter-Verified Paper Audit Trail (VVPAT) is accurate.** The VVPAT is only valuable as an auditing tool if the voter actually checks each vote printed on the VVPAT.

Increases in post-election auditing must take into account the statutory 28-day canvass period and the limited resources available to local elections officials.

Requiring a manual count of additional precincts or races beyond the 1% requirement will increase the burden on local elections officials. Many already find it difficult to complete their work in the time allotted. Some counties have reported that a new law requiring absentee (also

called "Vote by Mail") ballots to be included in the manual count has nearly doubled the time it takes to complete the manual count.

Finally, because the manual count is only one of many methods elections officials use to ensure accurate election results, the Working Group recommends the Secretary of State establish a long-term working group including a number of city and county elections officials to study and develop statewide standards to improve the post-election process as a whole. The last section of this report suggests several areas for future study.

1. California Law and the Purpose of Post-Election Auditing

California is one of a growing number of states that conduct post-election manual count audits of election results. According to a March 2007 report by electionline.org, "Case Study: Auditing the Vote," manual counts are now required in 15 states: Alaska, Arizona, California, Colorado, Hawaii, Illinois, Minnesota, Missouri, New Mexico, New York, North Carolina, Utah, Washington, West Virginia, and Wisconsin.

In California, elections officials conduct a manual count of 1% of the precincts in the election chosen at random. California's manual count law is now established in Elections Code § 15360.

> Elections Code § 15360 (a) During the official canvass of every election in which a voting system is used, the official conducting the election shall conduct a public manual tally of the ballots tabulated by those devices, including absent voters' ballots, cast in 1 percent of the precincts chosen at random by the elections official. If 1 percent of the precincts is less than one whole precinct, the tally shall be conducted in one precinct chosen at random by the elections official. In addition to the 1 percent manual tally, the elections official shall, for each race not included in the initial group of precincts, count one additional precinct. The manual tally shall apply only to the race not previously counted. Additional precincts for the manual tally may be selected at the discretion of the elections official.
> (b) If absentee ballots are cast on a direct recording electronic voting system at the office of an elections official or at a satellite location of the office of an elections official pursuant to Section 3018, the official conducting the election shall either include those ballots in the manual tally conducted pursuant to subdivision (a) or conduct a public manual tally of those ballots cast on no fewer than 1 percent of all the direct recording electronic voting machines used in that election chosen at random by the elections official.
> (c) The elections official shall use either a random number generator or other method specified in regulations that shall be adopted by the Secretary of State to randomly choose the initial precincts or direct recording electronic voting machines subject to the public manual tally.
> (d) The manual tally shall be a public process, with the official conducting the election providing at least a five-day public notice of the time and place of the manual tally and of the time and place of the selection of the precincts to be tallied prior to conducting the tally and selection.
> (e) The official conducting the election shall include a report on the results of the 1 percent manual tally in the certification of the official canvass of the vote. This report shall identify any discrepancies between the machine count and the manual tally and a description of how each of these discrepancies was resolved. In resolving any discrepancy involving a vote recorded by means of a punchcard voting system or by electronic or electromechanical vote tabulating devices, the voter verified paper audit trail shall govern if there is a discrepancy between it and the electronic record.

California law also requires that the audit be a "manual" count conducted in a "public process" and that precincts be selected "at random." (Elec. Code § 336.5) The purpose stated in law is "to verify the accuracy of the automated count." (Elec. Code § 336.5) Since the law took effect more than 40 years ago, there have been vast changes in the election process in California. None of the prescored punchcard and lever systems in use in 1965 are still in use today. They have been replaced by an array of complex, computerized systems including precinct-based and centralized optical scan systems, ballot marking devices, DREs augmented with VVPATs, and with the InkaVote system used in Los Angeles County. There has also been an enormous increase in the number of votes cast by mail. The complexity of these systems means there are

many more ways in which voting systems can fail to capture votes correctly, lose votes, miscount votes, and be manipulated to yield incorrect results.

Recent changes have improved the effectiveness and transparency of the audit. In 2005, a law was enacted requiring the VVPAT to be used as the ballot of record when conducting the manual count of DRE votes (Elec. Code § 19253). In 2006, a law was enacted requiring Vote by Mail ballots to be included in the 1% manual count (Elec. Code § 15360(a)). Also in 2006, laws were enacted requiring elections officials to select the precincts for the manual count using a random number generator and to provide a five-day notice of when and where the precinct selection and the manual count will take place (Elec. Code § 15360(c)-(d)). Beginning January 1, 2008, a new law will take effect requiring elections officials to report to the Secretary of State the manual count results for each election, including details about discrepancies and how they were explained or resolved (Elec. Code § 15360(e)).

Post-election auditing helps verify that the complex voting systems in use today are accurately recording and tallying ballots and that the winners of each election contest are called correctly. However, effective post-election auditing can also serve as a tool for:

o   Detecting voting system problems that may require further investigation;

o   Finding common errors voters make in marking ballots that might be avoided in future elections with voter education;

o   Identifying human errors in the tally process that can be remedied by new procedures in future elections;

o   Deterring fraud; and

o   Providing statistical data to improve future elections or audits.

As a result of the increasing sophistication of voting systems, the current flat 1% percent manual count is no longer sufficient for confirming election results and checking voting system accuracy and reliability.

2. Auditing Principles

In evaluating post-election audit models, it is helpful to examine how basic principles of financial auditing may apply to elections. Recent changes in the financial auditing industry have adopted a renewed focus on risk-based auditing. Below is a discussion of how each of the eight core financial auditing principles may be applied to post-election auditing.

Assess Risk.   The auditor must examine risks inherent in the election process that could cause an elections official unknowingly to certify incorrect election results. Risk assessment should examine the entire election process, including the election environment, risks inherent in the

voting technology used, internal control procedures, chain of custody requirements, and the separation of duties.

Determine Materiality. An omission or misstatement is *material* if it could change the outcome of an election. As discussed in Section 3 of this report, the materiality of an error is determined by the calculated confidence level that election results are correct. Materiality should be determined statistically to detect errors that could change the outcome of an election.

Design Audit Procedures to Address Identified Risks. Audit procedures should be designed to target risks identified during risk assessment. The following are two examples: 1) if a particular piece of voting technology is determined to be high-risk, audit procedures should be designed to address the identified risk, and 2) if the margin of victory in a particular race is close, audit procedures should be designed to detect small errors, because in a close race even small errors could change the outcome of the election.

Identify Audit Evidence. Audit evidence obtained from independent sources is more reliable. In all cases, audit evidence examined in the post-election audit should be the voter-prepared optical scan ballots and VVPATs. Substituted evidence, for example the substitution of a ballot image for a jammed VVPAT, should be made on a case by case basis and only when absolutely necessary.

Perform an Independent Audit. Financial auditors are required to be independent of the entity (including management and key employees) being audited. The purpose of independence is to allow the auditor to be objective and to create confidence in the results of the audit. Ideally, independence in post-election auditing can be achieved by employing independent vote counters and audit supervisors who have the ability to escalate audit issues and discrepancies. The Working Group encourages the development of procedures that will build more independence into the post-election audit process (see also Section 6). For example, vote counting in the post-election audit should be performed independent (i.e., "blind") of the summary results, meaning the counters should not know the results of the initial count in advance of the manual count.

Report Results. As noted in Section 1, a new California law will require elections officials to report manual count results, including discrepancy reports, to the Secretary of State. As detailed in Section 3, complete reporting of post-election audit results should ideally include the confidence level that each election was called correctly, given the discrepancies uncovered by the audit.

Develop Standard Audit Principles. The financial auditing industry has developed auditing principles and standards referred to as Generally Accepted Auditing Principles (GAAP) and Generally Accepted Auditing Standards (GAAS). Similar principles and standards should be developed for the post-election audit process to ensure that all counties are performing and reporting the post-election audit consistently and accurately.

Peer Review. The financial auditing industry has instituted a peer review process. This process requires certain auditing and accounting professionals to undergo a review of their work by an independent auditor. Elections officials should consider a peer review process in which elections

officials from one jurisdiction review the post-election audit work of another jurisdiction. Results of peer reviews should be reported to the Secretary of State. The peer review process should be followed by collaboration and development of procedures and best practices.

3. Sampling Models: Fixed and Adjustable Sample Sizes and Their Variations

This section discusses various ways to select the number of precincts to audit to assess the accuracy of elections. The Working Group reviewed a number of papers that examine this issue (see Bibliography), and in particular the Working Group has benefited from a draft of the soon-to-be-released report available from the Brennan Center (Norden, *et al.*) which examines in detail the strengths and weaknesses of several approaches. The Working Group also interviewed by phone a number of experts (Doug Chapin, electionline.org; Gary Poser, Office of the Minnesota Secretary of State; Professor Ronald L. Rivest, MIT; Roy G. Saltman; and Professor David Wagner, University of California, Berkeley) and elections officials (Conny McCormack, Los Angeles County, and Carolyn Crnich, Humboldt County), and heard presentations at a public meeting on July 2, 2007, from other experts (Professor David Dill, Stanford University; Joseph L. Hall, University of California, Berkeley), elections officials (Elaine Larson, Santa Clara County; Deborah Seiler, San Diego County; and John Tuteur, Napa County) and members of the public.

The Working Group and virtually all experts and presenters agree that the ideal audit procedure does not use a flat 1% random sample of precincts as current law requires. Instead, the ideal approach is based on a statistical random sampling procedure designed to control the level of *risk*. The risk to measure statistically – and require to be very small – is that of declaring the wrong winner in a contest. However, detecting and correcting problems with voting technology and deterring and detecting fraud are also key goals. Any approach to auditing should include not only an initial sampling procedure, but also rules for *escalating* scrutiny or taking other action if and when discrepancies between machine and hand counts are found, as discussed in Section 4. Section 3 focuses on how to sample precincts to manage risk. Section 5 discusses methods to make the audit process *transparent* (i.e., open and observable by candidates and the public).

To guarantee that the risk of declaring the wrong winner in a race is very small, the number of precincts that must be manually audited depends on a variety of factors, including the following:

1. The number of precincts in the race.
2. The manner in which the random sample of precincts is drawn. (It matters whether the sample is drawn at random from all precincts in the race, or separately from each jurisdiction involved in the race.)
3. The margin in each precinct.
4. The number of ballots cast in each precinct in the race, including overvotes and undervotes.
5. The number and nature of discrepancies between machine counts and manual counts found in the precincts that are audited.

6. Additional assumptions about the maximum possible number of discrepancies in each precinct.

Setting rules for determining the number of precincts to count by hand is only one part of a comprehensive audit approach. Elections officials also need to know what to do when the manual count reveals discrepancies. When can one be confident that errors in the machine count, whatever their source, are too small in the aggregate to change the election outcome? When should one doubt the election outcome and count more precincts – or all precincts – by hand? What constitutes evidence that the voting technology has a serious flaw? What should be done if such a flaw is detected? What constitutes evidence of fraud?

There is a growing literature on the mathematics of risk-based auditing of this kind. Unfortunately, as discussed below, there is as yet no satisfactory theory for determining the sample size required to be confident that errors are, in the aggregate, too small to affect the outcome of the election. The papers the Working Group has studied have a number of limitations, and it would be premature to focus on any one of them as the approach to setting the sample size for risk-based auditing. The major shortcomings the Working Group has seen are:

- o The literature does not frame the statistical problem in the best way:

    - Most of the papers address essentially this question:
      "If the machine count named the wrong winner, what is the chance we will see at least one error in the sample?"

  However, the Working Group believes the right question to ask is:
  "If the machine count named the wrong winner, what is the chance we would have seen more errors in the sample than we actually saw in the sample?" These questions are related. The answers to both depend on the number of precincts, the number of ballots cast in each precinct, the preliminary margins in each precinct, and other variables. The second question also involves the discrepancies that are observed. (See Section 3.2.3.)

- o None of the papers in the literature deal with the key fact that many races, including all statewide races, cross jurisdictional boundaries, whereas auditing is currently done independently by the jurisdictions. Because the published formulae assume that the sample is taken at random from all precincts in the race, rather than separately by jurisdiction, the sample sizes they dictate are not correct for races that cover two or more jurisdictions. This can make a major difference in the efficacy of the audit. A race covering two counties can have a razor thin margin when considered as a whole, but be very lopsided in opposite directions in the two counties considered separately. Such a situation might call for a very large sample because of the thin overall margin, but each jurisdiction separately will consider the margin within its boundaries as very large, and hence both will call for only small samples

- o None of the papers takes proper account of overvotes and undervotes or addresses multi-winner races in which voters choose *m* out of *n* candidates.

There is still work to do to get the mathematics right, and then further work to make it understandable to the public and feasible given the practical constraints elections officials face.

For these reasons the Working Group does not recommend *at this time* adopting rules to select the sample size based on numerical target of the risk.

Still, the Working Group suggests that any short-term measures taken to improve the audit process be designed with risk in mind, while remaining consistent with the statutory requirement of a 1% minimum random audit.   In considering short-term measures, it's important to keep in mind that – all other things being equal:

a) Races involving a smaller number of precincts generally need a larger *percentage* of precincts to be audited than races involving a larger number of precincts do; and

b) Races with slim apparent margins of victory need larger numbers of precincts to be audited than races with large margins of victory.

Next is a brief discussion of approaches to determining the number of precincts to count manually and some of their merits and shortcomings.

3.1 Fixed Number or Percentage of Precincts

In this approach, a fixed number of precincts in each jurisdiction, selected at random, are tallied manually. This is essentially California's current practice and appears to be the most common approach among states that perform post-election audits.

California law requires the entire ballot in each precinct in the 1% sample to be counted manually. (For the extra precincts added to ensure that every race is represented in the sample, only the otherwise unrepresented races need to be counted by hand.) Other states with fixed percentage audit laws require a specific number of races on the ballot to be manually counted. For example, Hawaii law requires a 10% audit of all precincts using DRE voting systems, however, the law only requires one race on each ballot to be manually counted. In Illinois, elections officials must audit 5% of precincts, but the law permits a hand count or a machine count audit. Humboldt County, California, manually audited 10% of precincts in 2006 within the 28-day canvass period. Connecticut recently conducted a pilot project involving a manual count of 20% of its 87 precincts that used an optical scan voting system (not 20% of all precincts). The Connecticut Legislature is considering a bill that would make the 20% manual count a statewide requirement in 2008.

Merits: Auditing a fixed number of randomly selected precincts is a good way to detect errors that affect a large percentage of precincts. It also imposes a predictable burden on jurisdictions, which makes it easy for elections officials to plan for the cost, space, staffing, and time needed to conduct the audit.

Shortcomings: Auditing a small percentage of precincts is not effective for finding problems that affect only a few precincts. Moreover, no fixed percentage (short of 100%) suffices to give high confidence that the apparent outcome of the election is correct. For that goal, the number of precincts that should be tallied manually depends on the margin in each precinct, the number of

ballots cast in each precinct, and other factors, including the number of discrepancies found in the precincts that are manually counted.

The forthcoming report from the Brennan Center (Norden, *et al*) recommends that states that audit a fixed percentage of precincts strengthen their audits by establishing clear escalation procedures, elevating the size of the audit in close races either by using targeted audits (see below) or requiring full recounts for very close races.[2] The study notes that Connecticut conducts a full *machine* recount for any race where the margin of victory is less than 20 votes or less than 0.5%. Arizona requires a full machine recount for races where the margin is less than 0.1%.

*Recommendation***:** The Working Group agrees that it is desirable to keep the manual count of 1% of precincts as a minimum requirement. However, there are practical and worthwhile ways to strengthen the manual count.

One way to strengthen the audit is to require counties to manually count a minimum of two precincts, even when one precinct would satisfy the 1% law. This suggestion is based on the principle that when the number of precincts in a race is small, a larger percentage of precincts needs to be audited to attain the same confidence that the machine count identified the true winner. For close races, the number of precincts could be increased. (See Sections 3.2.1-3.2.3 and 3.3.3.) If the manual count reveals a sizeable discrepancy in one or more races, the number of precincts could be increased. In any event, rules eventually need to be devised to decide, given the sample size and the discrepancies found, whether to trust the outcome of the election or to manually count more precincts to determine whether to trust the outcome.

One way to simplify the 1% manual count is to allow the Vote by Mail ballots to be sampled and audited separately from the ballots cast in polling places. Vote by Mail ballots could be audited in batches that do not correspond to precincts. The same 1% or greater sample size would be required, along with the same expectation of transparency and observability, and the same escalation procedures, but the Vote by Mail ballots would not have to be sorted into precincts for purposes of the audit.

3.2 Adjusting the Number of Precincts to Audit

There are a number of approaches to linking the number of precincts to manually count to the margin of a race and other variables. These approaches include "tiered audits," such as the one Congress is considering this year in H.R. 811 (Holt) in which the number of precincts to manually count depends only on the margin, as well as approaches such as SAFE where the

---

[2] Most states do not specify what to do when discrepancies are found. One state that does is Minnesota, where officials conduct a manual audit of two to four randomly selected precincts per county for three races (governor or president, U.S. Senate, and U.S. Representative). Minnesota law requires escalation to three additional precincts if there is a discrepancy rate greater than one-half of one percent (0.5%). If that discrepancy rate persists, the law mandates a full 100% recount. While this law is statistically arbitrary, it is appealing because it focuses on close races and it sets a statewide standard rather than leaving it up to local elections officials to decide what to do when discrepancies are found.

number of precincts to count depends on the margin and the number of precincts in the race, the minimum acceptable probability of detecting an outcome-changing amount of error, and other variables and assumptions. (McCarthy *et al*., 2007)

Generally, the goal of varying the number of precincts to audit is to ensure that there is a big chance that the manual count will find some error if the total error is large enough to change the outcome of the election.

3.2.1 Tiered Audits

H.R. 811 offers an example of a tiered audit. It would require that for all *federal* elections, the percentage of precincts that are manually counted be determined by the margin of victory in a race, as follows:

Margin of Victory	Percentage of Precincts
5% or greater	3%
Less than 2%	5%
Less than 1%	10%

Merits: This is a higher level of auditing than California law currently requires, and thus has a higher chance of discovering error. Even though the levels are somewhat arbitrary, it is at least consistent with the risk-based principle that slimmer margins require larger samples to achieve the same level of confidence. Moreover, it is transparent and easy to explain and implement. The fact that Congressional districts usually have approximately the same number of voters somewhat ameliorates the weakness that this audit schedule does not depend on the number of precincts in a jurisdiction.

Shortcomings: Any tiered audit schedule has a number of shortcomings:

o The rule is ad hoc. There is no statistical justification for the particular tiers; moreover, statistical theory dictates that the number of precincts in the sample – rather than the percentage of precincts in the sample – matters most. This can make a big difference in statewide races for U.S. Senate and President because the states vary tremendously in the number of voters.

o The rule is inefficient in some cases and inadequate in others. In some races, it requires far more auditing than is needed to discover any errors that might affect the outcome; in others, it does not require enough auditing. If there are many precincts in a race, a random sample of a smaller percentage of precincts suffices to determine whether there is enough error to change the outcome of the election. If there are very few precincts in the race, a random sample of a larger percentage of precincts is needed for the same degree of confidence in the outcome.

o The rule is incomplete: It does not say what to do when discrepancies are discovered. It does not tell whether to trust the outcome of the election, or to audit a larger number of precincts, or to do a 100% manual count.

Summary: The Working Group does not believe that tiered audits are the best way to select the sample size, especially in non-federal races where the number of precincts involved in a race can vary by factors of 1,000 or more. However, tiered audits can provide a higher level of scrutiny than a flat percentage, and they are easy to explain, so they might be a reasonable short-term measure while better methods – such as those discussed in Section 3.2.3 – are developed.

3.2.2 Adjusting Sample Size to Detect Error

The Working Group reviewed several papers that determine the sample size needed to guarantee that if the total error could change the outcome of an election, the chance that the sample will reveal one or more errors is at least 90% (or any other desired minimum). (McCarthy *et al.,* 2007). The number of precincts required to ensure this depends on many factors, for example:

- How the precincts are selected (e.g., at random from the state as a whole or at random from jurisdictions separately);

- The margin of victory, possibly separately in each precinct;

- The number and sizes of precincts (separately for each jurisdiction if the precincts are drawn at random from separate jurisdictions);

- The desired probability (e.g., 95%) that the audit will detect an error if, in the aggregate, the error is large enough to change the outcome; and

- Simplifying assumptions, such as the assumption that the same percentage of precincts will be drawn from each jurisdiction, the assumption that at most 20% (or some other percentage) of the reported votes could have been counted incorrectly, the assumption that there are no undervotes or overvotes, the assumption that all precincts are the same size, and so on.

The underlying mathematics can also be used to find the chance that any given size random sample of precincts – such as the 1% mandated by law – will turn up at least one discrepancy if the total error is large enough to change the outcome of a race. However, existing formulae apply only to races that are confined to a single jurisdiction and do not account for undervotes and overvotes. In close races – measured by the numbers and sizes of the precincts involved, as well as the margin – a 1% audit can have much less than a 50% chance of detecting any error even if the total errors were large enough to change the outcome.

Merits: Adjusting the sample size to ensure that there is a minimum probability of detecting an election-altering error provides a good *minimum* level of sampling, which can be much higher than the 1% sample California law mandates for small races and for close races. If the sample size is smaller than these approaches dictate, then even if no discrepancies are found by the manual count, one cannot be confident the outcome of the election was correct.

Shortcomings: However, despite the popularity of this approach the Working Group found several problems:

- It does not impose a minimum level of scrutiny for all races. The Working Group believes that California's 1% manual count law should remain in place as a minimum

level of auditing, with additional auditing if the outcome of the election is in doubt. Adjusting the number of precincts to detect error starting with a 1% floor remedies this shortcoming.

o California law requires elections officials to complete the canvass, including the post-election manual audit, within 28 days. The expense, staffing and facilities needed to conduct timely audits are less predictable when the amount of auditing depends on the margin in each race, and in some cases may be considerably larger than required for a 1% sample.

o If the sample size is at least as large as theory dictates and the manual count turns up no discrepancy, then one can be confident that errors are not large enough in the aggregate to change the election outcome. However, manual counts routinely turn up one or more discrepancies. Then, this approach does not show whether the outcome of the election is in doubt. In fact, the formulae do not use the observed discrepancies at all.

o Because this approach does not say what to do when discrepancies occur, it is incomplete. What is needed is a rule that tells whether to trust the outcome of the election, to distrust the outcome (and require a full recount), or to audit more precincts to tell whether the outcome should be trusted.

o The published formulae do not apply to all California elections, only to elections that do not cross jurisdictional lines, because the formulae are derived from the assumption that precincts are chosen at random from all precincts in the race. In California, precincts are chosen at random separately in each jurisdiction. In principle, formulae and algorithms could be devised for races that cross jurisdictional lines, such as statewide, Congressional, Legislative, and special district races, but 1) such algorithms do not yet exist, 2) they will likely be very complex, 3) they will require additional simplifying assumptions, and 4) they will still not tell us whether to trust the outcome, if any discrepancies are found.

o Published formulae in many cases depend on ad hoc assumptions, such as assumptions about the maximum amount of error in each precinct. They also omit undervotes and overvotes. These shortcomings are easy to remedy, but revised formulae have not been published and vetted.

Recommendation: The Working Group does not believe that this is the best way to select sample size. The method provides a good *minimal* level of sampling, but if it is used, it should be used with a floor on the sampling rate, such as the 1% sample that California law currently requires. The approach needs to be supplemented with rules to determine what to do when discrepancies are found. If any discrepancies are found, the approach does not give confidence that the election outcome is correct. To have confidence in the outcome when discrepancies are found, an even larger sample of precincts needs to be counted manually (see Section 3.2.3). Further research is needed to modify this approach to apply to California elections that cross jurisdictional lines.

3.2.3 Adjusting Sample Size to Estimate the Total Error

The Working Group agrees that adjusting the sample size to detect error (Section 3.2.2) is a step in the right direction. However, the Working Group also agrees that what elections officials, the Secretary of State, and voters need to know is: *"Should we trust the election preliminary outcome, given the number of discrepancies that were discovered in the hand count?"* The methods of Section 3.2.2 – suitably modified to take into account how California draws its random sample of precincts – answer "yes" only if the manual count does not turn up any discrepancy whatsoever. Unfortunately, in practice, manual counts find some discrepancies.

What is needed is a new approach to determine whether there is, for example, 99% confidence that the total discrepancy could not change the election outcome when discrepancies are found in the sample, and a rule for expanding the sample until either 1) there is 99% confidence that the errors are not large enough in the aggregate to account for the margin or 2) all the results have been recounted.[3]

If one can say with 99% statistical confidence that the total number of discrepancies is not large enough to change the election outcome, one could conclude that either 1) the outcome is correct or 2) a one-in-one-hundred sampling event has occurred. If one cannot say with 99% confidence that the error is not big enough to change the election outcome, then one should collect more data and try again. The methods of Section 3.2.2 (suitably modified for California) give a minimum sample size as a starting point: if the sample size is too small, one cannot be 99% confident that error does not account for the margin, even if the manual count finds no discrepancies at all.

Merits: This new approach has the merit of answering the most important question: *"Should we trust the outcome of the election?"* It takes into account the discrepancies that are found in the hand count and automatically increases the number of precincts audited when confidence in the outcome is low. It also provides a statistically justifiable procedure for repeatedly increasing the size of the audit sample until one can answer either *"Yes, we should trust the results now."* or until 100% of the ballots have been manually counted, whichever comes first.

Shortcomings: Unfortunately, there is as yet no formula or algorithm to implement this approach for California elections. There are methods from statistical auditing that would apply if precincts were drawn at random from California as a whole, with probability of selection proportional to the number of ballots cast in each precinct. Those methods might be adaptable to the full complexity of California elections, but the result would be complex and would require software to implement. The sample size selection would not be transparent – although in principle anyone could check that it is correct. Research on this approach is in its infancy.

Recommendation: The Working Group recommends that the Secretary of State support further research on this approach. This new approach is the most promising for providing a real answer

---

[3] In statistical language, we want a sequential test of the hypothesis that the election actually came out differently from the machine counts. If we can reject the hypothesis that the second-place candidate was the real winner (at significance level 1%, for the sake of argument), we can be confident that the election result is correct. If we cannot reject that hypothesis, we should continue to manually count more precincts.

to the question of whether to trust the outcome of an election and what to do when the manual count uncovers discrepancies.

3.3 Other Approaches

3.3.1 Complete Recount for Close Races

The Secretary of State might consider requiring an automatic 100% audit in races with small margins, either a hand recount or a machine recount using a different machine (in the case of optical scan ballots). Connecticut automatically conducts a full *machine* recount (not a hand recount) for any race where the margin of victory is less than 20 votes or less than 0.5%. Arizona performs a machine recount for races with margins of less than 0.1%.

Merits: Having a complete hand or machine count for close races adds confidence that the correct winner was named.

Shortcomings: If the race includes many precincts, auditing a sufficiently large random sample would suffice to give high confidence in the outcome of the election. Moreover, it is not currently possible (or even desirable) to count DRE votes by machine in a meaningful way. Changes to VVPATs to allow them to be scanned optically is a long-term possibility.

Recommendation: This approach might be useful for optical scan ballots.

3.3.2 Additional Auditing of High-Risk Machines or Voting Systems

If a particular voting system is discovered to have security risks or other weaknesses that cannot be resolved before the election, the Secretary of State could require additional hand-count auditing of a certain percentage of the high-risk machines used for casting and tallying ballots, above and beyond the minimum fixed percentage audit of election results and any other increase in the audit for races with small margins. To be complete, such a rule must also say what to do if discrepancies are found.

Merits: This approach targets effort where risk is greatest, augmenting the 1% manual count.

Shortcomings: There are no obvious rules for how much additional hand counting to perform. The approach would have to be supplemented with rules prescribing what to do if the manual count finds discrepancies.

Options: Additional scrutiny of high-risk technology is justified, but the Secretary of State will need to make ad hoc rules to set the level of additional scrutiny and will need to say what to do if discrepancies are found.

3.3.3 Targeted Selection of Additional Precincts

Some counties in California allow a "targeted audit" where candidates in narrow races select one or more precincts to be manually counted in addition to the 1% random sample. Because these precincts are chosen deliberately, rather than at random, it is difficult to say how this would increase the statistical confidence. However, it should increase the comfort of candidates, their supporters, and the public that the election results are correct. It would allow candidates to select precincts that look, for any reason, suspicious to them, even if those precincts are not selected for random audit. Deliberate selection of additional precincts could be used to expand the audit when the margin is small, and to expand the hand tally if the 1% random sample reveals discrepancies. However, it must be supplemented by a rule to determine what to do if the "suspicious" precincts actually do have a higher rate of discrepancies, and more generally, when to trust the election outcome after the targeted precinct counts, when to doubt the outcome and order a complete manual count, and when to expand the random sample further to determine whether to trust or distrust the outcome.

Merits: The additional burden of hand tallying a few precincts is small, but the increase to the comfort of the public and the candidates that the outcome was correct could be large. There is ample precedent for this process in California and in other states.

Shortcomings: The approach does not translate directly to a statistical statement of confidence. There is no obvious rule for determining when the candidates may select additional precincts, nor how many they should be allowed to select. The approach is incomplete unless there is a rule for deciding whether to trust the outcome of the election when discrepancies are found.

*Recommendation***:** The Working Group agrees that a step in this direction would be an economical way to increase public confidence in elections, provided it is supplemented by clear rules dictating what to do when the manual count finds discrepancies.

3.4 Changing the Unit of Audit

The approaches discussed so far assume that it is *precincts* that are drawn at random for auditing. But statistical theory shows that if it were possible to audit ballots in smaller groups, fewer ballots would need to be audited to attain the same statistical power to detect election-altering errors.

3.4.1 Ballot Sampling

If elections officials could directly audit a random sample of all *ballots* cast rather than auditing all the ballots in randomly selected *precincts*, the sample could be much smaller and still have the same chance of detecting an election-altering amount of error. However, there are three barriers to this method: 1) As in the polling audit model, in order to select a random sample of ballots cast, each ballot would have to be marked with a unique identifier which could compromise the anonymity of the ballot and privacy of the voter; 2) It would be difficult logistically to find and pull the ballots needed for the sample once random numbers are chosen

for the sample; and 3) Many more random numbers would have to be generated for the selection of a ballot sample than for a precinct sample, and transparent methods for generating random numbers are relatively slow. (See Section 5.)

It may be easier for elections officials to choose a precinct and manually count all the ballots in that precinct than it is to select a relatively small sample of ballots cast across various precincts, find those ballots in the stacks of all the ballots cast, gather and manually count those ballots, and then return them to the proper precinct boxes – and do all this in a secure and publicly-verifiable manner.

With Vote by Mail ballots, however, ballot sampling might be done efficiently in the future if the privacy concerns and logistical barriers noted above are removed Vote by Mail ballots might be gathered centrally and 1% chosen at random. Auditing Vote by Mail ballots in this manner might allow officials to sample ballots from throughout their jurisdiction, which could help identify errors in the machine count.

Merits: A smaller number of ballots need to be hand counted to get the same level of confidence that the election outcome is correct. The logistics of auditing Vote by Mail ballots could be simplified considerably.

Shortcomings: This approach is not currently feasible, and methods to make it feasible raise privacy concerns and logistical concerns (though not as much for Vote by Mail ballots). There still need to be rules to decide whether to trust the outcome of the election, given the discrepancies that are found. Further, as this method is a count of ballots rather than precincts, there are no totals to which to compare the sample count of ballots.

3.4.2 Polling Audit Model

The polling audit model involves sampling a small number of ballots in *every* precinct in the election. The model ideally requires the hand count to be completed in the precinct polling station on election night or the next day and matched against the machine count results for that precinct. This approach is not currently feasible, but is attractive because it distributes the work across precincts and requires less effort overall.

Merits: An overall *smaller* number of ballots need to be hand counted to get the same level of confidence that the election outcome is correct as compared to a precinct-based audit. Moreover, the burden is distributed across precincts.

Shortcomings: The approach is not currently feasible, and methods to make it feasible raise privacy concerns and logistical concerns. (See Section 3.4.1.) The approach is harder for the public to monitor because it is decentralized. There still need to be rules to decide whether to trust the outcome of the election given the discrepancies that are found.

3.5 Machine-Aided Audit

The audit models discussed in this report are based on checking election results and voting system function using a hand count. There has been some research into the effectiveness of auditing using a second machine count; however, experts have suggested machine audit results should themselves be manually audited. In addition, VVPATs currently generated by DREs are not machine-readable. (Some VVPATs may have barcodes, which are machine readable. But barcodes are not voter-verified, so they cannot be used for auditing.) For these reasons, using a separate machine count to audit the machine results from DREs is not currently an option for streamlining post-election audits in California; however, it might conceivably work for optical scan ballots.

Merits: A machine audit would save labor and time. Larger numbers of ballots could be counted a second time than with a manual process.

Shortcomings: "Independent" machines would be required, and they would themselves need to be audited. It is not currently possible to count VVPATs independently in an automated way; the method is best suited to optically scanned ballots.

4. Discrepancies and Escalation of the Audit

California law gives elections officials the authority and discretion to add precincts to the 1% manual count (Elec. Code § 15360(a)), and as noted in Section 1, beginning January 1, 2008, the law requires that those discrepancies and their resolution be reported to the Secretary of State. However, the law provides no specific requirements and standards for escalation of the manual count when discrepancies are found. The Working Group believes that uniform, technology-specific standards for the treatment of discrepancies and the escalation of the manual count should be developed and applied in California.

It is crucial to determine whether the discrepancies discovered during the manual count are so few that it is implausible that the total number of discrepancies – including those in unexamined precincts – could alter the outcome of the election. As discussed in Section 3, if the size of the sample is too small, then even if no discrepancies are found in the precincts that are hand counted, there could easily be enough errors overall to alter the election outcome. Thus, the selection of the initial sample size, the discovery of discrepancies, escalation of the manual count and confidence in the election outcome are inextricably intertwined.

In thinking about discrepancies, it is important to remember that the discrepancies found in the manual count of a random sample of precincts are not all the errors: the unexamined precincts may also have errors. Errors in the random sample of precincts are expected to occur at roughly the same rate as errors in the unsampled precincts.

4.1 Classification of the Causes of Discrepancies

There are at least four general causes of discrepancies between the machine and manual counts during auditing. In almost all cases these are errors and bugs, but fraud is always a possibility.

1. *Voter error:* Voters may mark optical scan ballots incorrectly, for example by using the wrong kind of pencil or ink, failing to fill in an oval completely, attempting to erase one vote and make a different choice, indicating their votes by circling their choices instead of coloring in an oval, or through smudging or staining Vote by Mail ballots, etc. With degraded ballots a voter's intent may be clear to the human eye, but the machines may misread the vote or misclassify it as an undervote or overvote.

2. *Official error:* Elections officials may mix up ballots from different precincts, misplace one or more ballots, miscopy an unreadable ballot, or make programming errors that result in misinterpreting scanned ballots; and poll workers may combine ballots from different precincts into one box.

3. *Voting system error:* Machine errors are technology-specific. DREs and optical scan machines may be miscalibrated, counts stored on memory cards may be corrupted, or software bugs may cause some ballots to be misinterpreted, either systematically or randomly. With DRE machines the VVPAT records may be smudged and unreadable (or lost entirely) due to a printer jam, or software errors may cause different information to be recorded in the electronic records from that printed on the VVPAT record.

4. *Counting errors:* Human auditors manually counting ballots or VVPATs may make occasional mistakes, which show up as discrepancies between the hand count and the machine count. However, safeguards, such as using three- and four-person counting teams and repeating the manual count if it disagrees with the machine count, eliminate almost all such errors. Discrepancies between the final manual count and the machine count determine whether to escalate the audit, even though the final manual count might not be perfect.

4.2 Handling of Discrepancies

The Working Group believes that there are three general responses that are appropriate when handling discrepancies:

1. Resolution and reporting of discrepancies;

2. Verification that all races have been called correctly, through escalated manual counts if necessary; and

3. Investigation and correction of systematic hardware, software, and procedural problems.

Below is a discussion of each in turn.

4.2.1 Resolving and Reporting Discrepancies

Elections officials should make every effort to determine the cause of each discrepancy that arises in the manual count, resolve each discrepancy and correct the official count if necessary. Generally, this is already done. While elections officials will report to the Secretary of State on discrepancies, the Working Group agrees elections officials should also make discrepancies and their resolution readily available to the public, even for local races not covered by the statute.

4.2.2 Verifying That Races Have Been Called Correctly; Escalating the Audit

Even when most or all discrepancies in the manual count of the initial random sample are satisfactorily resolved, there remains a fundamental question: *"Is the number of discrepancies found in the sample small enough that we can be very confident that the race has been called correctly, or is an escalation of the manual count warranted to increase that confidence?"*

In keeping with the principle that post-election auditing should control the risk of declaring the wrong winner, the Working Group believes that the process discussed in Section 3.2.3 is the best approach *in the long term*. Based on the number of precincts in the race, the number of precincts manually counted in the random sample, the number of ballots cast in each precinct and the margins in each precinct, elections officials would calculate the (smallest possible) probability that the observed discrepancies would be larger than observed if the outcome of the race were incorrect. If that probability is above some threshold, e.g. 99%, then one can be confident that the winner was correctly called. The audit could stop, and the winner could be certified.

However, if the probability is below the threshold, additional precincts for the manual count would be chosen at random to increase the sample size, after which a similar test would be performed. The process of escalating the sample size would continue until the confidence in the result exceeds the threshold or until all the precincts have been manually counted. This would generally result in larger manual counts for races with narrower margins of victory, and larger percentage manual counts for races involving a smaller number of precincts. It would also result in larger manual counts whenever, for any reason, the number of discrepancies in the initial sample is high – exactly as one would expect.

But as discussed in Section 3, the mathematics for a risk-based approach have not yet been completely worked out, let alone reduced to an easily understood and executed procedure, so the Working Group cannot recommend it *at this time*. Until the details have been filled in, there are a number of other escalation methods that the Working Group recommends. Although these options do not have as firm a statistical basis, they are still consistent with a risk-based approach.

Recommendations:

o   For very close races, candidates (or the proponents and opponents of propositions) should be permitted to choose one or two additional precincts for the manual count in addition to the 1% random sample. In those precincts, only the race in question would need to be counted, not the entire ballot. Registrars already have this option, but the Working Group

recommends that it be adopted uniformly since it increases the number of precincts manually counted for close races (consistent with the risk-based approach) and allows parties suspicious of the results in a particular precinct to satisfy their concerns. If discrepancies are found in the chosen precinct(s), more precincts should be counted manually. There should be some clear rule for deciding whether the discrepancies are small enough to stop the process and certify the election, or to continue to count more precincts.

o Assuming the discrepancies found in the manual count exist in similar numbers in all the precincts in a race and the total number would reduce the margin of victory in that race by one tenth, then additional precincts should be manually counted. Only the race in question needs to be hand counted in those precincts. This is a crude but easily implemented approximation to a risk-based escalation strategy. (As noted previously, if the initial sample size is too small, there can be a large chance that the sample will show no discrepancies at all even if there is enough error overall to alter the election outcome.)

o If discrepancies are found that are due to hardware, software, or ballot definition error, additional precincts should be manually counted. Depending on the nature of the problem and whether it can be fixed, up to a 100% manual count or a machine count may be warranted.

Options:

o If any voting system is found to be less accurate, reliable, or trustworthy than others, the Secretary of State could require a higher minimum manual count percentage across the board for that voting system until the problems are fixed or the system's trustworthiness is established.

o The minimum number of precincts to count manually in each jurisdiction might be increased from one to two (plus additional precincts for races not included in the 1% sample). This is consistent with the idea that races involving a small number of precincts should have a larger percentage of their precincts counted manually.

o In the case of optical scan systems, the Secretary might commission a study to determine the circumstances, *if any*, when a manual count might be supplemented by an additional machine count using a second optical scanner, either from the same vendor or another. Such a procedure may provide additional confidence in the election outcome at a lower expense than additional manual counting. (The Working Group does *not* suggest eliminating the manual count in favor of machine counts using a second machine.)

The Working Group is acutely aware of the time, space, and cost of manually counting more than the 1% minimum required by statute, particularly in large counties. The recommendations suggest that escalation triggered by discrepancies involve hand counting *only the races affected* by those discrepancies, not the entire ballot in the additional precincts. If the discrepancies are due to a voting system problem or official error that can be remedied, then the problem should be

21

fixed (e.g., ballots re-counted by a properly functioning scanner or tally machine), and the new machine count results should be compared against the results of the manual count.

4.2.3 Investigating and Correcting Systematic Hardware, Software, and Procedural Problems

If a manual count turns up errors that are the result of hardware, software, or systematic procedural problems, the Working Group believes additional special actions may be warranted.

Recommendation: The Secretary of State should be immediately notified of *uncorrectable* hardware or software problems or systematic procedural problems that are discovered during a manual count. The Secretary of State in turn should immediately notify elections officials in all jurisdictions that might be affected by the same problems so that they may address it.

In the case of DREs, the software is supposed to guarantee that the VVPAT records always match the electronic records exactly. Any discrepancy that is not due to an obviously mechanical cause (usually a paper jam, paper inserted incorrectly, or a faulty printer) must be due to a software problem in the DRE.

This is an extremely serious matter that may affect votes all over the country and could be evidence of fraud. Hence, any such discrepancy, no matter how small, and regardless of whether it can change the outcome of the election in question, should be thoroughly investigated by the Secretary of State, including a possible forensic investigation and code review.

Recommendation: If *any discrepancy at all* is found between manually counted VVPAT records and corresponding electronic vote counts that cannot be accounted for by some obvious mechanical problem, then all relevant VVPAT records, memory cards and devices, and DRE machines should be preserved; the Secretary of State should be immediately notified; and a thorough technical investigation should be conducted to determine the cause of the problem and assess its effects on the election.

5. Random Selection of Precincts

California law requires the selection of precincts for the post election manual count to be random and requires elections officials to use a "random number generator or other method specified in regulations" adopted by the Secretary of State to randomly choose the precincts. The requirement that a "random number generator" be used to choose the random sample of precincts for the manual count raises questions about the use of computers. Generally, so-called "random number generators" in computers do not generate random numbers, but rather, pseudo-random numbers. True random number generators rely on unpredictable physical phenomena, such as radioactive decay. In contrast, pseudo-random numbers are predictable.

Any use of computers – even computers with true "hardware" random number generators – reduces the transparency of selecting the sample of precincts because software cannot be

observed directly, as Cordero, Wagner, and Dill (2006) argue. Cordero *et al.* also examine several physical methods of drawing random samples and their weaknesses. For example, choosing precinct numbers out of a hat is not transparent because observers cannot verify that every precinct number is in the hat exactly once. Finally, the Cordero paper describes a simple, transparent way to make the random selection, namely, to use 10-sided dice of different colors, with each color representing the ones, tens, or hundreds places, and a printed look-up table of precincts numbered from 0 to N-1, with special provisions for situations in which there are more than 1,000 precincts. Other selection methods may work as long as: 1) the method produces true "simple random samples," i.e., every subset of precincts is equally likely to be chosen, 2) there is no possibility of human interference, and 3) observers can see every detail of the process.

Recommendations:

- o **Develop regulations.** The Secretary of State should develop regulations for randomly choosing precincts. The regulations should only allow the use of 10-sided dice and other physical methods for generating random samples. The Secretary should consider amending Elec. Code § 15360(c) to remove the requirement to use a "random number generator" and to specifically prohibit computer-generated "pseudo-random" number generators.

- o **Provide a legal opinion.** The Secretary of State should provide legal clarification on whether the statutory meaning of "random number generator" includes 10-sided dice and other manual methods of random selection and whether computer "pseudo-random number generators" satisfy the statutory requirement.

- o **Establish standards.** The Secretary of State should establish standards permitting one random sample of precincts to be drawn to hand count election day results and an independent random sample to be drawn to hand count Vote by Mail ballots. By drawing two independent random samples – one of the precincts as soon as they are counted and the other for Vote by Mail ballots, which take longer to process – the manual count could begin as soon as the ballots from all precincts have been counted. This would make it much easier for jurisdictions to meet the 28-day limit on certifying the canvass and still ensure that all ballots have the same chance of being audited.

6. Additional Options to Improve the Manual Count

The following options would improve the manual count process in California:

Chain of custody. The Secretary of State should evaluate and standardize the procedures used to secure the physical custody of the ballots and VVPATs. The Working Group recommends including as the use of a two-person rule, cameras, locks, and chain of custody logs.

Make preliminary counts available to the public immediately. Preliminary electronic counts in *all* precincts (or other audit units) should be published as soon as they are available. The

sooner they are published (e.g., election night), the smaller the opportunity is for preliminary electronic counts to be altered.

Start the manual count as soon as possible after the random drawing of precincts. As soon as possible after the random sample is drawn, all of the ballots and VVPATs for the manual count should be brought into public view and the counting should begin. This minimizes the opportunity, both real and perceived, to manipulate the paper records between the time the sample is chosen and the time it is counted.

Ballot tracking. Require tracking and reporting of undervotes, overvotes, and cancelled DRE ballots as part of the manual count process.

Unreadable VVPATs. Current law requires if there is a discrepancy between the electronic record and the VVPAT, the VVPAT is to be used during the manual count. (Elec. Code § 19253(b)(2)) However, if the discrepancy is the result of an unreadable VVPAT, the law has the unintended consequence of disenfranchising voters. The law should be amended to allow elections officials to print copies of electronic ballots, but only for unreadable or spoiled VVPATs, and to use those copies in the manual count. The electronic record is not voter-verified, but this is the only way to include those voters' ballots.

Independence of the manual counters. The following options can make manual counts more independent: 1) four-person counting teams, which can help catch misreadings of ballots; 2) hiring temporary help from the local community to count ballots, instead of using permanent employees; and 3) appointing community members rather than regular city or county employees to ballot inspection boards and manual count boards.

Blind counting. People performing the manual count must not know the preliminary electronic count, nor should they be told whether or not their result matched the electronic count (in case the manual count must be repeated). Observers, who perhaps know the preliminary count because it has been published, should be reminded not to tell the counters whether the manual count matches the electronic count.

Establish uniformity in public reporting. As noted above, beginning January 1, 2008, elections officials must report manual count results to the Secretary of State, including discrepancies found and measures taken to resolve discrepancies.

7. Increasing the Transparency of the Manual Count and the Post-Election Process

The post-election audit is *transparent* if the public can observe all parts of the process, see that the audit was properly conducted, and see that the outcomes of the races are correct. Transparency in the auditing process is crucial for public confidence. Jurisdictions can take the following measures to make California's post-election audit process more transparent:

   o   Make preliminary counts immediately available to the public.

- Draw precincts for the manual count in a random, verifiable, publicly observable manner.

- Begin the manual count as soon as possible after precinct selection. (See Sections 5 and 6.)

- Conduct the manual count in public view. Observers must not be permitted to touch ballots or interfere in any way with the counting, but also should not be kept at a distance that keeps them from verifying the count.

- Log all unusual events and decisions made during the manual count process and make that log public. A public log of findings and departures from written protocol reduces opportunities to manipulate the outcome and provides a permanent record that can be used to improve the audit process itself.

- Transfer ballot data from the DREs to storage media and secure the data electronically for at least 22 months. Security measures may include redundant copies, chain of custody logs, a two-person rule for accessing the data, digital signatures applied to the data, and the use of hash codes.

- Make pre- and post-election Logic and Accuracy test results available for public viewing immediately after they are conducted.

- Set up election advisory committees with representatives of various constituencies, similar to those established in Alameda, Humboldt, and Marin counties.

The Secretary of State can do the following to increase the transparency of the manual count:

- **Post key manual count dates.** The Secretary of State could ask local elections officials to report the dates, times, and locations of random selection of precincts and manual counts and post that information on the Secretary of State Web site, providing the public convenient access to this time-sensitive information.

- **Post manual count discrepancy reports.** The Secretary of State should publish manual count discrepancy reports on the agency's Web site to give the public convenient access to statewide manual count results.
  (For an example, see the Minnesota Secretary of State's 2006 Post Election Voting System Review, http://www.sos.state.mn.us/home/index.asp?page=544.)

- **Consider increasing the use of parallel monitoring.**

8. Public and Poll Worker Education About the Voter Verified Paper Audit Trail

Post-election audits of DREs assume that the VVPAT shows voters' true intent for each vote cast. To ensure the assumption is correct, voters must check the VVPAT before casting their ballot. Informal studies suggest that many voters do not do review the VVPAT. Many poll workers are also unaware of the significance of the VVPAT.

Voters and poll workers could be better educated about the importance of checking VVPATs. Poll worker training could address the importance of VVPATs. Signs inside DRE voting stations could remind voters to check VVPATs, as could news advisories, public service announcements, and official state and local voter guides.

Voter education can also help alleviate problems that may result from VVPATs, such as printer jams, paper outages, or illegible print. If voters understand the importance of VVPATs, they will more likely bring such problems to poll workers' attention.

The most serious election day problem that could occur with a VVPAT is a discrepancy between the printed record of the ballot and the summary screen or VVPAT record on the DRE screen. If this kind of problem arises, voters should be encouraged to notify poll workers so that the DRE can be examined before further use.

The Secretary of State should work with local elections officials in counties using DRE systems to educate poll workers and voters about the importance of the VVPAT and urge voters to check the VVPAT before casting their ballots.

9. Recommendations for Further Study

As mentioned earlier, the Working Group recommends that the Secretary of State establish a long-term working group that includes a number of state, county, and city elections officials to study and develop statewide standards to improve the canvass period as a whole.

The following areas should be included in the scope of further study:

Confidence levels. When a practical algorithm is available, elections officials should calculate and publish the statistical confidence that each race was called correctly, given the number of precincts audited, the number and sizes of precincts in the race, the margin of victory in each precinct, the discrepancies found during the manual count, and other key factors.

Chain of custody. Evaluate and standardize procedures used to secure the physical custody of the ballots and VVPATs. The Working Group recommends including standards such as the two-person rule, cameras, locks, and chain of custody logs.

Reconciliation. Create uniform reconciliation methods, including procedures for:
o Reviewing tapes from individual DREs and optical scanners and comparing those totals both to the overall vote totals and to the manually counted totals.
o Comparing vote totals from the election night results tapes from each precinct to the precinct totals on the statement of vote to make sure they match.
o Comparing the total number of ballots issued to the precinct to the number of ballots voted, not voted, and spoiled and provisional ballots to ensure completeness of the count.
o Comparing the number of voted ballots to the number of signatures on the precinct roster.

Tracking, reporting, and auditing for spoiled ballots and provisional ballots. Because provisional ballots are not part of the manual count process, methods should be developed to track, report, and audit the machine count of provisional ballots.

Manual counting methods. Evaluate and standardize procedures used in the manual count. For example, some counties verify the manual count by repeating the manual count twice and stopping if the two manual count results match. Other counties have two teams count the same ballots manually at the same time.

Counting the VVPATs. Determine the best way to count VVPATs. Some counties cut VVPAT rolls before counting individual ballots, while others leave the rolls intact. Explore whether it makes sense to require VVPATs to be machine-readable in the future, so that a machine count could be conducted as part of an audit.

Auditing of duplicated ballots. In order to prevent fraud or mistake in the duplication of ballots that are not machine readable, duplicated ballots should be audited. One method is to check duplicated ballots that come up in the manual count against the originals of those ballots.

Auditing of each voting system component. Evaluate and standardize procedures used to verify the accuracy of all technology involved in the election, including optical scanners used to read paper ballots, DREs used to record electronic ballots, and vote-counting software used to count ballots and report results.

Transparency. Develop procedures for public reporting of the entire process and results of the manual count and reconciliation. Current law only requires public reporting of the manual count results.

Secretary of State oversight. Explore whether the Secretary of State should increase oversight of the post-election process. Oversight measures could include establishing a state canvassing board, monitoring the manual count, or sending additional observers to monitor counties that experienced problems in prior elections.

Training and peer review. Develop training and peer-review programs for elections officials to ensure uniform implementation of audit procedures. This could include establishing mechanisms to gather feedback from elections officials following each election to continually refine procedures. "Round robin" audits could be used as a peer review method and would involve one jurisdiction conducting the manual count audit for another jurisdiction instead of a jurisdiction auditing its own election results.

Appendix I: Definitions

DRE:
Direct Recording Electronic voting machine.

Escalation:
Escalation as referred to in this report means expanding the size of the manual count. Generally speaking, however, escalation of an audit may also include other forms of investigation into a problem found and the reporting of the problem to the Secretary of State.

Logic and accuracy testing:
Basic testing that voting equipment functions properly before and after an election.

Post-election audit:
Post-election manual process of manually counting ballots or VVPATs, comparing the manual counts to machine count totals, and resolving discrepancies. The 1% manual count called for in current California law is a post-election audit. The terms "manual count" and "post-election audit" are used interchangeably in this report.

Reconciliation:
Verification at every stage in the canvass, from the precinct closing to the final count, that the number of ballots counted is the same as the number of ballots cast, including all categories of ballots, for example Vote by Mail and provisional ballots, blank ballots, spoiled ballots, etc. Many steps in the audit process presume that reconciliation has already been done, so that there should be few discrepancies in the *numbers* of ballots of each kind, even if there are discrepancies in the votes counted during the audit.

Recount:
Non-routine post-election recounting process to resolve a close race or to resolve an election challenge. It differs from an audit in that the candidates are directly involved and the proceedings may be adversarial.

Software Independence:
Informally, "software independence" refers to "any approach where the integrity of the election outcome is not dependent on trusting the correctness of complex software." (Rivest & Wack, 2006)

VVPAT:
Voter Verified Paper Audit Trail.

Appendix II: Common Causes of Discrepancies

Causes of discrepancies between the manual and machine count include:

- o Jammed VVPATs

- o Multiple scans of the same optical ballots or ballots that don't get counted by scanner

- o Counting errors by staff doing the manual count

- o Poll workers mixing counted and uncounted ballots together after the polls close

- o Contests omitted from ballot (usually discovered before the election)

- o Difference in results between what was reported from the machine at the polls and what was uploaded into the server on election night

- o Missing memory cards or ballots

- o Mismatches between ballot order of candidates/races and order in the software used to count results

- o Omitted precincts or machines

- o Software errors in DREs and scanners

Specific examples of errors that have caused discrepancies in past elections are:

Napa County, March 2004 Statewide Primary Election. In Napa County, officials discovered that their Sequoia optical scanners had not been properly calibrated to detect the various types of ink with which ballots were marked. Reference: "Lost E-Votes Could Flip Napa Race," by Kim Zetter, *Wired News*, March 12, 2004.

San Diego County, March 2004 Statewide Primary Election. In San Diego County, Diebold's software awarded nearly 3,000 votes cast on Vote by Mail ballots for Democratic Presidential candidate John Kerry to Dick Gephardt, who by that time had dropped out of the race. Reference: "Some votes miscounted in primary, officials say," by Luis Monteagudo Jr. and Helen Gao, *San Diego Union Tribune*, April 4, 2004.

Alameda County, October 2003 Statewide Recall Election. Alameda County discovered that Diebold's optical scan software attributed 9,000 votes to a Socialist Recall candidate that should have gone to Cruz Bustamante. Reference: "Diebold reports multiple problems: Registrar wants reason for e-voting," by Ian Hoffman, *Tri-Valley Herald*, April 13, 2004.

San Bernardino, November 2001 Local Elections. San Bernardino County's elections department failed to test its software prior to tabulating the election results. While the employee charged with this duty claimed the software had been tested, the county did not confirm this claim prior to utilizing the software to tabulate the ballots. After ballots had been counted and a slate of winners announced, the elections department discovered the software malfunctioned. The program was fixed and the paper ballots were recounted. As a result, the elections department had to notify 13 candidates who had previously been told they won that they in fact lost. Reference: "Election Foul-Up Leads to New Slate of Winners," by Buck Wargo, *Los Angeles Times*, November 14, 2001.

Appendix III: Bibliography

ALEXANDER, K., Public verification of software vote counts and California's manual count law. California Voter Foundation, February 2006.
http://www.calvoter.org/issues/votingtech/manualcount.html

BERTELSEN, J., 1% Manual Tally Observer Report, Congressional District 11, January 10, 2007. http://www.countedascast.com/docs/CD11_Manual_Tally_Report_Jan01.pdf

BICKEL, N., COULTER, J., COULTER, T., and CHAPMAN, B. for the Alameda County Council of the League of Women Voters, Counting Our Votes on Paper and Electronically: The League of Women Voters Observes the Election Process in Alameda County in the November 2006 General Election, June 26, 2007. http://lwvbae.org/ACC%20-%20Vote%20Count%June%202007.pdf

CORDERO, A., WAGNER, D., and DILL, D., The role of dice in election audits — extended abstract, June 2006. http://www.cs.berkeley.edu/~daw/papers/dice-wote06.pdf

DOPP, K., and STENGER, F., The Election Integrity Audit, September 2006.
http://electionarchive.org/ucvAnalysis/US/paper-audits/ElectionIntegrityAudit.pdf

DOPP, K., Tiered Election Audits: Based on Margins between Candidates, January 23, 2007.
ElectionArchive.org. http://electionarchive.org/ucvAnalysis/US/paper-audits/FourTierAudit/TieredElectionAudits.pdf

ELECTIONLINE.ORG, Case Study: Auditing the Vote, March 2007.
http://www.electionline.org/Portals/1/Publications/EB17.pdf

FELTEN, E.W., Machine-assisted election auditing. Center for Information Technology Policy and Department of Computer Science, Princeton University; Woodrow Wilson School of Public and International Affairs, Princeton University.

GABRIEL, M., Election Monitoring: Observing and Reconciling Election Counts, California, September 29, 2006. http://www.countedascast.com/docs/ElectionMonitoringV1.0.pdf

GOGGIN, S.N., An examination of the auditability of voter verified paper audit trail (VVPAT) ballots. Department of Psychology, Rice University, Houston, Texas. This paper will appear in the Electronic Voting Technology 2007 Conference Proceedings and will be made publicly available on August 6, 2007.

HALL, J.L., A quick primer on the mathematics of post-election audit confidence. School of Information, University of California Berkeley, March 20, 2007.
HALVORSON, M., and WOLFF, L., Report and Analysis of the 2006 Post-Election Audit of Minnesota's Voting Systems. Citizens for Election Integrity Minnesota, April 4, 2007.
http://ceimn.org/files/CEIMNAuditReport.pdf

LOS ANGELES COUNTY, Summary Reports; Tally Sheets Report; Roster Reconciliation Operation Manual; Touchscreen, Snap, and Absentee Tally Reports; Manual Count Cost Summary, November 2006.

MARIN COUNTY, Marin County Registrar of Voters, Procedures for One Percent Manual Tally. Marin County, California.
http://www.co.marin.ca.us/depts/RV/Main/election_info/Procedures/One_Percent_Manual_Tally.pdf

MCCARTHY, J., STANISLEVIC, H., LINDEMAN, M., ASH, A., ADDONA, V., and BATCHER, M., Percentage-Based Versus SAFE Vote Tabulation Auditing: A Graphic Comparison. Verified Voting Foundation.
http://www.verifiedvotingfoundation.org/auditcomparison

MCNERNEY ELECTION PROTECTION TASK FORCE, Final Report on Findings and Recommendations, November 6, 2006, California General Election, March 2007.
http://www.jerrymcnerney.org/pdfs/EPTF-FINAL-Mar28-07.pdf

NORDEN, L., BURSTEIN, A., HALL, J., and CHEN, M., Post-Election Audits: Restoring Trust in Elections, July 2007. http://brennancenter.org/subpage.asp?key=38&proj_key=76

RANDOM.ORG, Introduction to Randomness and Random Numbers
http://www.random.org/randomness/

RIVEST, R. L., On Estimating the Size of a Statistical Audit, November 14, 2006.
http://people.csail.mit.edu/rivest/Rivest-OnEstimatingTheSizeOfAStatisticalAudit.pdf

RIVEST, R., and WACK, J., On the notion of 'software independence' in voting systems, July 2006. http://vote.nist.gov/SI-in-voting.pdf

SALTMAN, R. G., Effective Use of Computing Technology In Vote-Tallying, Tech. Rep. NBSIR 75–687, National Bureau of Standards (Information Technology Division), March 1975.
http://csrc.nist.gov/publications/nistpubs/NBS_SP_500-30.pdf

VERIFIED VOTING FOUNDATION, Manual Audit Requirements, August 18, 2005.
http://www.verifiedvoting.org/downloads/Manual_Audit_Provisions.pdf

Appendix IV: Testimony and Oral Consultation

Jerry Berkman, Alameda County, California (July 2, 2007)

Judy Bertelsen, Alameda County, California (July 2, 2007)

Nancy Bickel, Berkeley, California (July 2, 2007)

Doug Chapin, Director, electionline.org   (June 27, 2007)

Carolyn Crnich, County Clerk-Recorder-Registrar of Voters, Humboldt County (July 10, 2007)

Cathy Darling, County Clerk-Registrar of Voters, Shasta County (July 2, 2007)

David L. Dill, Professor, Department of Computer Science, Stanford University (July 2, 2007)

Michelle Gabrielle, Alameda County, California (July 2, 207)

Joseph Lorenzo Hall, School of Information, University of California, Berkeley (July 2, 2007)

Elaine Larson, Assistant Registrar of Voters, Santa Clara County (July 2, 2007)

Conny McCormack, Registrar-Recorder-County Clerk, Los Angeles County (July 9, 2007)

Kenneth Mostern, TrueBallot, Inc. (July 2, 2007)

Gary Poser, Director of Elections, Office of the Secretary of State, Minnesota (June 29, 2007)

Ronald L. Rivest, Professor, Department of Electrical Engineering and Computer Science, MIT; member, Technical Guidelines Development Committee (TGDC) to the U.S. Elections Assistance Commission (EAC) (July 9, 2007)

Roy G. Saltman, author and independent consultant, election technology and voting integrity (July 5, 2007)

Deborah Seiler, Registrar of Voters, San Diego County (July 2, 2007)

Jim Soper, Alameda, California (July 2, 2007)

Richard Tamm, Berkeley, California (July 2, 2007)

John Tuteur, Assessor-Recorder-County Clerk, Napa County (July 2, 2007)

David Wagner, Associate Professor, Computer Science, University of California at Berkeley; member, Technical Guidelines Development Committee (TGDC) to the U.S. Elections Assistance Commission (EAC); Principle Investigator, CA Secretary of State's Top to Bottom Review (June 2007)

Stephen L. Weir, County Clerk-Recorder-Registrar of Voters, Contra Costa County; President, California Association of Clerks and Election Officials (July 2, 2007)

Appendix V: Post-Election Audit Standards Working Group Charter

MISSION

The mission of the Post-Election Audit Standards Working Group was to evaluate various ways to conduct post-election audits; provide options to the Secretary of State for use in the planning, development and implementation of manual count audit standards that will enhance the security, accuracy and reliability of election results; and develop best practice guidelines for reconciling manual audit and machine count results.

BACKGROUND

While many states have no requirement to manually audit election results, California Elections Code § 15360 requires counties to conduct a manual tally of 1% of precincts following each election. Congress is considering legislation to set a national standard that could require higher percentages of election results to be audited in a post-election manual count. Because post-election manual auditing can help ensure the security, accuracy and reliability of voting systems, it makes sense to take a comprehensive look at auditing standards, models and other processes that may be appropriate to implement in California.

OBJECTIVES

The Working Group was asked to provide perspectives and an evaluation in a report to be submitted to the Secretary of State by July 27, 2007. The Working Group was not directed to come to a consensus nor develop formal recommendations. Rather, it was asked to weigh the advantages and disadvantages of various audit methods and evaluate the reliability and accuracy of each method. All comments are welcome and encouraged. The Secretary of State will consider all comments and recommendations from the Working Group.

ROLES AND RESPONSIBILITIES

All Working Group members were expected to:

o Demonstrate a commitment to open discussion and the sharing of ideas, as well as the desire to work in partnership with other members to identify issues and possible solutions.

o Review materials distributed prior to the meetings and attend meetings prepared to engage and provide input, and make recommendations on the issues being discussed.

o Raise issues and deliberate with other members.

o Provide feedback and recommendations regarding project issues as requested.

o Help develop the Working Group's report to the Secretary of State.

MEMBERSHIP

The Working Group was comprised of representatives of various interests as specified below. Secretary of State staff facilitated meetings of the Working Group.

- o David Jefferson (Chairperson), Center for Applied Scientific Computing, Lawrence Livermore National Laboratory

- o Elaine Ginnold, Registrar of Voters, Marin County

- o Kathleen Midstokke, City Clerk, City of Downey

- o Kim Alexander, President, California Voter Foundation

- o Philip Stark, Professor of Statistics, University of California, Berkeley

- o Amy Lehmkuhl, Certified Public Accountant, Ueltzen & Company, LLP

MEETINGS

The Working Group held one public forum to hear presentations from experts and comments from members of the public.

COMMUNICATION

- o Communication was handled via scheduled meetings and email.

- o Drafts of the Working Group's report to be submitted to the Secretary of State were kept confidential.