A preliminary version of this paper will appear in *Proc. of the 2007 USENIX/ACCURATE Electronic Voting Technology Workshop (EVT'07)*, August 2007

Machine-Assisted Election Auditing

Joseph A. Calandrino[*], J. Alex Halderman[*], and Edward W. Felten[*,†]

[*]Center for Information Technology Policy and Dept. of Computer Science, Princeton University
[†]Woodrow Wilson School of Public and International Affairs, Princeton University
`{jcalandr,jhalderm,felten}@cs.princeton.edu`

Abstract

Election audit procedures usually rely on precinct-based audits, in which workers manually review all paper ballots from selected polling places, but these audits can be expensive due to the labor required. This paper proposes an alternative audit strategy that allows machines to perform most of the work. Precincts are audited using auditing machines, and their output is manually audited using efficient ballot sampling techniques. This strategy can achieve equal or greater confidence than precinct-based auditing at a significantly lower cost while protecting voter privacy better than previous ballot-based auditing methods. We show how to determine which ballots to audit against the auditing machines' records and compare this new approach to precinct-based audits in the context of Virginia's November 2006 election. Far fewer ballots need to be audited by hand using our approach. We also explore extensions to these techniques, such as varying individual ballots' audit probabilities based on the votes they contain, that promise further efficiency gains.

1 Introduction

Security analyses of computerized voting systems, including DREs and optical scan machines, have exposed numerous vulnerabilities that could compromise the integrity of elections performed using these devices (see [8, 6] and references therein). One proposed defense against such attacks is to produce voter-verified paper records and audit them to ensure that they support the totals claimed by the machines. This defense is possible when using standard optical scan machines or DREs with receipt printers.

The most common auditing method is the precinct-based audit [3, 10, 11, 13, 14], in which workers count all paper ballots from selected precincts and compare the results to the reported precinct tallies.[1] Unfortunately, performing precinct-based audits can require considerable time, labor, and expense. These costs are multiplied by the complexity of the ballots in many elections, which may include dozens of contests. In a trial recount of a DRE paper trail performed in Cobb County, Georgia, workers took an average of 5 minutes per ballot to audit 976 votes at a total cost of nearly $3,000 [5]. Unless efficiency can be improved, performing a similar recount of 3% of precincts in New Jersey could cost more than $200,000. Slow, expensive manual audits limit the level of confidence that can be achieved within a fixed election budget, and they may delay the detection of errors until well after election results have been announced and losing candidates have conceded.

In this paper we propose an alternative audit strategy that substantially reduces these costs by using specialized machines to automate most of the work of auditing paper ballots followed by a manual audit of the machine results. The problem with machines, of course, is that the ones used for the audit are not necessarily more trustworthy than the ones used in the initial count. They may be useful for catching inadvertent errors (especially if they use a different technology and independently developed software), but a determined attacker could still target both sets of machines. What we desire is software independence—an assurance that any tampering with the machines will not cause undetected changes to the election outcome [12]. To achieve this, we pair audit machines with efficient statistical auditing techniques that allow humans to confirm that the election outcome is correct.

Precinct-based audits require officials to select precincts at random and perform full manual audits in those precincts. A manual precinct tally that differs

---

[1] H.R. 811, now under consideration in Congress, would mandate a complete manual audit of 3%, 5%, or 10% of precincts, depending on the margin of victory. However, the bill permits alternative audit methods so long as they provide an equivalent level of confidence. [1]

from the electronic tally may raise suspicion. Statistical "ballot-based" audits are an alternative to manually auditing every ballot from selected precincts. Workers instead sample from all the paper ballots in all precincts and use the sample to assess the accuracy of the original count. Depending on the specific implementation, officials may either compare selected paper ballots directly to corresponding electronic ballots, or they may compare the sample tally to the total electronic tally. We focus on the former type of ballot-based auditing in this paper.

Ballot-based audits tend to be more efficient than traditional precinct-based audits [9], since fewer ballots need to be audited to achieve the same level of confidence in the result. For example, in a state-wide race in New Jersey, fewer than one ballot per precinct (4,599 ballots total) would need to be sampled to achieve 99% confidence that the outcome had not been shifted by more than 0.2%. By contrast, over 150,000 ballots (6.9% of precincts) would need to be audited using standard precinct-based audits (e.g., [14]) to achieve the same confidence.

Neff [9] and Johnson [7] were among the first to propose combining ballot-based audit techniques with electronic voting. Neff assumes that the voting machines link each paper ballot to its electronic counterpart using, for example, a unique identifier printed on the paper ballot and stored with the electronic ballot. When voting is complete, each precinct commits to its set of electronic ballots, then demonstrates that the paper ballots in a given random sample match the corresponding electronic ballots.

The primary weakness of this method is that it establishes the link between electronic and paper ballots at the time that votes are cast. This raises problematic voter privacy issues. For example, if the ballots are linked using sequentially increasing serial numbers, observers could correlate votes with the order in which they were cast, which can reveal the identity of voters. Opaque, random-looking cryptographic identifiers printed on ballots might protect privacy, but they could also potentially provide covert channels for leaking voter identities. Even if used securely, they might aid malicious parties who seek to intimidate voters by undermining their confidence in the secrecy of the ballot. Our audit strategy postpones linking paper and electronic records until the audit phase, which allows it to achieve equivalent confidence without jeopardizing privacy or resorting to cryptography.

Johnson alternatively proposes delaying both vote tallying and serial number printing until after all ballots are submitted, allowing voting machines to be simple, memory-less ballot printers [7]. Voters submit their ballots, which, once polls close, are randomized and scanned/tallied. The tallying machine is therefore able to print serial numbers while scanning without privacy risk. Unlike Johnson, we assume that the voting machines maintain an electronic tally, which helps deter traditional attacks against paper-based voting, such as ballot-box stuffing, and, as we will show, provides opportunities for improving the efficiency of the audit.

Our main contributions are:

- We propose a novel audit approach wherein ballots are audited using auditing machines, and the machines' output is manually audited by humans using ballot-based auditing techniques. (Sections 2 and 3)
- We evaluate the efficiency of our method using data from Virginia's November 2006 elections, and we find that it enjoys significant gains compared to the traditional precinct-based approach. (Section 4)
- We suggest several extensions to address practical considerations and to further improve efficiency, including means of using knowledge of ballot contents to reduce the sample size. (Section 5)

2 Machine-Assisted Auditing

We propose replacing manual precinct-based audits with machine-assisted audits. Poll workers, rather than auditing ballots manually, feed them through a specialized audit machine that functions like a combined optical scanner and printer. As it scans the contents of each ballot, it prints a unique serial number that is stored along with the ballot contents. At the end of the scanning process, the machine outputs a list of votes on each ballot together with the ballot's serial number. If the audit tallies differ from the initially reported electronic count, discrepancies clearly exist and a wider investigation should be conducted.[2] If both tallies match, the workers perform a secondary audit to check the accuracy of the machine's audit. They first quickly flip through the pile of numbered ballots to ensure that the serial numbers increase sequentially from one to the reported ballot total without repeats.[3] They then take a random sample of the electronic ballot records, retrieve the corresponding paper ballots, and verify that they match.

Since the ballots are serialized and fed out of the machine in order, retrieving a particular ballot for verification requires very little effort. The most significant labor

---

[2] Depending on circumstances, an appropriate response might be to inspect the corresponding machines, other machines of the same model, other ballots in that precinct, etc. A comprehensive set of precise responses to various circumstances is beyond the scope of this paper, but we note the importance of establishing procedures to avoid partisan bickering if discrepancies arise.

[3] This check helps protect against collusion between voting and audit machines, as described shortly. A more efficient means of ensuring that the number of paper and electronic ballots match may exist.

required may be to check for repeats, which given sequential ordering, is a rapid single-pass process.

In practice, separate devices may be used to perform the printing and scanning functions of our proposed audit machine. When voting is complete, a printer device could place serial numbers on the ballots, and then a separate scanner could read the numbers along with the votes. In precincts utilizing optical scan machines, properly designed machines could perform both the initial count and the audit: this option decreases costs but reduces redundancy. If the same machine performs counts, audits, and printing, officials must have some means of mechanically disabling the printer while polls are open, such as removal of the printer head. Printers also must be physically unable to alter the record of the vote on the ballot. They could be designed so that they cannot reach outside of a predefined empty margin on ballots, or they could utilize a kind of ink that would be immediately apparent when ballots were inspected.

Correctness Our algorithm is given a set of electronic ballots and a set of paper ballots. The goal of the algorithm is to ensure that if the electronic tally would declare a different winner than a full manual recount of the paper ballots would declare, then with high probability (i.e., a given confidence level) the audit will find a discrepancy. In this section we sketch a proof that this property holds.

Given the electronic tally, we can find a number B such that the difference between the sets of electronic and paper ballot (assuming that the electronic ballots support the initial tally) must be at least B ballots[4]. For example, in a two-candidate race where the candidates' electronic tallies differ by x votes, $B = \lceil \frac{x}{2} \rceil$.

Assuming the machine audit gives the same vote total as the original electronic tally (otherwise a discrepancy is reported), if the differences between the paper and electronic tallies would require modification of at least B ballots, then the paper and machine audit ballot sets will also differ by at least B ballots.

Now consider the i'th precinct. Let there be n_i electronic ballots in this precinct, and let b_i be the number of ballots that differ between the paper and electronic ballot sets in this precinct. If the number of paper ballots is at most n_i (otherwise a discrepancy will be reported), it follows that there are at least b_i electronic ballots that have no matching paper ballot.

Summing over all precincts, we find that as long as, in every precinct, the number of paper ballots is at most the number of electronic ballots, overall there will be at least B electronic ballots that have no corresponding paper ballot. It follows that our sampling algorithms, which are designed to find at least one bad (electronic) ballot with high confidence whenever there are at least B such bad ballots, are sufficient to detect a discrepancy with high probability if one exists.

Privacy Our technique avoids many of the privacy issues inherent in some earlier ballot-based audit methods that involve placing identifiers on ballots during the voting process. In our technique, the ballots do not receive serial numbers until the audit phase, so they are likely to become at least partially reordered before being numbered. Well-designed ballot boxes and cut-and-drop paper trail systems assure that the papers are somewhat shuffled as they are inserted. Since voters widely trust these methods to frustrate correlation with voter check-in times, this provides significant practical privacy benefits. Should alternative ballot shuffling methods offer greater protection, officials may substitute such methods without modifying the audit process. In any case, the audit machine has no more information about the order of votes than would workers performing a manual audit.

Another benefit of this technique is that a voting machine need only maintain tallies rather than electronic copies of individual ballots. These tallies must include the total number of ballots submitted and the total number of votes for each option. Thus, voting machine designers do not need to worry about properly shuffling electronic ballots to protect voter privacy or about maintaining storage for those ballots. However, if the same machines perform counts and audits, they must have some means of using extra memory during the audit for storing the ballot scan results.

3 What to Audit

Due to the popularity of plurality voting systems in the U.S. we exclusively consider those systems, though machine-assisted audits may be useful in many other voting systems. With plurality voting, voters may choose a number of candidates equal to the number of seats available.[5] If k seats are available, voters may select up to k candidates, and candidates receiving the top k vote totals are the victors. This definition is an extension of the familiar single-seat contest.

An audit process need only sample enough ballots to confidently detect the minimum amount of fraud that would have affected the election's outcome.[6] To modify the fewest ballots while changing the outcome, an

---

[4] A difference of one ballot means one ballot added, removed, or modified.

[5] This is a mild misuse of the term plurality system: other forms of plurality voting for multiple candidates exist [2].

[6] As a quality-control measure, officials might want to examine more ballots than are strictly necessary following a landslide victory. Such quality-control audits might help to detect any machine malfunctions before they had a chance to affect a future (close) election.

adversary would swap the positions of the losing candidate with the most votes and the victor with fewest votes. Switching votes directly between these candidates requires the fewest ballot changes, as each switch alters the relative difference by two. To do so, the adversary would take ballots with votes for candidate A but not B and change them to contain votes for B but not A. Therefore, we need only audit enough ballots to discover fraud that alters a number of ballots equal to half the difference (rounded up) in vote totals between the "just losing" and "just winning" candidates.

We describe two techniques for selecting which ballots and precincts need to be audited. The first technique has the benefit of a constant sample size given the number of ballots, the margin of victory, and the desired level of confidence. Sample size may vary with the second approach, but that approach is more amenable to extensions that we propose later.

3.1 Constant Sample Size Method

The hypergeometric distribution describes the number of bad ballots an auditor can expect to find when sampling without replacement. Assume that auditors desire a confidence level c that no fraud significant enough to change the election's outcome occurred. By [10], given N total ballots and a minimum of B incorrect ballots, the probability mass function of the hypergeometric distribution dictates a minimum sample size, n, of:

$$n = \min \left\{ u \mid 1 - \prod_{k=0}^{u-1} \frac{N-B-k}{N-k} \geq c \right\} \qquad (1)$$

A simple computer program can rapidly, verifiably calculate n for any practical value of N.

After all precincts report their audit results and scanned ballots, state officials randomly select n ballots to check. To do so, officials assign each ballot an equivalent portion of the output of a pseudorandom generator.[7] Representatives for all candidates or issues in a race may assist in randomly generating a seed for the function (for example, consider [4]).

Because officials select ballots at random with respect to any given race, officials may use the same ballot from auditing one race in auditing any other race appearing on that ballot, provided that all voters eligible to vote in the latter race are also eligible to vote in the former. This reduces the number of ballots to retrieve. Note that the correlation between votes on a given ballot prevent us from gaining additional assurance from using the same ballots for multiple races, but officials still gain confidence c in the results of each race.

A machine audit of a precinct is only necessary if a ballot will be selected for manual verification in that precinct. Thus, given accurate precinct ballot counts, auditors could use the initial electronic tallies to perform a mock ballot selection before the machine audit. Any precinct which would have contained a chosen ballot given the mock selection will undergo a machine audit. Following the machine audit, representatives must generate a new seed for the pseudorandom generator.[8] Officials may then randomly select a single ballot from each audited precinct and randomly draw the remaining required ballots from the full pool in all audited precincts.

3.2 Varying Sample Size Method

Rivest [11] proposes an efficient precinct-based auditing technique in which, rather than drawing a given-size sample from the population of precincts, auditors instead randomly select each precinct with a given probability. The same idea is also useful in the context of ballot-based auditing. Assume that, to change the results of an election, the set of ballots must contain a minimum of B bad ballots. To achieve a confidence level of c that at least one bad ballot will be sampled, auditors may select each ballot with probability p chosen such that $(1-p)^B \leq 1-c$, or $p \geq 1-(1-c)^{1/B}$.

Officials may follow the same process as before for generating a seed and may apply a pseudorandom generator to a unique identifier for each ballot (for example, 1 to N, where N is the total number of ballots in all precincts voting on the given issue), mapping the result back to $[0,1]$ to determine whether to check the ballot.

To determine which precincts need to be audited, we may calculate the probability that one or more of the v_i ballots in precinct i will be sampled as $1-(1-p)^{v_i}$. Auditors may select each precinct based on the probability that it contains a sampled ballot. If so, officials perform a machine audit in that precinct. Given that at least one ballot is sampled in a precinct, the probability of sampling k ballots in that precinct is:

$$\frac{\binom{v_i}{k} p^k (1-p)^{v_i-k}}{1-(1-p)^{v_i}} \qquad (2)$$

Following the machine audit, officials randomly select the precinct's sample size based on this distribution. As

---

[7] A pseudorandom generator accepts a truly random "seed" as input, and produces a stream of outputs that can safely be treated as random for certain purposes. The details of which pseudorandom generator to use and how to use it are beyond the scope of this paper, but we note that it is crucial to get these details right.

[8] Because the output of a pseudorandom generator depends deterministically on the random seed it is given, an adversary with knowledge of the seed prior to the machine audit could determine which serial numbers will be sampled for manual review following the machine audit. Such an adversary could collude with the audit machine to hide fraud under serial numbers that will not be sampled.

before, officials should generate a new seed immediately following the machine audit of selected precincts. The resulting distribution of ballots chosen is the same as it would have been had we simply chosen each ballot (independently) with probability p.

We would expect the methods in this section to yield a slightly greater number of manual ballot reviews than the methods in Section 3.1. Intuitively, this is because Section 3.1 guarantees some fixed number of manual ballot reviews. The methods in this section offer no such guarantee, so we cautiously sample a greater expected number of ballots. Section 4 supports this conclusion, but the difference is practically insignificant.

3.3 Comparison to the Method of Rivest

Assume use of the audit method in Section 3.2, and let $p = 1 - (1 - c)^{1/B}$. The probability that precinct i requires a machine audit is therefore $1 - (1 - c)^{v_i/B}$. If an adversary can steal any number of votes in a precinct without generating suspicion, Rivest [11] proposes a logistic precinct-based approach that yields the same precinct audit probability. For machine-assisted auditing, however, auditors need only manually review a subset of the machine-audited ballots while the precinct-based approach requires manual review of all ballots in those precincts.

Rivest presents his logistic approach as a non-optimal heuristic [11], so the usefulness of this link seems limited. Furthermore, the percentage of votes in a precinct that one may steal without generating suspicion is more likely 10–20% than the 100% assumed here. In light of this, a performance comparison between Rivest's optimal precinct-based techniques and our methods under realistic circumstances would be informative.

4 Evaluation

To evaluate the efficiency of machine-assisted auditing (and ballot-based auditing in general) versus precinct-based auditing,[9] we consider both techniques in the context of available data from Virginia's November 2006 elections, both for local and statewide races.[10] In all cases, we seek a 99% confidence level.

Virginia contains 2,599 precincts and approximately 4.6 million registered voters, nearly 53% of whom cast ballots during the November 2006 election. The general election decided nineteen issues: four statewide issues, including a U.S. Senate race and several statewide initiatives, and fifteen smaller races, such as U.S. House races. In addition, voters considered numerous local ballot issues, ranging from city council elections to school constructions projects [15]. Because auditing is typically both more important and more labor-intensive in closer races, we focus on such races, excluding consideration of races for which modification of 10% or more of the ballots would have been necessary to change the outcome. This choice rules out many of the races but leaves a set of 49 remaining. Seven of those remaining were general election issues and forty-two were local issues.

The remaining general election issues include a U.S. Senate race with a margin of victory of 0.39%, four U.S. House races, a race for the Virginia House of Delegates, and a state constitutional amendment. For those races, machine-assisted auditing using the methods in Section 3.1 would require a manual review of approximately 437 ballots on average—0.06% of the 796,469 average total ballots (see Table 1). Only the smaller House of Delegates race would require review of greater than 1% of the ballots (1.05%), and five of seven races require audit rates under 0.1%. The expected manual audit size when using the methods in Section 3.2 is 439 ballots on average—still 0.06% of the average total ballots. Precinct-based auditing would review approximately 177,849 ballots on average—22.33% of the average total ballots. In all cases, precinct-based auditing requires an expected hand count of more than 40 times as many ballots. The closely contested U.S. Senate race would require review of 2,337 of 2,370,445 ballots with machine-assisted auditing and 1,141,900 ballots on average with precinct-based auditing.

While less overwhelming, the results for local ballot issues are highly favorable as well. In this case, machine-assisted audits by Section 3.1 would review approximately 224 ballots on average—2.28% of the 9,842 average total ballots. If using machine-assisted audits by Section 3.2, the expected number of ballots manually reviewed is approximately 226 on average—2.29% of the average total ballots. Precinct-based audits would require manual review of approximately 3,928 ballots on average—39.91% of the average total ballots. Only five of the forty-two races would require a manual review of more than 50% of the ballots with machine-assisted audits (using either sampling method). In contrast, only six of the forty-two races would require a review of *less* than 50% of the ballots on average with precinct-based audits. Precinct-based audits would require a complete audit in more than half of the cases.

---

[9] For precinct-based auditing, we use the methods and assumptions in [14]: auditors choose precincts uniformly at random, an adversary may switch no more than a set percentage of the votes in a precinct without arousing suspicion (we use 10%), and the adversary may switch votes in the largest possible precincts.

[10] We consider all races from the available Virginia data [15]. Some local races are absent, so we ignore those. Due to minor absences in the data set, we assume that no voter submitting a ballot abstains from voting on an issue and that voters for multi-seat races submit multiple ballots rather than a single ballot with multiple selections. While these assumptions slightly affect the realism of the tests, they likely had only a minor impact on the overwhelming results.

The races that are particularly difficult for machine-assisted auditing are town council, city council, and school board races with 7/492, 5/849, 12/769, 7/246, and 3/2409 margins of victory—requiring manual review of 68.3%, 78.4%, 53.4%, 68.3%, and 90.0% of ballots respectively (by Section 3.1). In each of these cases, precinct-based auditing would require a full audit.

If comparing machine-assisted audits and precinct-based audits purely on the number of manual ballot reviews, these results indicate a conclusive advantage for machine-assisted audits.

5 Extensions

In this section, we consider a number of methods for increasing the efficiency, practicality, and utility of machine-assisted audits.

5.1 Handling Misreadings

With some small probability, auditors might misread a paper ballot and falsely conclude that it either does or does not match the corresponding electronic ballot. Accidentally concluding that the two versions of a ballot do not match is not an issue: auditors would certainly immediately double-check any such ballots. The opposite error would be more serious. We would expect its probability to be low, however, especially in larger elections. In that case, the number of ballots to check per precinct will often be relatively small, meaning that auditors are less likely to become careless. In addition, the state may request and double-check copies of the paper ballots against the reported electronic ballots.

If auditor error is a serious risk, Johnson [7] offers a starting point for adapting sample sizes to overcome such errors, assuming use of the audit techniques in Section 3.1. If officials are instead using the methods of Section 3.2, these errors are easy to manage. Suppose that an auditor misclassifies a mismatch as a match with probability m. In this case, the true probability of detecting a bad ballot will not be p but will instead be $p(1-m)$. Thus, p must be chosen such that $p \geq \left[1-(1-c)^{1/B}\right]/(1-m)$.

5.2 Early Returns

A variety of circumstances may result in delayed reporting from certain precincts. Precincts that report in a timely manner might wish to begin the audit process without waiting hours or days for a complete initial tally. Given partial returns, auditors may assume reasonable or worst case scenarios from the remaining precincts and begin the audit under those assumptions. Once all precincts have reported, unexpected results might force additional sampling from previously reported precincts, but the bulk of the audit process may already be complete.

An intuitive explanation will help to convey how this approach works. Imagine that one group of precincts reports vote tallies early, and a second group reports later. Once the first group reports, we guess how much auditing will ultimately need to be done in the first group. We then audit the first group at this level, while we wait for the second group of precincts to report. When the second group reports, we will have the full results, so we will know whether we did enough auditing in the first group. If we did enough auditing, then we need only complete the audit by auditing the second group. (We might have done a bit more auditing in the first group than turned out to be necessary, but this extra auditing does not harm.) If it turns out that we did not do enough auditing in the first group, then we will have to return to the first group and do some additional auditing to make up the difference. Regardless, we will end up having done at least enough auditing in every precinct, so the result will be correct. Whether this approach is faster and cheaper than our original scheme depends on the accuracy of our guess about how much auditing to do.

The details depend on whether we are following the constant sample size methods of Section 3.1 or the variable sample size methods of Section 3.2.

Assuming the constant sample size methods of Section 3.1, auditors would estimate both the proportion of ballots cast in previously reported precincts (q) and the necessary sample size (n'). They may then select $d \geq n'q$ ballots from the already-reported precincts. No harm beyond the additional labor is done if auditors sample more ballots than necessary, so officials may conservatively choose a larger value of d. After reporting is complete, auditors could compute the true sample size, create a one-to-one mapping between all reported ballots and $\{1, \ldots, N\}$, and randomly select n values from that set. If d or fewer of the selected values correspond to ballots in previously reported precincts, no additional sampling is necessary in those precincts. If more than d values correspond to those precincts, that number minus d additional ballots must be drawn from the precincts. Similarly, auditors must select a number of ballots from the late reporters equal to the number of selected values corresponding to ballots in those precincts.

Using the variable sample size methods of Section 3.2, auditors would calculate p' based on the expected number of switched ballots required to change the outcome and begin sampling. Once all results are reported, officials may calculate the true value of p and use it for newly reported ballots. If the final margin of victory is smaller than expected, they also must sample previously reported but unsampled ballots with probability $p'' = 1 - \frac{1-p}{1-p'}$, yielding an overall selection probability of p for those ballots.

5.3 Varying Probability by Precinct

In Section 3.2, the sampling process selects each ballot with equal probability. That process need not do so. For example, officials may prefer to reduce the probability that ballots are selected in extremely small precincts, thereby reducing the probability that machine audits (and audit machines) will be necessary for a large number of small precincts. The only constraint that the audit process must satisfy is that, given any set of ballots of size B with corresponding selection probabilities $p_1, \ldots, p_B$, $(1 - p_1) \ldots (1 - p_B) \leq 1 - c$. Thus, auditors may entirely ignore small precincts in some cases provided that they increase the probability of sampling ballots in other precincts to compensate.

5.4 Hybrid Strategies

The line between precinct-based auditing and machine-assisted ballot-based auditing need not be so fine. States could use machines to perform audits in randomly selected precincts then audit the machine results.[11] This is similar to a dial that auditors could turn. Assume a desired confidence level c. At one extreme, machines could audit all precincts, and auditors could sample ballots such that the overall probability of uncovering a bad ballot—if enough are bad to impact the outcome—is c. This is a machine-assisted ballot-based audit approach.

At the other extreme, auditors could select precincts such that the probability that at least one selected precinct contains a bad ballot—if enough are bad to impact the outcome—is c, and auditors could manually check all ballots in those precincts. This is precinct-based auditing. In either case, the probability of detecting fraud significant enough to affect the election's outcome is c. Between these two possibilities, one could trade a greater expected number of machine audits for a smaller expected proportion of manual audits and vice versa to achieve a confidence level of c. Depending on the costs and benefits of each, states may choose whatever balance is most appropriate for their specific circumstances.

5.5 Considering Ballot Contents

Consider a two-candidate mayoral race in which the electronic results indicate that Alice beat Bob 11,000 to 10,000. Traditional audit techniques would require that officials consider ballots containing votes for either candidate even though the primary objective is to discover whether any votes for Alice should have been for Bob. Using the method of Section 3.2 and considering all ballots, we would expect to audit 193 ballots to get 99% confidence in the election result. Examining only ballots reported to contain votes for Alice could cut auditors' work nearly in half, as auditors seek to discover an equivalent amount of fraud in a far smaller pool of ballots. In the example race, this option would reduce the expected number of ballots audited to 101, a 48% reduction.

In general, by considering the contents of ballots, officials may reduce the number of manual verifications required. For the remainder of this section, we assume use of the audit process in Section 3.2.

Generalizing this idea, we describe below an algorithm that bases the probability of auditing a particular ballot on the (electronically reported) contents of that ballot, in a race with any number of candidates and seats. Our algorithm is correct in the presence of undervotes (ballots marked for fewer than the maximum allowed number of candidates), and it copes with overvotes (ballots marked for too many candidates) and other types of spoiled ballots by treating them as if they were undervoted ballots with no candidates marked.

Assume a race in which n candidates are competing for k seats, and let $v_1, \ldots, v_n$ be the electronically reported vote totals for the candidates in decreasing order. Therefore $v_1, \ldots, v_k$ correspond to winning candidates. Because a single ballot may contain votes for up to k candidates, we need to consider the combination of votes on each ballot.

Given a ballot, let C_s, where $1 \leq s \leq k$, be the winning candidate with the lowest vote total that received a vote on the ballot. (Let C_s be null if the ballot does not contain votes for any winning candidate.) Let C_t, where $k + 1 \leq t \leq n$, be the losing candidate with the highest vote total that did not receive a vote on the ballot. (Let C_t be null if the ballot contains votes for all of the losing candidates.) We must consider each of the following four cases.

- If C_s and C_t are both null, then there is no need to audit this ballot. Intuitively, this ballot is already helping the (reported as) losing candidates as much as it can and helping the (reported as) winning candidates as little as it can, so any error in this ballot cannot have helped a (reported as) winning candidate wrongly take a seat from a (reported as) losing candidate.

- If C_s is non-null, then we need to audit this ballot with probability at least $1 - (1 - c)^{1/b_1}$, where $b_1 = v_s - v_{k+1}$. Intuitively, one possible result-changing

---

[11] Mock ballot selection and Section 5.3 technically do this, but the probability of selecting a precinct need not be directly based on the probability of selecting its underlying ballots.

scenario involving an error in this ballot would be to add $v_s - v_{k+1}$ incorrect votes for candidate s.

- If C_t is non-null, then we need to audit this ballot with probability at least $1-(1-c)^{1/b_2}$, where $b_2 = v_k - v_t$. Intuitively, one possible result-changing scenario involving an error in this ballot would be to remove $v_k - v_t$ true votes for candidate t.
- If C_s and C_t are both non-null, then we need to audit this ballot with probability at least $1-(1-c)^{1/b_3}$, where $b_3 = \min\left(b_1, b_2, \lceil \frac{v_s-v_t}{2} \rceil\right)$. Intuitively, one result-changing scenario involving an error in this ballot would be to transfer $\lceil \frac{v_s-v_t}{2} \rceil$ votes from candidate t to candidate s. In addition, we include b_1 and b_2 as the two previous scenarios are still possible in this case.

If more than one case applies, we must audit the ballot with at least the highest minimum probability dictated by any individual case.

If ballots contain votes for more than one race, we must manually check ballots with the maximum probability necessary for any individual race/vote combination on that ballot.

5.6 Considering Initial Returns

Similar tricks may also be useful given only reported initial electronic vote tallies and accurate counts of the number of ballots per precinct. A precinct in which initial tallies indicate that all ballots contain votes for Bob could not have contributed to discrepancies affecting the election's outcome, so both machine-assisted and precinct-based auditing could ignore that precinct entirely. In a single-seat race, a ballot may contain a single vote at most, so we may determine the precise ballot contents for that race from the initial electronic tallies alone and use that to calculate an appropriate probability of manually verifying each ballot.

By inferring ballot contents for the 2006 Virginia general election, we are able to apply the methods in Section 5.5. This reduces the expected number of manual ballot reviews to an average of 227—0.03% of the average total ballots and a 47.96% decrease from the methods of Section 3.1 (see Table 2). In addition, this option reduces the expected number of precincts requiring machine audits by 35.26% to 157 on average and the expected number of ballots to be machine audited by 31.93% to 202,431 on average. When applied to the close U.S. Senate race, the expected number of ballots to be manually reviewed decreases by 49.56%, the expected number of precincts to be machine audited decreases by 36.87%, and the expected number of ballots to be machine audited decreases by 33.24%.

As we increase the number of available seats, competing candidates, and races on a ballot, inferences tend to become more difficult and less beneficial. The added complexity is a result of an increase in the number of possible vote combinations on a ballot. One may still draw inferences from the ballots, however. For example, if 43% of ballots contain a vote for the "just losing" candidate, we know that 43% of the ballots could not have had a vote for that candidate removed or switched to another candidate. As vote totals in complex multi-seat races become more tightly clustered, the complication of drawing inferences might counterbalance the increasingly minor benefits of those inferences. A test of these methods on real elections data might help to better establish which cases benefit from these techniques.

5.7 Write-Ins

Machine-assisted audits can easily handle the case in which all write-in candidates are put in a machine-readable form, whether by voters or by election officials. They also may handle the case in which voters may not write-in a candidate already appearing on the ballot by treating all write-in votes as votes for a single additional candidate. Otherwise, if the number of write-in votes is insufficient to affect the outcome of the election (given initial tallies), we may assume whatever combination of write-in votes results in the closest possible election and search for the necessary quantity of fraud in ballots not containing write-in votes.

If write-in votes could change the results of the election, a count of those votes will ultimately be necessary before certification of the election. If the count of ballots with write-ins is manually performed, we may simply manually audit the remaining, machine-audited ballots to discover any fraud large enough to affect the outcome. If write-in ballots are machine-scanned, we may add a serial number printer to that scanner and sample ballots from the full pool. Given the relatively small quantity of write-ins in many elections, we expect this to rarely be a significant issue in practice.

5.8 Machine Malfunction

Presumably, some percentage of audit machines will occasionally fail. While a failure would cause a delay in the audit process for the affected precinct, the delay would only be for that precinct. All other precincts could proceed normally, and the affected precinct could wait for repairs or obtain an audit machine from a completed precinct. If voting machines fail, the failures could delay initial tally reports from the affected precincts. As Section 5.2 indicates, such a delay need not hold up other precincts.

5.9 Candidate Assurance

To give candidates additional assurance that the audit process did not miss or under-sample precincts in which fraud seems apparent, Appel explores the idea of allowing candidates to select a small number of additional precincts and pay for full manual audits of those precincts (candidates are reimbursed if errors are uncovered) [3]. The proposals in this paper can be easily combined with this assurance process. Even if the possibility is unlikely, any process short of a total audit could occasionally miss fraud obvious to a human. In addition, candidate selection of precincts provides those skeptical towards the audit process with an alternative route of uncovering fraud. Appel [3] explains and motivates this idea further.

6 Conclusion

A well-designed audit process assures the public that an election's outcome is the product of voters' intentions, not fraud or voting machine flaws. By adding a novel machine-assisted audit procedure to ballot-based audits, we can enjoy the efficiency benefits of those audits while avoiding privacy concerns and retaining the security benefits of combined paper/electronic solutions. Our tests using data from Virginia's November 2006 elections confirm the efficiency advantages of machine-assisted audits, and the extended techniques that we propose promise to reduce even further the number of ballots that need to be inspected by humans.

Though future work is needed to better estimate the costs of machine-assisted audits and to assess other practical challenges that election officials face, we believe that the techniques in this paper offer a promising alternative to traditional precinct-based auditing and warrant further study.

Acknowledgments

We thank Andrew Appel, Ari Feldman, Shirley Gaw, David Wagner, Harlan Yu, and Bill Zeller for helpful discussions. We also thank Ronald Rivest and Joe Hall for pointing us towards useful papers and resources and the anonymous reviewers for helpful comments.

Calandrino performed this research while under appointment to the Department of Homeland Security (DHS) Scholarship and Fellowship Program under DOE contract number DE-AC05-06OR23100. All opinions expressed in this paper are the authors' and do not necessarily reflect the policies and views of DHS or DOE.

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] 110TH CONGRESS. H.R. 811: Voter confidence and increased accessibility act of 2007.

[2] ACE ELECTORAL KNOWLEDGE NETWORK. Plurality/majority systems, 2006. http://aceproject.org/ace-en/topics/es/esd/esd01/.

[3] APPEL, A. W. Effective audit policy for voter-verified paper ballots in New Jersey, February 2007. http://www.cs.princeton.edu/~appel/papers/appel-nj-audits.pdf.

[4] CORDERO, A., WAGNER, D., AND DILL, D. The role of dice in election audits - extended abstract. In *IAVoSS Workshop on Trustworthy Elections 2006*.

[5] DUNN, S. Voter verifiable paper audit trail pilot project, Cobb County, Georgia, November 2006. http://www.gaforverifiedvoting.org/docs/Cobb_county_pilot_report.pdf.

[6] FELDMAN, A., HALDERMAN, J. A., AND FELTEN, E. Security analysis of the Diebold Accuvote-TS voting machine. In *Proc. 2007 USENIX/ACCURATE Electronic Voting Technology Workshop (EVT '07)*.

[7] JOHNSON, K. C. Election certification by statistical audit of voter-verified paper ballots, October 2004. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=640943.

[8] KOHNO, T., STUBBLEFIELD, A., RUBIN, A., AND WALLACH, D. Analysis of an electronic voting system. In *Proc. 2004 IEEE Symposium on Security and Privacy*, pp. 27–42.

[9] NEFF, C. A. Election confidence: A comparison of methodologies and their relative effectiveness at achieving it, December 2003. http://www.votehere.net/papers/ElectionConfidence.pdf.

[10] RIVEST, R. L. On estimating the size of a statistical audit, November 2006. http://people.csail.mit.edu/rivest/Rivest-OnEstimatingTheSizeOfAStatisticalAudit.pdf.

[11] RIVEST, R. L. On auditing elections when precincts have different sizes, April 2007. http://people.csail.mit.edu/rivest/Rivest-OnAuditingElectionsWhenPrecinctsHaveDifferentSizes.pdf.

[12] RIVEST, R. L., AND WACK, J. P. On the notion of "software independence" in voting systems, July 2006. http://vote.nist.gov/SI-in-voting.pdf.

[13] SALTMAN, R. G. Effective use of computing technology in vote-tallying. Tech. Rep. NBSIR 75-687, National Bureau of Standards, March 1975.

[14] STANISLEVIC, H. Random auditing of e-voting systems: How much is enough?, August 2006. http://www.votetrustusa.org/pdfs/VTTF/EVEPAuditing.pdf.

[15] VIRGINIA STATE BOARD OF ELECTIONS. General election — November 7, 2006. http://www2.sbe.virginia.gov/web_docs/Election/results/2006/Nov/htm/index.htm.

General Election

Issue	Totals			Mch-Asst Auditing (by Sect. 3.1)			Mch-Asst Auditing (by Sect. 3.2)			Pct-Based Auditing		
	# Pct	# Votes	Diff	# Pct	# Bal (Mch)	# Bal (Man)	# Pct	# Bal (Mch)	# Bal (Man)	# Pct	# Bal (Mch)	# Bal (Man)
U.S. Senate	2,599	2,370,445	0.39%	1,350	1,615,277	2,337	1,351	1,615,721	2,339	1,252	-	1,141,900
Const. Amnd.	2,599	2,328,224	14.12%	62	86,134	63	64	88,872	65	9	-	8,062
U.S. House	158	173,159	2.82%	123	153,729	325	123	153,816	327	57	-	62,469
U.S. House	325	212,079	19.19%	42	40,072	46	43	41,469	48	3	-	1,958
U.S. House	197	241,134	16.36%	45	71,490	54	47	73,745	56	5	-	6,120
U.S. House	163	235,280	11.88%	57	100,202	76	58	101,365	77	9	-	12,991
Delegate	17	14,963	5.75%	14	14,897	157	14	14,899	159	13	-	11,442
Average	865	796,469	8.11%	242	297,400	437	243	298,555	439	193	-	177,849

Local Issues

Issue	Totals			Mch-Asst Auditing (by Sect. 3.1)			Mch-Asst Auditing (by Sect. 3.2)			Pct-Based Auditing		
	# Pct	# Votes	Diff	# Pct	# Bal (Mch)	# Bal (Man)	# Pct	# Bal (Mch)	# Bal (Man)	# Pct	# Bal (Mch)	# Bal (Man)
Average	10	9,842	10.54%	8	8,717	224	8	8,741	226	5	-	3,928

Table 1: Machine-Assisted vs. Precinct-Based Auditing on VA Election Data (note: number of precincts for machine-assisted auditing by Section 3.1 calculated as an average over 10,000 trial simulations)

Issue	Precincts		Ballots (Mch-Audit)		Ballots (Man-Audit)	
	Number	Decrease	Number	Decrease	Number	Decrease
U.S. Senate	853	36.87%	1,078,423	33.24%	1,179	49.56%
Const. Amnd.	37	40.48%	48,927	43.20%	37	40.94%
U.S. House	93	24.14%	123,739	19.51%	172	46.97%
U.S. House	30	28.83%	28,149	29.75%	32	30.95%
U.S. House	32	29.48%	51,209	28.37%	36	33.16%
U.S. House	39	32.15%	71,793	28.35%	47	38.12%
Delegate	13	3.93%	14,777	0.81%	88	43.82%
Average	157	35.26%	202,431	31.93%	227	47.96%

Table 2: Machine-Assisted Auditing when Inferring Ballot Contents vs. Machine-Assisted Methods of Section 3.1