National Election Data Archive
P.O. Box 682556 Park City, UT 84068
Phone: 435-658-4657

History of Confidence Election Auditing Development
(1975 to 2008)
& Overview of Election Auditing Fundamentals

Kathy Dopp M.S. Reviewers: Ron Baiman Ph.D., Frank Stenger Ph.D.

October 22, 2007 updated 11/07, 12/07, 2/08, 4/08, 8/3/08, 9/08

CONTENTS

1975 to 2005 – Confidence-Level Election Audits - Recommended but Ignored....3
2005 – Verified Voting, National Election Data Archive, Carter-Baker Commission, and US GAO....4
2006 – Brennan Center, NSF Project ACCURATE, League of Women Voters US, and US EAC Technical Guidelines Development Committee....6
2007 – American Statistical Association and California Secretary of State Bowen....9
2008 – Probability Proportional to Error Bound Methods for Election Auditing....16
Agreement on Election Auditing Fundamentals....18
Order of Steps for Conducting Election Audits....19
What Remains to Be Done to Implement "Confidence-Level" Election Audits?....21
CONCLUSION....21
Acknowledgments....23
Appendix A: Incorrect Auditing Recommendations That Would Not Detect Faulty Outcomes....24
Appendix B: How Many Miscounted Vote Counts Could Alter An Election Outcome?....26
Appendix C: Certain Audit Calculations Must be Based on "Ballots Cast", Not "Votes Counted"....29
Appendix D: 100% Hand-Counts Are Rarely Needed To Confirm Election Results....31
Appendix E: Discrepancy Conditions Requiring Expanded Audits or Full Counts....32
Appendix F: County-Level Election Auditing....33
Appendix G: The Debate over Which Value to Use for Max Miscount per Vote Count....34
Appendix H: Derivation of the Probability Proportional to Margin Error Bound (PPMEB) Method... 35

This document is found online at
http://electionarchive.org/ucvAnalysis/US/paper-audits/History-of-Election-Auditing-Development.pdf

Definition: *Confidence-level election audits* are post-election manual "eye-to-hand" counts of voter-verified paper ballots to check the accuracy of unofficial machine counts where the audit sample size is calculated to ensure that if vote miscount alters an election outcome as reported in the unofficial results, then that miscount is detected so that it can be corrected and the winners of elections can be certified with confidence.[1] In precise mathematical language, "Confidence-level" means 100% minus the maximum chance that an incorrect outcome is certified. For example, certifying an election outcome at 95% confidence-level means that either the outcome is correct, or something with chance at most 5% occurred.

Abstract

Developers of election auditing methods are asking two questions:

1. What election audit sample size would detect, to any desired confidence level, the smallest amount of vote miscount which could result in an incorrect election outcome?
2. Given the discrepancies found in an election audit, what confidence level is there that the election outcome is correct, and should we certify the election or expand the audit?

This paper provides a chronological review of the development of confidence-level election auditing methods as well as descriptions of

- fundamentals that election auditing experts agree on,
- work that must be done before verifiable, efficient, and effective confidence-level election audits can be implemented,
- incorrect election auditing recommendations made by some authors, and
- methods for conducting confidence-level election audits.[2] Some election audit methods are described for the first time in this paper.

Election auditing works are listed in chronological order in rows of a table. The work's author(s), name, date, and URL are in the left column, and a review of each work is in the right column.

[1] This paper does not discuss other types of audits that are designed to prevent and detect errors in machine programming and errors in ballot definition files, including pre-election ballot definition audits, pre-election, parallel, and post-election voting machine testing, or audits of machine programming code. Nor does this paper discuss post-election ballot sampling which are akin to taking exit poll samples of ballots in order to make statistical, rather than precise comparisons with the machine counts.

[2] Some methods are shown in the "Agreement on Election Auditing Fundamentals" section. Appendix A shows which candidate margins to use for calculating audit samples, even in multi-candidate and multi-winner races and describes a short numerical algorithm for how to precisely calculate the minimum number of miscounted auditable vote counts (precinct, batch, or machine counts) which could reverse an election outcome, which can be easily implemented in a program or in a spreadsheet. Appendix B provides the method for calculating candidate margins using the number of ballots cast, and shows why this works correctly, as opposed to calculating margins as a percentage of votes cast.

1975 to 2005 – Confidence-Level Election Audits - Recommended but Ignored

Author, Work, Date	Remarks
1. Roy G. Saltman, *Effective Use of Computing Technology in Vote-Tallying*, Technical Report NBSIR 75--687, National Bureau of Standards, Information Technology Division, (March 1975), available at: http://csrc.nist.gov/publications/nistpubs/NBS_SP_500-30.pdf	Proposes election audit sample sizes based on margins between candidates. Derives a formula for calculating the minimum number of miscounted vote counts[3] which could be used to wrongly alter an election outcome. However, Saltman's work assumes that there are no overvotes or undervotes. His work was largely ignored until July 2006 when re-derived in a way that handles overvotes and undervotes. (See Appendix B)
2. Walter R. Mebane, Jr., Jasjeet S. Sekhon, and Jonathan Wand, *Detecting and Correcting Election Irregularities*, (October 9, 2003), available at: http://wand.stanford.edu/research/detecting.pdf	A policy paper recommending election auditing, suggests "rather than relying on ad hoc detection and litigation of electoral problems" that we "systematically monitor and audit elections in a preventative fashion."
3. C. Andrew Neff, *Election Confidence---A Comparison of Methodologies and Their Relative Effectiveness at Achieving It* (Revision 6), (December 17, 2003), available at: http://www.electionmathematics.org/em-audits/US/Neff-ElectionConfidence.pdf	Espouses a "vote receipt methodology" described as "issuing indisputable receipts for either "real ballots", or "test ballots" to "establish election confidence". Although this paper overlooks the fact that an electronic fraudster can overcome any "test ballot" auditing, Neff's point is true that smaller unit sizes (in his case single ballots) are more efficiently audited. However, today's voting systems are not auditable on the individual ballot level.[4]
4. Jonathan Wand, *Auditing an Election Using Sampling: The Impact of Bin Size on the Probability of Detecting Manipulation* (version as of February 17, 2004), (February 2, 2004), available at: http://wand.stanford.edu/elections/probability.pdf	Shows mathematically that the efficiency and effectiveness of election audits depends upon the size of vote counts (audited units or "bin size"). Auditing the accuracy of a larger number of smaller counts (containing fewer ballots in each) is more efficient than auditing the accuracy of fewer numbers of larger vote counts.

[3] An auditable "vote count" is a count of any collection of ballots reported in the unofficial election results, for instance precinct or machine vote counts, or a batch of mail-in ballots counted together – as long as their counts are recorded and publicly available prior to the audit.

[4] Ballot-level auditable voting systems are not in use today to my knowledge. However some planned voting systems will randomly generate machine and human-readable tracking numbers which will be printed on a paper ballot when a voter deposits the ballot into a ballot box (or casts the ballot and it automatically drops into a ballot box in the case of ballot marking devices for voters with disabilities or foreign language needs). The ballot tracking numbers may be used to randomly identify and check individual ballots for auditing. The central tabulator and voting machines would have to be capable of printing reports of the vote counts on each ballot with its tracking number.

Author, Work, Date	Remarks
5. Douglas W. Jones, *Auditing elections*, Communications of the ACM 47:10 46-50, (October 2004), available at: http://doi.acm.org/10.1145/1022594.1022622 http://www.cs.uiowa.edu/~jones/voting/cacm2004.shtml and some Miami-Dade documents: http://www.electionmathematics.org/em-audits/FL/	Gives logical arguments for auditing elections and describes some fundamental requirements for election auditing to be effective.
6. Kenneth C. Johnson, *Election Certification by Statistical Audit of Voter-Verified Paper Ballots* (October 31, 2004) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=640943	Analyzes two election audit processes: a conventional statistical recount, and a statistical error-count that detects discrepancies between ballot database records and corresponding paper ballots; and recommends using the error-count method. This recommendation requires that voting systems are auditable on the individual ballot level, which most are not yet.[5]

2005 – Verified Voting, National Election Data Archive, Carter-Baker Commission, and US GAO

Author, Work, Date	Remarks
7. *Testimony of David L. Dill, Professor of Computer Science, Stanford University and Founder of the Verified Voting Foundation and VerifiedVoting.org before the Commission on Federal Election Reform* (The Carter-Baker Commission), American University, Washington D.C. (April 18, 2005) http://www.verifiedvotingfoundation.org/article.php?id=5987	Recommends routine auditing of elections by choosing a small random sample of ballots and manually counting them[6] to "catch procedural, equipment, and personnel problems before they affect an election outcome." Says precinct-count optical scan voting, in which a voter marks a paper ballot that is counted by a machine in the polling place, is "highly accurate, and much less costly than touch-screen voting", and "computerized ballot-marking devices make the system fully accessible to voters with disabilities."
8. Kathy Dopp & Ron Baiman, *How Can Independent Paper Audits Detect and Correct Vote Miscounts?* (version as of July 25, 2006), (June 30, 2005), available at: http://electionarchive.org/ucvAnalysis/US/paper-audits/Paper_Audits.pdf	Explains election auditing mathematics in a short handout.[7]

[5] Ibid footnote #4.

[6] Ballot sampling is incapable of detecting actual discrepancies between manual and machine counts unless voting systems are designed to be auditable on the ballot level. It is unclear whether Dill was recommending exit-poll like statistical comparisons of ballot samples with unofficial results, or machine testing by recounting sampled ballots, or meant to recommend choosing a random sample of unofficial vote counts rather than ballots to audit.

[7] This handout was widely distributed on the Internet and at conferences of election activists, and to State election officials at a NASS and NASED conferences, and submitted to the US GAO and submitted as testimony to the Carter-Baker Commission's Director Robert Pastor.

Author, Work, Date	**Remarks**
9. *North Carolina Session Law SL 2005-323 The "Public Confidence in Elections Act"* (ratified August 26, 2005); and SL 2006-192 HB 1024 (ratified July 27, 2006) Section 7.(a) G.S. 163-182.1(b) and Section 7.(b) G.S. 163-182.2(b): http://www.ncleg.net/Sessions/2005/Bills/House/HTML/H1024v7.html http://www.ncleg.net/Sessions/2005/Bills/Senate/HTML/S223v7.html http://www.ncvoter.net/downloads/Kalsbeek_May2006_Recount_Assessment.pdf	Requires hand-to-eye auditing of "ballots of full precincts, full counts of absentee ballots, and full counts of early voting sites." "The size of the sample of each category shall be chosen to produce a statistically significant result and shall be chosen after consultation with a statistician"[8] ..."If the discrepancy…is significant, a complete hand-to-eye count shall be conducted." However, audit sizes depend on the definition of "significant" of one statistician whom authorities hire, and only one state-wide race (the presidential race in presidential election years) is audited.
10. Ellen Theisen, *Auditing Election Equipment --- The Real Scoop!* (August 27, 2005) http://www.votersunite.org/info/auditingissues.pdf Movie available at: http://homepage.mac.com/sheltonlankford/.Public/RandomSample.mov and excel spreadsheet at: http://www.votersunite.org/info/AuditEffectivenessCalculator.xls	Recommends against auditing elections saying "Each election has a unique set of contests and candidates, and each county prepares many different, unique ballots … a 'statistically significant' sampling is a meaningless phrase in this context."
11. *Building Confidence in U.S. Elections – Report of the Commission on Federal Election Reform*, Center for Democracy and Election Management, American University (September 2005) http://www.american.edu/ia/cfer/report/full_report.pdf	The Carter-Baker Commission calls for "regular audits … to double-check the machines' vote tallies" (p. 36)
12. United States Government Accountability Office (US GAO), *Federal Efforts to Improve Security and Reliability of Electronic Voting Systems are Under Way, but Key Activities Need to be Completed* (September 2005) http://www.gao.gov/new.items/d05956.pdf	The US GAO recommends "A post election audit of voting systems should be conducted to reconcile vote totals and ballot counts, even if there is no recount scheduled." (page 46)

[8] William D. Kalsbeek Ph.D. , a professor in the department of Biostatistics and director of the Survey Research Unit at the UNC School of Public Health was hired by North Carolina to calculate their audit amounts. Kalsbeek is a fellow of the American Statistical Association.

Author, Work, Date	Remarks
13. Paul Walmsley, *Requirements for Statistical Live Auditing of Optical-Scan and VVPAT Records – and for Live Auditing for Vote Tabulation*. (September 30, 2005) http://www.booyaka.com/~paul/ea/eac-20050930/live-audit-overview.txt http://www.booyaka.com/~paul/ea/eac-20050930/interpretation-live-audit.txt http://www.booyaka.com/~paul/ea/eac-20050930/tabulation-live-audit.txt	Recommends "live" audits conducted during elections; provides requirements for voting systems to be auditable at the ballot level[9], and provides procedures for conducting election audits in two stages: (1) verifying the accuracy of ballot casting, scanning and interpretation and (2) verifying the accuracy of the vote tabulation component.[10]

The Carter/Baker Election Reform Commission and the U.S. Government Accountability Office, in the fall of 2005, join Verified Voting, and the National Election Data Archive in recommending routine election audits.

2006 – Brennan Center, NSF Project ACCURATE, League of Women Voters US, and US EAC Technical Guidelines Development Committee

Author, Work, Date	Remarks
14. Brennan Center Task Force on Voting System Security (Lawrence Norden, Chair), *The Machinery of Democracy: Protecting Elections in an Electronic World*, (March 7, 2006) http://www.brennancenter.org/programs/downloads/Full%20Report.pdf	Recommends parallel voting machine testing, using the concept of maximum "non-suspicious" vote miscount per voting machine[11]; and recommends manual audits of machine counts.

[9] Ibid footnote #4.

[10] Walmsley's two stage election auditing method suggests a similar solution that would increase the efficiency of auditing certain voting systems which, although not ballot-level auditable as his method requires, could nonetheless be audited more efficiently by breaking the audit into two stages - a verification of machine records and a separate verification of central tabulation – provided that the same machine records were used in both stages of the audit - although the sample sizes for each audit stage would need to be calculated separately whenever different auditable vote count units are used in the two stages. Walmsley's suggestion to conduct "live" auditing during elections, without disrupting the election process, merits serious consideration because that would reduce concern over securing ballots and election records between the election and the audit.

[11] This is similar to Saltman's concept (1975) of "maximum level of undetectability".

15. Dopp and Joycelynn Straight, *Legislative Vote Count Audit Proposal* – Written for Utah but never sponsored, (May 2006, updated in November 2006 to require confidence level election auditing sample sizes and in November 2007 for showing to Utah legislators) http://electionarchive.org/ucvAnalysis/US/paper-audits/VoteCountAudit-UT.pdf	A legislative proposal for independent election audits requiring procedures that would make election outcomes (the winners of elections) publicly verifiably accurate.
16. The League of Women Voters of the United States (LWVUS), *Resolution on U.S. Voting Systems under HAVA* (as amended and passed at the June, 2006 Convention) http://uscountvotes.org/ucvInfo/US/LWVUS-Resolution2006.pdf	The League of Women Voters adopts a position in favor of voting systems with voter-verifiable paper ballot as the official record of the voter's intent that the voter can verify during the process of voting; and routine audits in randomly selected precincts.
17. Arel Cordero, David Wagner, and David Dill, *The Role of Dice in Election Audits*--- Extended Abstract, IAVOSS Workshop On Trustworthy Elections (WOTE 2006), (June 29, 2006) http://www.cs.berkeley.edu/~daw/papers/dice-wote06.pdf	Recommends that the random selection method for election audits must be transparent to observers in order for voters to have confidence in the audit's results. Provides usable suggestions for transparent methods of randomly selecting election audit samples.
18. Dopp & Baiman, *How Can Independent Paper Audits Detect and Correct Vote Miscounts?* (June 30, 2005, version as of July 25, 2006) http://electionarchive.org/ucvAnalysis/US/paper-audits/Paper_Audits.pdf Note: This news was publicly released on the Open Voting Consortium, National Election Data Archive and other email lists and then published in this revised paper, while Dopp and Stenger began deriving exact numerical solutions for calculating confidence-level audit sample sizes.	Provides a formula for how much vote miscount could wrongly alter election outcomes (similar to Saltman's 1975 formula but works correctly with miscounted undervotes and overvotes if margins are correctly calculated) and provides a spreadsheet to calculate confidence-level (confidence in the outcome – who wins) election audit sample sizes.
19. Stanislevic, *Random Auditing of E-Voting Systems: How Much is Enough?* (August 9, 2006, revised August 16, 2006) http://www.votetrustusa.org/pdfs/VTTF/EVEPAuditing.pdf	Replicates Saltman's work and provides a loose algorithm for calculating the minimum number of miscounted precincts which could reverse an election outcome without noting, as Saltman did, that his method does not handle cases with miscounted undervotes and overvotes.[12].

[12] Stanislevic recommends auditing calculations that do not provide sufficient sample sizes to detect outcome-altering errors caused by misrecorded under-votes or over-votes. His imprecise methods (for adjusting for precinct-size variation and for calculating election audit sample sizes) sometimes slightly miscalculate the minimum audit sample size necessary to achieve a desired confidence level, even assuming no overvotes or undervotes. In email conversation with Dopp, Stanislevic insisted that his recommendations ignoring undervotes and overvotes were correct; and he continues to make the same recommendations as recently as the July 2007 Verified Voting paper.

20a. Election Science Institute DRE Analysis for May 2006 Primary, Cuyahoga County, Ohio, (August 2006) http://bocc.cuyahogacounty.us/GSC/pdf/esi_cuyahoga_final.pdf	An evaluation of administrative procedures, pre-election programming and testing of the voting machines, voter and booth worker interaction, and counting and auditing procedures of the May 2006 Primary in Cuyahoga County.
20. Dopp and Frank Stenger, *The Election Integrity Audit* (version as of September 25, 2006) http://electionarchive.org/ucvAnalysis/US/paper-audits/ElectionIntegrityAudit.pdf (A computer program by Frank Stenger, Kathy Dopp, & Brian Julin for calculating audit amounts http://electionarchive.org/auditcalculator/eic.cgi)	Presents a precise numerical solution to calculate the minimum election audit sample size required to ensure that an election outcome is correct to a desired confidence level, including a precise numerical method for calculating the minimum number of miscounted vote counts which could alter an election outcome, taking vote count sizes into account; provides audit procedural recommendations, and correctly handles miscounted switched votes as well as undervotes and overvotes.
21. Lobdill, *Election Audit Sampling Plan Design— It's Not Just About Sampling Without Replacement* (October 9, 2006) http://vote.nist.gov/Election-Audit-Sampling-Plan-Design-Its-Not-Just-About-Sampling-Without-Replacement-10-09-06.pdf	Describes a method that an attacker could use, and provides an imprecise algorithm for calculating election audit sample sizes that does not handle miscounted undervotes and overvotes.[13]
22. Ron Rivest, *On Estimating the Size of a Statistical Audit*, unpublished draft (September 19, 2006, version as of November 14, 2006) http://theory.lcs.mit.edu/~rivest/Rivest-OnEstimatingTheSizeOfAStatisticalAudit.pdf	Provides a formula to estimate the exact audit sample sizes found by using the Dopp-Stenger numerical method (or a trial and test method). This formulaic estimate has the advantage of being simply calculated in a spreadsheet or with a calculator; and takes as input the exact number of miscounted precincts that could alter an election outcome as found precisely by using the numerical method in Dopp-Stenger's work.[14]
23. Stanislevic, *Electronic Vote Counting: Where Did We Go Wrong?* (November 21, 2006) http://www.votetrustusa.org/index.php?option=com_content&task=view&id=2067&Itemid=26	Lauds Saltman's earlier discovery of the formula for confidence-level election auditing which does not account for overvotes or undervotes.[15]

[13] In July 2006 Lobdill announced by sending emails to the Open Voting Consortium (OVC) email discussion list that he (Lobdill) had found the numerical solution for election audit sample size calculations which had been provided to him previously via email by Frank Stenger and Dopp. After Lobdill's attempt to take credit for Stenger's work, and his making vital mischaracterizations of Dopp's work to the OVC discussion list - Lobdill was removed from Dopp's election auditing discussion group. In this paper Lobdill fails to cite the earlier Dopp-Stenger work.

[14] This algorithm is described, with a slight modification to correct for an unusual possibility, in Appendix A of this paper.

[15] Stanislevic fails to mention that Dopp had previously re-derived and publicized a variation of Saltman's earlier formula which correctly handles undervotes and overvotes (as long as the margin is correctly calculated), and had provided a spreadsheet to calculate confidence-level audits based on candidate margins. Stanislevic also fails to mention that Dopp-Stenger previously provided the exact numerical solution for calculating confidence-level election audit sample sizes even though he had reviewed the prior Dopp-Stenger work in draft stages (and made suggestions regarding its citation of his own work.)

24. Lobdill, *Considering Vote Count Distribution in Designing Election*, (November 26, 2006) http://vote.nist.gov/Considering-Vote-Count-Distribution-in-Designing-Election-Audits-Rev-2-11-26-06.pdf	Describes and elucidates confidence-level election auditing principles but recommends imprecise calculation methods which also do not correctly handle undervotes or overvotes.[16]
25. Stanislevic, *Larger Audits Required to Confirm 2006 US House Races*, (December 20, 2006) http://www.votetrustusa.org/index.php?option=com_content&task=view&id=2135&Itemid=26	Stanislevic's stated results may be believable, although he provides no supporting data or analysis. The rows in his table of values have column headings which appear to state that the "margin %" for each row is both more than and less than the same value, which is of course impossible.

In 2006, The Brennan Justice Center, Project ACCURATE (funded by The National Science Foundation), and The League of Women Voters U.S. join Verified Voting, The National Election Data Archive, the Carter-Baker Commission on Election Reform, and The U.S. Government Accountability Office, in recommending routine election audits. In 2006, The U.S. Election Assistance Commission's Technical Guidelines Development Committee, chaired by the Director of the National Institute of Standards and Technology (NIST), recommends "software independence" of voting systems which it defines as: "A voting system is software-independent if an undetected change or error in its software cannot cause an undetectable change or error in an election outcome." Manual auditing of voter verified paper ballots to check machine count accuracy makes voting systems software independent.

2007 – American Statistical Association and California Secretary of State Bowen

Author, Work, Date	**Remarks**
26. Dopp, *How Big Should an Election Audit Be? Fixed Rate Audits Do Not Work for Elections*, (January 17, 2007) http://electionarchive.org/ucvAnalysis/US/paper-audits/ElectionAuditEstimator.pdf	Shows how to derive and apply Rivest's estimate for estimating audit sample sizes to elections using as input the number of miscounted vote counts that could wrongly alter an election outcome; and provides a spreadsheet for the layperson to calculate confidence-level election audit sample sizes using Rivest's estimate to refine a trial and test method.
27. Dopp, *Tiered Election Audits Based on Margins between Candidates*, (January 23, 2007) http://electionarchive.org/ucvAnalysis/US/paper-audits/FourTierAudit/TieredElectionAudits.pdf	Explains tiered election audits[17]; provides a spreadsheet to calculate tiered election audits; and shows that minimum audit *amounts*, not just percentages, must be required for each tier if the audit is to achieve a desired confidence level down to a minimum margin.

[16] Lobdill credits himself and Stanislevic for developing election auditing calculation methods for determining "the smallest number of precincts that can be corrupted that will produce a reversal" of a tally, although the only precise numerical solution to this problem was previously publicly released by Dopp and Stenger. See footnote 12.
[17] This paper was written in response to Larry Norden's proposal that federal legislation require 3%, 5%, 10% tiered audits, depending on candidate margins in the unofficial results.

28. Norden, Burstein, Hall, Dill, Hoke, Mebane, Oakley, Rivest, and Wagner, "*Thoughts on Mandatory Audits*" (February 1, 2007) http://www.brennancenter.org/dynamic/subpages/download_file_47860.pdf	An analysis supporting the ad hoc proposal for 3%, 5%, and 10% tiered election audits for Federal elections in Rush Holt's H.R.811
29. Andrew Appel, *Effective Audit Policy for Voter-Verified Paper Ballots in New Jersey*, (February 22, 2007) http://www.cs.princeton.edu/~appel/papers/appel-nj-audits.pdf	Based on analysis, Appel recommends (for NJ) a mandatory audit of 1% of precincts, with candidates permitted to select up to 7 additional precincts for audit. He assumes precincts are equal-sized. This plan for NJ seems sound, except that the non-random selections by candidates put his probability calculations, which are based on random selection, into question. His suggestion that candidates pay for audits is problematic.[18]
30. Dopp et Al., *One-Page Concept Proposal for Election Reform Legislation,* (March 2007) http://electionarchive.org/ucvInfo/US/EI-FedLegProposal-v2.pdf	Makes recommendations for federal election reform legislation including manual election audit amounts and procedures, auditable voting equipment, independent audit and recount committees, etc.
31. Dopp, *Fool Me Once: Checking Vote Count Integrity*, (March 3, 2007), available at: http://electionarchive.org/ucvAnalysis/US/paper-audits/TierElectionAuditEval.pdf	Shows how to use a graphical method for evaluating the effectiveness of election audits and uses it to evaluate three proposed election audits; explains why fixed percentage audits fail to ensure confidence; and responds to the Brennan Center analysis by showing how to do a more comprehensive analysis of any election auditing proposal and correcting some of its arithmetic.
32. P. Smith and B. Kibrick, *Manual Audit Requirements*, Verified Voting Foundation (March 17, 2007) http://www.verifiedvoting.org/downloads/StateManualAuditProvisions-03-07.pdf an updated version (October 2007) will be available at http://verifiedvoting.org/audits	Provides the text of legislative election audit requirements in Alaska, Arizona, California, Colorado, Connecticut, Florida, Hawaii, Illinois, Kentucky, Minnesota, Missouri, New Mexico, New York, North Carolina, Oregon, Pennsylvania, Washington, and West Virginia.[19]
33. Dopp, *Federal Election Audit Costs,* (March 21 2007) http://electionarchive.org/ucvAnalysis/US/paper-audits/FederalAuditCosts.pdf	Shows that confidence-level election audits are more effective and less administratively burdensome than tiered election audits. Estimates the cost and average effectiveness of three election audit proposals using federal election results from 2002 and 2004 US House and Senate elections.

[18] Note: For real-world audit sample sizes any decimal or fractional precincts must be rounded up to the next integer.
[19] Some of these States' requirements will be changing. Note: Pennsylvania election officials ignore the State's audit requirements in counties by using paperless direct-recording electronic (DRE) voting machines that are not auditable. According to Marybeth Kuznik of VotePA.us there has been no legal challenge to PA DREs yet on this statutory basis.

34. Stanislevic and Mark Lindeman, *H.R.811: Fact and Friction*, (March 4, 10, & 28) http://e-voter.blogspot.com/2007/03/hr811-fact-friction-part-i.html http://e-voter.blogspot.com/2007/03/hr811-fact-friction-part-ii_10.html http://e-voter.blogspot.com/2007/03/hr811-fact-friction-part-iii.html	(i.) Evaluates the original March version of House Resolution 811. (ii.) Evaluates the "Norden" letter (above) (iii.) Compares H.R. 811 audits to confidence-level audits,[20] failing to make data or analysis available to independently check their results.
35. Stanislevic, *Amend H.R.811 to Allow States to Use Audits to Confirm ALL Federal Election Outcomes* (April 7, 2007) http://e-voter.blogspot.com/2007/04/amend-hr811-to-allow-states-to-use.html	Recommends legislative language for calculating election audit sample sizes. However this language, if adopted, would allow as little as 10% confidence-level (ineffective) election auditing.[21]
36. *Collaborative Public Audit of the November 2006 General Election pursuant to the charge from the Cuyahoga County Board of Elections*, (April 18, 2007) by The Collaborative Audit Committee, coordinated by The Center for Election Integrity, Cleveland State University, with audit methodology and statistical analysis by The Northern Ohio Data and Information Service (NODIS) http://electionmathematics.org/em-audits/OH/2006Audit/cuyahoga_audit_report.pdf	Possibly the first actual confidence-level election audit conducted basing the sample size on candidate margins. This audit uncovers information about the inaccuracy and inauditability of Diebold voting machines, and reveals obstacles to election auditing, particularly in Ohio. Uses the Dopp-Stenger numerical method to exactly calculate sample sizes.[22]
37. Rivest, *On Auditing Elections When Precincts Have Different Sizes,* (April 29, 2007) http://people.csail.mit.edu/rivest/Rivest-OnAuditingElectionsWhenPrecinctsHaveDifferentSizes.pdf	Gives a new method of adjusting for precinct-size variation by using weighted rather than uniform random selection of audited units (larger precincts are more likely to be selected). However, the precise inputs needed for accurately using this somewhat complex selection method are unknown. It also provides a couple of other formulaic estimates that may be useful to election auditing.

[20] Stanislevic recommends randomly auditing smaller units such as individual DREs or optical scanners rather than entire precincts to increase the confidence level without increasing cost, but fails to mention that today's DRE voting systems are usually not auditable on the DRE level because the central tabulators are incapable of reporting individual DRE vote counts for each race and ballot issue. Stanislevic does not inform readers of a prior evaluation of the "Norden" letter by Dopp (3/3/07); and both Stanislevic and Lindeman fail to mention Dopp's prior (3/26/07) similar analysis of actual federal election results which had arrived at essentially the same conclusion, and which they had reviewed prior to publishing their own analysis.

[21] Stanislevic's recommendations for legislative language state that the audit sample sizes must be "at least as effective as section 322(a) in ensuring that for each federal election held in the state, a 100% manual recount would not alter the outcome of the election". The problem with this language is that the audits prescribed in 322(a) are only 10% effective in some cases, and give less than a 50% chance to protect even some US Senate races from fraud in small states.

[22] The Cuyahoga County auditors treat audit discrepancy results as though their sample is designed to predict the election results, rather than designed to detect at least one miscounted precinct in the case that sufficient miscount exists to alter the election outcomes; and the Cuyahoga auditors do not include the amount of missing and damaged paper ballot records in calculating the amount of discrepancies.

38. Dopp, *First US Scientific Election Audit Reveals Voting System Flaws But Questions Remain Unanswered – Critique of the Collaborative Public Audit of Cuyahoga County Ohio's November 2006 Election* (May 7, 2007) http://electionarchive.org/ucvAnalysis/OH/CuyahogaElectionAudit.pdf	Evaluates the Cuyahoga County Election Audit. (above)
39. *New Jersey Election Audit Legislative Proposal* (June 5, 2007) p. 19 – 25 http://www.njappleseed.net/entity_pdfs/175.pdf	A legislative proposal requiring 99% statistical power (confidence level) audits for federal and statewide elections, public random selections, audits of all ballot types, setting qualifications for an "independent" audit team. However the proposal nonsensically requires randomly selecting districts rather than precincts or other vote counts, and is unclear procedurally or has loopholes in vital areas.[23]
40. Dopp, *Comparison of Proposed Federal Election Audits* (June 12, 2007) http://electionarchive.org/ucvAnalysis/US/paper-audits/ComparisonFederalElectionAuditProposals.pdf	Estimates the effectiveness of three election auditing proposals using actual Federal election results from 2002 and 2004 U.S. House and Senate races.
41. *The American Statistical Association* urges the U.S. Congress to require confidence-level election audits (July 24, 2007) http://www.amstat.org/news/2007ASAElectionLettertoDFeinstein.pdf	The American Statistical Association recommends 99% confidence-level election auditing.
42. D. Jefferson, K.Alexander, E. Ginnold, A. Lehmkuhl, K. Midstokke, and P.B. Stark, California Secretary of State Bowen's *Post-Election Audit Standards Working Group Report - Evaluation of Audit Sampling Models And Options for Strengthening California's Manual Count,* (July 27, 2007) http://www.sos.ca.gov/elections/peas/final_peaswg_report.pdf	Recommends election auditing standards for California and brings attention to issues that need further investigation and explanation, but makes some mistaken conclusions such as "the statistical models that exist for determining the size of the random sample … require the random sample to be selected from the state as a whole … These models do not currently take into account overvotes and undervotes, nor … multi-winner elections."[24]

[23] NJ's chief election official and county officials conduct the audits under this proposal.

[24] *The Election Integrity Audit* (September 2006) by Dopp-Stenger had previously addressed some of these issues. Perhaps the California post-election auditing task force was relying on the auditing works of Saltman, Stanislevic, Lobdill, and Verified Voting which failed to address them.

43. McCarthy, Stanislevic, Lindeman, Ash, Addona, & Batcher, *Percentage-based vs. SAFE Vote Tabulation Auditing: A Graphic Comparison,* (July 28, 2007 revised and republished Nov. 2, 2007) http://www.verifiedvoting.org/downloads/SAFE-Auditing-Nov-2-Final4.pdf a copy of an earlier (not the first) published version is posted here: http://electionarchive.net/docs_pdf/info/US/SAFE-Auditing-July-26-Final.pdf	Justifies confidence-level election audits, repackages the method of using Saltman's calculation method as an input to the Aslam, Popa, & Rivest estimate. It liberally employs and touts graphical methods developed earlier by Dopp to evaluate election audits without citation[25]; employs some confusing misnomers, gives incorrect election auditing formulas & some incorrect recommendations, fails to note that its recommended methods do not account for overvotes or undervotes;[26] and fails to make data available to check its findings.
44. Dopp, *Response to Post-Election Audit Standards Working Group Report – Evaluation of Audit Sampling Models and Options for Strengthening California's Manual Count,* (August 3, 2007) http://electionarchive.org/ucvAnalysis/US/paper-audits/Response2CA-SOS-Audit-Report.pdf	Responds to some objections to confidence-level election auditing made by the CA post-election auditing task force report. This paper was sent in draft form to CA SOS audit group members, and happily, the CA Secretary of State Bowen required "confidence level" election audits, despite the prior objections.
45. California Secretary Of State Bowen, *Withdrawal Of Approval/Conditional Reapproval* Of Diebold Election Systems, Inc., Hart InterCivic, & Sequoia Voting Systems, (August 3, 2007) http://www.sos.ca.gov/elections/elections_vsr.htm	CA Secretary of State Debra Bowen requires confidence-level "post-election manual count auditing requirements, including: Increased manual count sample sizes for close races, … where the size of the initial random sample depends on … factors, including the apparent margin of victory, the number of precincts, the number of ballots cast in each precinct, and a desired confidence level that the winner of the election has been called correctly."

[25] The graphical method for evaluating election audit proposals that Verified Voting's paper employs and touts in its title was demonstrated earlier in Dopp in *Fool Me Once: Checking Vote Count Integrity*, (March 3, 2007). Verified Voting's authors John McCarthy and Howard Stanislevic failed to cite Dopp's work even though they were specifically asked to do so and their paper uses the exact same graphical method to illustrate the concepts of confidence-level auditing in five of its eight graphical illustrations. Another of Verified Voting's authors is Mark Lindeman who claims that the 2004 presidential exit poll discrepancies are inconsistent with vote fraud and was Farhad Manjoo's source for the Salon article claiming to disprove R.F. Kennedy's Rolling Stone article about the 2004 presidential election. See *2004 Presidential Election – Compendium of Attempts to Dismiss "Vote Fraud*
http://electionarchive.org/ucvAnalysis/US/IncorrectElectionDataAnalysis-06.pdf

[26] This Verified Voting paper appears to a casual reader to be presenting an all-new new election auditing method, but is describing confidence-level election auditing methods developed previously by Saltman and Aslam, Popa, & Rivest. Unlike Saltman, they fail to inform the reader that their methods assume that there are no undervotes or overvotes. Unlike Aslam, Popa, & Rivest the first publicly released version of this paper put parentheses in the wrong places in its formulas, giving impossible results, although they may have fixed and reposted their paper after being notified of the errors.

46. Lawrence Norden, Aaron Burstein, Joseph Lorenzo Hall, and Margaret Chen *Post-Election Audits: Restoring Trust in Elections,* (August 1, 2007) Brennan Center for Justice with the Samuelson Law, Technology & Public Policy Clinic http://www.brennancenter.org/dynamic/subpages/download_file_50227.pdf Executive Summary http://www.brennancenter.org/dynamic/subpages/download_file_50228.pdf http://www.brennancenter.org/dynamic/subpages/download_file_50109.pdf	This paper is a mix of well-thought-out, correct, incorrect, and ambiguous recommendations. Some of the Brennan Center's recommendations are misleading to a lay reader. [27] Tries to give a comprehensive overview of election auditing methods, but recommends methods for calculating insufficient audit sample sizes in cases of miscounted overvotes and undervotes.
47. Javed Aslam, Raluca Popa and Ronald Rivest, *On Estimating The Size and Confidence of a Statistical Audit,* written (June 30, 2007) presented (8/6/2007) at EVT07 conference http://people.csail.mit.edu/rivest/publications.html http://people.csail.mit.edu/rivest/AslamPopaRivest-OnEstimatingTheSizeAndConfidenceOfAStatisticalAudit.pdf	Provides a highly accurate formula to estimate the confidence-level election auditing sample sizes that are found precisely by using the Dopp-Stenger numerical method, over-estimating the audit sample sizes necessary to achieve a desired confidence by at most one auditable vote count unit.[28]
48. Stanislevic, *Audit Sample Size Calculator Tool* http://mysite.verizon.net/evoter/AuditCalcNY.htm	Internet tool, using the Aslam, Popa, & Rivest formula for calculating election audit sample sizes; does not adjust for auditable unit (precinct) size-variation.[29]

[27] For example, on page 10 they recommend machine-level audits as being more efficient (less effort) than precinct audits (without informing the reader that most digital recording electronic (DRE) voting systems today are not designed to be machine-level auditable). Then on page 25, they contradict the statement on page 10 by incorrectly claiming that auditing on the ballot-level requires selecting more records than auditing on the precinct-level. They also incorrectly credit Stanislevic, rather than crediting earlier authors such as Neff, Wand, and Johnson who first brought attention to the fact that smaller-sized more numerous auditable units are more efficient (less effort) to audit than fewer larger-sized units. They incorrectly claim that "If precinct size varies within a Congressional District, it will be more difficult to obtain a high level of confidence in the results, unless precinct selection is weighted by the number of voters in each precinct." Methods to weight precinct selection can be complex and difficult to determine probabilities. Precinct-size variation can be adjusted for more easily by using the auditable report to calculate the precise minimum number of miscounted largest precincts which could alter the outcome and using it as input to determine the sample size and an exact method exists for calculating the number of miscounted precinct or machine vote counts which could wrongly alter election outcomes. There is a fundamental difference between "amount" and "percentage". Specifying audit percentages (of the number of total auditable vote counts) does not provide a desired confidence level because, for the same percentage, the amount of an audit varies widely depending on the total number of auditable vote counts in each race which varies for each state, jurisdiction, and race.

[28] The closeness of the Aslam, Popa, & Rivest estimate was determined by empirical trials never over-estimated the exact election audit sample size by more than one auditable vote count unit. This estimate provides an easy way to calculate confidence-level election audit sample sizes without the need for either a computer program or trial and testing The exact minimum number of miscounted auditable vote counts that could alter an election outcome, as given precisely in the Dopp-Stenger *Election Integrity Audit* paper, is a necessary input to the Aslam, Popa, & Rivest formula.

[29] The program asks for one too many inputs from the user which could conflict with at least one of the other inputs if all of the inputs are not correctly calculated by the user to be consistent with the other inputs. However, possibly the program checks the consistency of the inputs for the user.

49. Dopp, *Amendment Suggestions for the Senate Ballot Integrity Act of 2007 (S1487) To Increase Public Verifiability of Election Outcome Accuracy & Eliminate Unfunded Mandates,* (September 2007) http://electionmathematics.org/em-legislation/S1487Amendments.pdf	Suggests legislative language (pages #3 – 11) for election auditing requirements that are sufficient and efficient, and publicly verify election outcomes by taking H.R.811's audit requirements and trying to improve on them and combine them with S.1487's requirements.
50. Philip B. Stark, *Conservative Statistical Post-Election Audits,* (October 16, 2007 – revised November 24, 2007) http://statistics.berkeley.edu/~stark/Preprints/conservativeElectionAudits07.pdf	Presents a new method for analyzing election audit discrepancies that depends on the margin between candidates, the sample size, and the maximum value of the sum of the absolute value of the difference between audit counts and initial counts for all candidates out of sampled precincts. Recommends beginning with an audit sample and increasing the audit amount until the election can be certified with confidence. Stark is still working on revisions to his method and this paper.[30]
51. Dopp, *Preliminary Evaluation of Philip Stark's Conservative Statistical Post-Election Audits* (October 27, 2007) http://electionarchive.org/ucvAnalysis/US/paper-audits/Corrections2PhillipStark.pdf	Brief preliminary review of Stark's paper above.
52. Joseph A. Calandrino, J. Alex Halderman, and Edward W. Felten, *Machine-Assisted Election Auditing,* Center for Information Technology Policy and Dept. of Computer Science, Princeton University Woodrow Wilson School of Public and International Affairs (August 2007) http://itpolicy.princeton.edu/voting/audit07full.pdf	Preliminary review based on authors' abstract: Because today's voting machines are often auditable only at the precinct level and it is more efficient to audit at the ballot level, this paper presents a new method for auditing precincts using auditing machines which output ballot-level auditable reports. If ballots are run through the auditing machines, this strategy would achieve equal confidence as precinct based auditing at a lower cost – by auditing fewer ballots - while protecting voter privacy. (This review will be modified after I have studied the paper.)

[30] Stark's paper addresses the question "Given the observed audit discrepancies, should we certify at 99% confidence?" i.e. he asks, "Given the discrepancies in the audit, what is the probability that the election outcome is correct?" Stark somewhat misrepresents prior work by saying "Previous papers ask the question '… what is the smallest … sample … [which] has chance of 1-α of containing at least one miscounted *ballot*.'" and "These approaches do not say whether to trust the reported election results." Stark's example on pp. 24-28 uses the same fundamental probability calculation as prior authors to calculate election sample sizes. Stark uses a trial and test method to obtain the audit sample size, employing a calculated bound for the amount of total miscount possible in each precinct. Stark's choice of error bound is much higher than the 15% or 20% bound used by prior authors, making Stark's audit sample size more conservative (larger) for a given race with a specific margin and total number of vote counts than that of some authors. A smaller upper bound using the amount of error that could alter an election outcome is suggested in Appendix B of this paper, the number of ballots minus the number of votes for the overall losing candidate with the most votes in any precinct. Stark's example of a close race having only nine total precincts other authors' methods also require a 100% sample size. Stark solves the problem of how to handle overvotes and undervotes when auditing multi-winner races, and his paper provides an additional method for evaluating election audit discrepancies to decide whether to expand an audit or to certify the election.

In 2007, The American Statistical Association joins Verified Voting, the National Election Data Archive, the Carter-Baker Commission on Election Reform, and The U.S. Government Accountability Office, The Brennan Justice Center, The National Science Foundation's Project ACCURATE, and The League of Women Voters U.S. in recommending routine election auditing. The California Secretary of State Debra Bowen requires state-wide 'confidence-level' election auditing and 100% manual audits of most direct-recording electronic (DRE) voting machine counts.

2008 – Probability Proportional to Error Bound Methods for Election Auditing

53. *On Auditing Elections When Precincts Have Different Sizes* Javed A. Aslam, College of Computer and Information Science, Northeastern University, Raluca A. Popa and Ronald L. Rivest, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology (Draft of January 17, 2008) http://people.csail.mit.edu/rivest/AslamPopaRivest-OnAuditingElectionsWhenPrecinctsHaveDifferentSizes.pdf	Provides a method for selecting vote count audit samples that is approximately 30% more efficient than prior methods by making weighted random selections of auditable vote counts (precincts or other reported vote counts) using the upper bound for the possible margin error within each auditable vote count. This selection method increases the effectiveness of any vote count audit sample size by tending to select vote counts contributing potentially larger error.
54. *Percentage-Based versus Statistical-Power-Based Vote Tabulation Audits,* John McCarthy, Howard Stanislevic, Mark Lindeman, Arlene S. Ash, Vittorio Addona, And Mary Batcher, The American Statistician, (February 2008), Vol. 62, No. 1 15 http://verifiedvoting.org/downloads/TAS_paper.pdf	This paper does a good job of selling the concept of confidence-level audits but fails to provide complete, correct descriptions of election auditing methods developed by other authors. This paper adds one original method for estimating the effect of precinct-size variation by using the 2004 Ohio's congressional district #5 election results that is unnecessary to use during any audit because the actual vote counts must be made publicly available prior to calculating audit sample sizes. Their original method is often inaccurate because not every election contest has the same precinct-size variation pattern as Ohio's congressional district 5.

55. *Election audits by sampling with probability proportional to an error bound: dealing with discrepancies* Philip B. Stark, Department of Statistics, University of California, Berkeley, CA 94720-3860, (DRAFT 20 February 2008) http://statistics.berkeley.edu/~stark/Preprints/ppebwrwd08.pdf	Stark's paper discusses sampling vote counts using probability proportional to a bound on the possible margin error in each vote count (PPMEB) to answer "How large a PPMEB sample should be drawn to have confidence that the election outcome is correct if the sample includes only "small" errors?" and "What is the confidence that the outcome of the election is correct, given the discrepancies a PPMEB audit uncovers?" and "How should one increase the PPMEB sample size if discrepancies are found, and when can one stop auditing?" Stark's paper also discusses how to use PPMEB with stratification.
56. *Critique of the Verified Voting McCarthy/NJ Election Audits*, Dopp (February 21, 2008) http://electionarchive.org/ucvAnalysis/US/paper-audits/Critique-McCarthy-NJ-Audits.pdf	Discusses some specific mathematical errors in the most recent recommendations published by the American Statistical Association which were made by Verified Voting/McCarthy and incorporated into N.J. election auditing statute.
57. *Mandatory Post-Election Vote Count Audits-How to Determine Initial Sample Size for 95% PPMEB Confidence-Level Vote Count Audits* Kathy Dopp (December 23, 2007 updated 1/17/2008, 2/25/2008) http://electionarchive.org/ucvAnalysis/US/paper-audits/UT/VoteCountAudits.pdf and *Derivation of the formula for the number of selection rounds for the Probability Proportional to Margin Error Bound (PPMEB) method for determining samples for Vote Count Audits* Kathy Dopp, (4/28/2008) http://electionarchive.org/ucvAnalysis/US/paper-audits/PPMEB-Auditing.pdf	These documents provide a simpler explanation of the concepts and methods for calculating probability proportional to error bound (PPMEB) vote count audit sample sizes derived originally by Aslam, Popa, & Rivest. The first document also resolves the debate over what upper bound to use for the available margin error within each auditable vote count which could contribute to altering the outcome –thus providing more accurate audit sample sizes. (See Appendices B and G in this paper re. prior debate.) The second document provides a derivation for the PPMEB method that is easier to understand.
58. *A Sharper Discrepancy Measure for Post-Election Audits* Philip B. Stark Department of Statistics, Berkeley, CA 94720-3860 To appear in The Annals of Applied Statistics http://statistics.berkeley.edu/~stark/Preprints/pairwise08.pdf	Provides a method to analyze the discrepancy between machine counts and a hand tally of votes in a random sample of precincts to infer whether error affected the electoral outcome, and to calculate a P-value for the hypothesis that a full hand tally would find a different winner. Stark's approach uses the same fundamental approach taken by other researchers to calculate correct election audit sample sizes (to size the sample to detect outcome-altering miscount), but in my opinion, in the context of audit discrepancy analysis, even if correctly calculated, Stark's method only provides a lower bound for the probability which some times tells us when "not" to certify an outcome, but can not tell us when "to" certify an election outcome.

59. Mandatory Vote Count Audit - A Legislative & Administrative Proposal, Kathy Dopp, (February 2006 Revised 9/2006, 12/2007 to 7/ 2008) http://electionarchive.org/ucvAnalysis/US/paper-audits/legislative/VoteCountAuditBillRequest.pdf	A legislative post-election auditing proposal which could be applied to any State with minor modifications to meet unique State statutes. This proposal has been updated to include the latest mathematical methods developed for calculating audit sample sizes and making random selections and will be edited soon to describe how to analyze the audit discrepancies.
60. Realities Mar Instant Runoff Voting, 17 Flaws and 3 Benefits, Kathy Dopp (June 10, 2008, updated June 25, 2008 and August 03, 2008) http://electionarchive.org/ucvAnalysis/US/RCV-IRV/InstantRunoffVotingFlaws.pdf	Includes a section on pp. 7-8 on how to conduct post-election audits for election contests which were counted using the instant runoff voting (IRV) method.
61. CAST: Canvass Audit by Sampling and Testing, Philip B. Stark, Department of Statistics, University of California, Berkeley, CA DRAFT 16 September 2008 http://statistics.berkeley.edu/~stark/Preprints/cast08.pdf	Uses a similar method of adding up audit unit upper margin error bounds to calculate the minimum number of corrupt audit units necessary to alter an election outcome as recommended earlier by Dopp but uses a formula for random selection with replacement size rather than without replacement to determine the sample size; allows an arbitrary "tolerable level" of error; normalizes the error bounds to percentages; and uses the maximum value of error bounds over all candidates in his sums to calculate minimum number of corrupt audit units that could alter the outcome. Stark states that conclusions can be drawn from the value of his test statistic for the audit discrepancy results that do not seem adequately justified to this author.

Agreement on Election Auditing Fundamentals

Virtually all experts who support confidence-level election auditing agree that:

- Election audits should achieve a high confidence, such as 95% to 99% probability, that the voters' intended choices have determined the election outcomes.
- A report of all the auditable unofficial vote counts that might be selected for auditing must be publicly released prior to beginning the audit because otherwise unaudited vote counts could be manipulated after the audit to match erroneous totals.
- All ballot types (mail-in, early, provisional, Election Day, etc.) should be manually audited.
- For election audits to be publicly transparent and verifiable
  - election records necessary to verify the integrity of the audit are made available to auditors and to the public,
  - the manual audits and random selections are publicly observable,
  - ballots and other election records are secured using procedures that are open to public scrutiny and verification.

- The audit sample must be randomly selected using a scientific definition. For example, "Random selection or random drawing" means that items are selected by using probability sampling methods including simple random samples or stratified random samples.
- Election audits can also detect voting machine malfunction and voter disenfranchisement[31].
- Election audit sample sizes should be based on
    - the margin between the least-scoring winning candidate and the best-scoring losing candidate,[32]
    - the total number of auditable vote counts (e.g. precinct, batch and machine counts),
    - an upper bound for the margin error in each auditable vote count and possibly a maximum assumed rate of availble margin error that a fraudster might target, and
    - if the maximum rate of margin error is assumed to be less than 100% of the actual upper error bound, then candidates should be able to select some additional discretionary vote counts for the audit sample, or alternatively, a calculation should be done to determine if any vote counts appear suspicious and also need to be audited.
- Smaller vote counts, even down to the ballot level, are more efficient to audit. However most current voting systems are not auditable on the individual ballot or even on the DRE machine level because the central tabulators are incapable of printing an auditable report of the individual ballot or DRE vote counts because they do not retain sufficient detailed data in the central tabulators.
- A minimum of 1% of auditable vote counts should be randomly selected for audit in order to detect any distributed machine error, and if an auditable vote count is not randomly selected in each jurisdiction having separate election administration, then one additional auditable vote count should be randomly selected for audit from any "missed" jurisdiction for each race being audited to ensure the detection of ballot programming and other errors.

Order of Steps for Conducting Election Audits

1. Publicly release a report of unofficial vote counts including machine serial numbers, the number of ballots cast, the type and date of the count (early, Election Day, mail-in, provisional, etc.), and the counts "for" or "against" any issue, judge, or candidate within each auditable vote count.[33]
2. For confidence-level election audits, use the report from step one to

    Uniform Method

    a. determine the minimum number of auditable vote counts which could alter the election outcome[34], and

[31] Independent audits detect probable voter disenfranchisement by doing ballot reconciliation to determine if the number of total ballots recorded is equal to the number of voters recorded; and determining the number and rates of under and over-votes, provisional ballot use, and provisional ballots counted, voting equipment errors, voter errors, etc.
[32] Using these margins works for multi-candidate and multi-winner races. Margins are calculated as shown in Appendix B.
[33] The sum of the number of ballots reported in each auditable vote count must equal the total number of ballots cast in each race or issue; and the sum of the vote counts listed for each auditable vote count must equal the totals unofficial vote count.
[34] A calculation method is supplied in appendix B in this paper, but there is some debate over what method to use, as described in appendix G in this paper.

b. use the amount calculated in step 2a to calculate the election audit sample size using either trial and error, the Aslam-Popa-Rivest formulaic estimate, or the exact Dopp-Stenger numerical algorithm.[35]

Probability Proportional to Margin Error Bound Method (approximately 30% more efficient)

a. determine the selection probabilities for each vote count and the total possible margin error

b. determine the number of selection rounds[36]

The PPMEB method was described first in the Aslam, Popa, Rivest paper, and also explained in the Stark and Dopp papers.

3. Publicly select the random sample of auditable vote counts.[37]
4. Publicly conduct the manual counts of voter-verified paper ballots for each auditable vote count in the sample.
5. Compare the manual counts with the report of unofficial counts in step one and record and investigate the discrepancies. Make all possible corrections to the unofficial results.
6. Use the number of discrepancies found between manual and unofficial machine counts to calculate the confidence-level for the election outcome. If the calculated confidence is less than the target confidence, expand the sample size for the manual audit, possibly to a full hand count.[38] Certify the election outcome when the target confidence-level is obtained or call for another election for any races left in doubt (this could become necessary if sufficient voter-verified paper ballots are missing or damaged).
7. Publicly release the official election results and the results of the manual audits, including the amount and type of discrepancies and any actions taken to resolve or explain any discrepancies between the initial unofficial counts and the manual audits.

[35] Confidence-level election audit sample sizes may be calculated in several ways: 1. use Dopp's method for calculating margins and the number of miscounted vote counts which could reverse an election outcome and then use the numerical method given by Dopp-Stenger for calculating the sample size; or 2. use Dopp's method for calculating margins and the number of miscounted vote counts which could reverse an election outcome and then use the formulaic estimate given by Aslam, Popa, Rivest for calculating the sample size. In trials, the Aslam-Rivest-Popa formula over-estimates the audit sample size at most by one auditable vote count. In addition Stark suggests using an upper bound equal to the total possible error in any precinct to calculate the minimum number of miscounted precincts which could reverse an election outcome. As shown in Appendix B in this paper Dopp suggests using a smaller upper error bound that is equal to the most possible error that could reverse an election outcome, i.e. the number of ballots minus the number of votes for the highest scoring runner-up in a race in each precinct, in conjunction with a maximum level of undetectability. Such debates over calculation details affect how large the election audit sample sizes are. County-level audits and higher error bounds result in overall larger sample sizes than state-level audits and smaller error bounds do.

[36] A complete derivation for the formula for the number of selection rounds is posted at: http://electionarchive.org/ucvAnalysis/US/paper-audits/PPMEB-Auditing.pdf

[37] Research is ongoing in the best methods for random selections. Aslam, Rivest, and Popa - and Cordero, Wagner, and Dill suggest methods using ten-sided transparent dice.

[38] Appendix E in this paper and Stark's paper specify methods of deciding whether to expand the manual audits or certify.

What Remains to Be Done to Implement "Confidence-Level" Election Audits?

- Complete some method(s) for calculating efficient confidence-level initial audit sample sizes.[39]
- Describe, for the lay person, possible random sampling methods for selecting auditable vote counts.[40]
- Specify all discrepancy conditions that put an election outcome into question and require expanding the audits; and provide efficient strategies for expanding audits, including methods which reduce the unnecessary expansion of audits.[41]
- Publish an illustrated manual & a toolkit for the layperson with clear explanations and step-by-step instructions for conducting election auditing, including training materials[42] and descriptions of options for auditing unofficial election outcomes and securing ballots and election records for various voting and election systems. Include clear instructions for calculating the confidence-level obtained from the audit discrepancy results, regardless of the initial audit sample size. [43]
- Pass Legislation: Many States need to increase the time period between Election Day and the date when election results are certified to allow adequate time for post-election audits[44] and may need to pass legislation to fund independent confidence-level manual election audits.
- Organize conferences and conduct workshops and training sessions on election auditing.

CONCLUSION

In order to assure that voters accurately decide who governs them, election reform advocates, computer scientists, mathematicians and statisticians, election officials, voting machine vendors, and legislators must implement publicly-verifiable confidence-level election audits.

[39] Calculation methods could be developed for detecting "suspicious" auditable vote counts that should be audited in addition to the randomly selected counts. This is necessary in order for the audit sample size to be logically consistent with the use of an assumed "maximum level of undetectability" by miscount (MLU) per vote count. (In other words it is necessary to ensure that our assumptions for the calculations are correct.) Auditable vote counts with suspicious-looking vote counts need to be audited in addition to the random sample calculated by currently agreed-upon confidence-level methods. "Suspicious" can be defined by a calculation based on: The MLU, unofficial vote counts, and another reliable measure of voter preference such as partisanship where available. As an alternative, candidates should be permitted to select some discretionary auditable vote counts for the audit.

[40] One of the best random selection methods may be the translucent ten-sided die described by Cordero, Wagner, and Dill.

[41] Dopp-Stenger recommend designing audit samples to verify the outcomes in each county to reduce the likelihood of having to expand audits in all counties, when miscounts are detected in one county.

[42] Including spreadsheet & programming tools and descriptions of accurate efficient methods to hand count paper ballots.

[43] Most states conduct flat-rate or tiered percentage audits rather than confidence-level audits. This is an inefficient approach when trying to obtain a target confidence-level in the election outcomes. i.e. Flat rate initial audit sample sizes will often be too small to achieve a desired confidence and require numerous auditing rounds. Flat rate audits, such as flat-rate x% audits and tiered, or adjustable percentage, audits such as the 3%, 5%, 10% audits that depend on unofficial margins will achieve low initial confidence-levels when races have a small number of total vote counts and/or very close margins. Adjustable percentage audit sample sizes most often require auditing too small or too large sample sizes than is necessary for 99% confidence in outcomes.

[44] Some states allow 30 days between Election Day and the date when results must be certified, but some other states, like Florida, have very short canvassing periods of two weeks or less.

There has been confusion, even among experts, over how to correctly implement election audits. Election audit sample sizes should be sufficient to determine, for instance, if close election contests with high amounts of under-votes are correctly decided.[45]

Due to such confusion among experts it is clear that we need

- a correct, understandable election auditing manual and easy-to-use tool-kit (preferably before the 2008 election) that completes and corrects the derivation of "confidence-level" election auditing methods, including a clear explanation of how to interpret discrepancy amounts to determine if outcomes are in doubt; and that provides pictures, forms and toolkits; and explains election auditing in simple-to-follow terms for the lay person. In order to be successful, so that election jurisdictions do not have to hire statisticians to plan every post-election confidence-level audit, such a project will require professional manual creators, and program designers, to create an easy-to-use manual and tool-kit in collaboration with election officials, security experts, and election auditing experts.[46]
- conferences which bring together actual experts who have developed correct election auditing methods together with State election officials and State and Federal legislators.

Funding is needed for collaborative efforts which work with the experts who have been originating correct new election auditing procedures in order to develop and disseminate the knowledge, resources and skills necessary to implement routine election audits.

About the Author and Reviewers

Kathy Dopp, M.S. Mathematics, Executive Director, National Election Data Archive

Bios: Dopp founded the National Election Data Archive; derived a mathematical function to allow exit poll discrepancies to be evaluated for consistency with vote miscount or exit poll response bias; and mathematically proved that attempts to dismiss the exit poll evidence of vote miscount in the 2004 presidential election were incorrect.[47] Dopp's analyses of the 2004 presidential election results, with NEDA's statistician volunteers, were cited by Representative John Conyers, Mark Crispin Miller, and Robert F. Kennedy. The Dopp-Stenger "confidence level" election auditing calculation method was used in the Cuyahoga County Ohio Collaborative Election Audit. kathy@electionarchive.org

Frank Stenger, Ph.D. Mathematics, University of Utah School of Computing

Bios: Stenger has written over 150 papers and several text books, is an expert in computations, developer of a new method of computation called Sinc methods, is one of 13 people selected by the Society of Industrial and Applied Math (SIAM) who has contributed most to numerical analysis in the 20th century, winner of the University of Utah distinguished research award, and is editor of six scientific journals. stenger@cs.utah.edu

Ron Baiman, Ph.D. Economist, Loyola University in Chicago

[45] In the 2006 Sarasota County, Florida Congressional District 13 race, there were 18,000 missing votes (undervotes) in a Democratic-leaning county recorded on paperless ES&S DREs in a tight election. Statistics show that these undervotes probably altered the outcome and caused Christine Jennings to lose the U.S. House race.

[46] This project could take roughly six months and cost roughly from $50,000 to $100,000 to accomplish.

[47] See "*Mathematical Proof that Election Sciences Institute's Test to Rule Out Vote Fraud Is Logically Incorrect*" http://electionarchive.org/ucvAnalysis/US/exit-polls/ESI/ESI-hypothesis-illogical.pdf and "*Vote Miscounts or Exit Poll Error? New Mathematical Function for Analyzing Exit Poll Discrepancy*" http://electionarchive.org/ucvAnalysis/US/Exit-Poll-Analysis.pdf and "*2004 Presidential Election – Compendium of Attempts to Dismiss 'Vote Fraud'*" http://electionarchive.org/ucvAnalysis/US/IncorrectElectionDataAnalysis-06.pdf `

Bios: Baiman (Vice President of National Election Data Archive) - Ph.D. and M.A. in Economics, B.S. in Mathematics and Physics – Baiman is a Policy Research Project Development Analyst at Loyola University in Chicago, and a visiting assistant professor teaching global economics at the University of Chicago. He is a recipient of a Russell Sage Foundation Grant and the Choice Award for Best Academic Title. baiman@sbcglobal.net

Acknowledgments

Thanks to Paul Dopp and Wayne Tripp for advising me to write this and for reviewing it. Thanks to The National Election Data Archive's Board of Trustees and to University of Utah's Frank Stenger and to Ron Baiman of Loyola University Chicago for reviewing this. Thanks to all the authors whose work is reviewed for their work and for bringing attention to the subject of election auditing, and thanks to authors who made suggestions on my reviews of their work. Thanks to Andy Bardwell of Bardwell Consulting and to University of Berkeley's Arthur Keller for describing ballot-auditable voting systems. Thanks to David Dill of Verified Voting for instructive criticisms and thanks to both Dill and Pam Smith of Verified Voting for suggesting additional election auditing works to add to the chronology; and to University of Berkeley's Philip Stark for prompting the addition of appendices to prove that some statements are true and for helpful reviews of some of my work. Thanks to Massachusetts Institute of Technology's Ron Rivest for contributing to Appendix C. Thanks to University of Berkeley's Joseph Lorenzo Hall for his web page election auditing bibliography.[48] Thanks to Javed Aslam, Raluca Popa, Ron Rivest and Philip Stark for their work on the Probability Proportional to Margin Error Bounds methods. Thanks to Andy Bardwell for encouraging me to keep working.

This document is found online at
http://electionarchive.org/ucvAnalysis/US/paper-audits/History-of-Election-Auditing-Development.pdf

[48] Joseph Lorenzo Hall *Post-Election Manual Auditing of Paper Records* Bibliographyhttp://www.josephhall.org/nqb2/index.php/2006/11/26/pea_biblio

Appendix A: Incorrect Auditing Recommendations That Would Not Detect Faulty Outcomes

The errors listed below are followed by the numbered works in the chronology above which make the error.

- Inputs to calculations using incorrect inputs (for measuring both margins and vote count size) that would result in insufficient audits to detect situations when miscounted undervote or overvote errors result in incorrect election outcomes. This error creates a strategy for fraudsters to use who want to avoid detection. (19 p. 15), (21 pp. 1, 4), (23), (24 pp. 2, 5), (34), (36 p. 45 #6), (43 p. 27), (46 p. 5, 6, 21, 22, & Appendix B)
- Incorrect formulas – (43 Appendix B gave two incorrect formulas in the original version of the paper by putting parentheses in the wrong places giving impossible results)[49]
- Legislative language with loopholes that would allow for ineffective, insufficient election audits (25) & (43 p. 27 suggests requiring a specific formulaic estimate rather than allowing States to use precisely correct calculation methods, and specifies language which would result in insufficient sample sizes not accounting for undervotes or overvotes)[50]
- Faulty procedures that could miss detecting erroneous election outcomes by allowing the audits to be easily manipulated (19 p. 15), & (21), (23), (24), (34) by virtue of faulty calculations – see above, (36 - ignores missing and damaged ballot records in its discrepancy calculations p.40#4 & p.45#9), (39 pp. 2-3 - provides several ways to manipulate audits including does not clearly specify that auditable vote counts shall be publicly reported prior to the audit, disregards missing or damaged paper ballot records in calculating discrepancies, allows races with fewer than 100 precincts to be insufficiently audited, etc.), (43), (46 pp. 6, 7, 25, 32, 33 – some corrected on pp. 33, 39)
- Rounding incorrectly for real-life elections (a precise algorithm must include rounding rules or will often give slightly incorrect results-(19), (21), (24), (23), (29 p. 7 table 2), (43 p. 15)
- Conducting audits without confirming the accuracy of election results. Audit procedures currently used in some states do not confirm the accuracy of unofficial machine counts.[51]

---

[49] The formulas were later corrected after one of Verified Voting's A.S.A. authors was alerted to the problem.

[50] Recommendations for legislative language for election auditing, based on improvements to Holt's H.R.811 requirements are provided in *Amendment Suggestions for the Senate Ballot Integrity Act of 2007* http://electionmathematics.org/em-legislation/S1487Amendments.pdf

[51] For some examples: Utah compares DRE voter-verifiable paper ballot records with another record printed by the same DRE memory card and never compares voter verifiable ballot records with the unofficial vote counts on the central tabulators. Many Utah counties leave opaque plastic covers over DRE paper roll ballot records, and Utah's instructions do not tell voters to verify their paper ballot roll records, so most voters are unaware of the need to verify paper ballot records. According to current interpretation of Utah statutes, an auditable report of vote counts on each DRE machine is prohibited from being publicly released and polling place totals are *not* publicly posted on election night, making it impossible to verify the manual audits of DREs; and records needed to reconcile the number of voters with number of ballots and with voter registration rolls are withheld. See http://utahcountvotes.org. Ohio statute prohibits independent auditors from handling ballot records during the canvass period. Some Nevada counties compare post-election memory card printouts with unofficial results, never examining the voter-verifiable paper roll ballot records. John Gideon of Voters Unite and Kathy Dopp independently called some of Nevada's county election officials and confirmed that some Nevada counties are not manually counting any voter verifiable paper records but were merely reprinting machine counts and comparing those with machine counts. Pennsylvania has a law requiring audits which are impossible to perform on its paperless digital recording electronic devices. Other states, including Florida, conduct audits after the election results are certified. Hawaii does not publicly post polling place totals and sends all electronic ballot records to a central state office.

- Testing ballots or testing machines rather than checking vote counts neglects the fact that fraudsters can easily rig programs to count correctly at any time except during an actual election, or counts can be manipulated post-election on the central tabulator, and that current voting systems are not designed to be auditable on the individual ballot level. (3 p. 1-12)
- Interpreting audit discrepancy results incorrectly or misunderstanding what the maximum level undetectability means (36 p. 15 & 45), (43 p.22), (46 p.23)

Some authors use misnomers which add to the confusion.

- Work #43 p. 14 renames what Saltman in 1975 properly calls the "maximum level of undetectability by miscount" (MLU) and the Brennan Center and Dopp-Stenger in 2006 call "maximum vote shift" by incorrectly renaming MLU "Within Precinct Miscount" (WPM). WPM is a misnomer because MLU applies to all auditable vote counts, not just precincts. Further WPM is reminiscent of "Within Precinct Discrepancy" (WPD), the signed difference between candidate margins in exit polls versus vote count margins – a concept that is unrelated to MLU. This paper also potentially confuses a casual reader by renaming "confidence-level" audits to "SAFE" audits; copiously using "we" rather than mentioning the persons who did the original work; and failing to cite prior work.
- Work #46 p. 9 uses the term "adjustable percentage" audits which is a misnomer for confidence-level audits. This paper says that "The mathematics behind this [adjustable percentage] method involves calculating a sample size given the margin in the closest race on the ballot and a desired confidence level for detecting error or fraud." However, it contradicts itself by claiming that the tiered approach as proposed "under the Holt proposal" is a variation of the adjustable-percentage audit model. But in truth such models [as the Holt 3%, 5%, and 10% audits] are not designed to achieve any particular confidence level[52] and in virtually all cases, result in insufficient audit amounts or, more usually, in unnecessarily large audit amounts. It is the *amount* of the audit, *not the percentage* of the total number of auditable vote counts that determines the probability (confidence) for an election audit to detect outcome-altering miscount. To understand why, consider a 10% audit in an 800-precinct U.S. House district which gives what may be a sufficient audit of 80 precincts, whereas a 10% audit in a 150-precinct U.S. House district gives an audit of 15 precincts which would *not* be sufficient in any close race where fewer than 1 in 15 miscounted precincts could alter the outcome.

[52] Although in practice Holt's audit performs pretty well, it requires auditing more vote counts overall for a lower overall average effectiveness than confidence-level audits would. For an analysis of the effectiveness of the Holt audits based on actual prior election data see Dopp, *Federal Election Audit Costs* (March 21 2007) http://electionarchive.org/ucvAnalysis/US/paper-audits/FederalAuditCosts.pdf

Appendix B: How Many Miscounted Vote Counts Could Alter An Election Outcome?

Roy Saltman's Calculation in 1975

In Saltman's own words in 1975 [slightly altered for readability only] he assumes that there are just two opposing candidates and no overvotes or undervotes and that:

p is the total number of precincts, and
x is the maximum level of undetectability by observation is given as a fraction, and
d is the difference in the candidates' votes plus one (in a two-candidate race) is also given,
n is the total number of votes cast

Then
f, the minimum number of voteswitched precincts that wi11 overturn the contest is easily computed.

First, d/2 (plus 1/2 if d is odd) is the minimum number of votes that must be switched in order to reverse the election,. Then, n/p is the number of votes per precinct, (assuming an equal number of votes in each precinct) and nx/p is the maximum number of votes in each precinct that can be switched without detection by observation. Then the minimum number of vote-switched precincts required to reverse the election is: $f = \frac{p}{x} \bullet \frac{d}{2n}$

where d/n is the fractional difference between the candidates and d/2n is the minimum fractional difference between the candidates that needs to be switched in order to reverse the election. As f must be an integer number of precincts, if it is not, the next highest integer is selected.

Analysis of Saltman's Calculations

Saltman's calculations do not correctly handle situations with miscounted undervotes and overvotes because of his use of n, the total number of votes, rather than using the total number of *ballots,* cast. Unfortunately Verified Voting, and the Brennan Center ascribe to Saltman's method, but fail to note, as Saltman correctly does, that that the method assumes that there are no incorrectly recorded overvotes or undervotes.

Saltman's assumption of only two candidates in a race is unnecessary. Even if the race has multiple winners and multiple candidates, the margin to use is simple to calculate. The margin used for calculating the audit sample size is always the margin between the lowest performing winning candidate and the highest performing losing candidate. This margin is the correct one for the calculations because to alter any election outcome in the race, the margin has to be at least this large.

Kathy Dopp's Calculations in 2006 – Using Ballots

If
N = the total number of precincts (or other auditable machine vote counts)
v = the maximum percentage of corruption expected on any one voting machine (undetectability by observation) as a percentage of ballots
m = the margin between the two leading candidates (as a percentage of ballots cast)
then X = the number of corrupt counts that could alter an election outcome is calculated as

$$X = \frac{Nm}{2v}$$

Dopp's initial recommendation in July 2006 was silent on the question of how to calculate margin percentages. However, Dopp and Stenger recommended, in their September 2006 paper, that the number of ballots cast, not the number of votes counted, be used as a base for election auditing calculations.[53] This makes sense because every ballot cast could be counted incorrectly in any race. Every voter who casts a ballot wants that ballot accurately counted, not just voters whose ballots recorded votes in the initial machine tally (See appendix B). Dopp's initial recommendation in July 2006, that the smallest margin between any candidates be used was overly conservative and could result in unnecessarily large sample size in multiple-candidate races.

The margin used for election auditing calculations in any race should be the margin between the lowest performing winning candidate and the highest performing losing candidate, as a percentage of ballots cast in the race. This method works even when there are multiple candidates and winners in a race. Dopp and Stenger conservatively recommend conducting county-level audits and using the minimum calculated margin between the State and County margins. This ensures that the election outcome in all counties is correct, as well as the State-wide outcome.

December 2007: Adjusting Audit Sample Size for Variation in Sizes of Auditable Vote Counts - Calculating the Exact Minimum Number of Vote Counts Required To Alter an Election Outcome

If the sizes (in number of ballots) of auditable vote counts (machine, precinct, or batch) vary significantly, it is possible for miscount to be targeted to a small number of the largest precincts, making larger audit samples necessary. Therefore we must directly calculate the minimum number of corrupt (precinct or other) vote counts that could alter election outcomes. Note: To conduct a valid, publicly verifiable election audit, a report of all vote counts for all candidates, and the number of ballots cast in each auditable vote count must be publicly released prior to the random selection of vote counts for audit.

To adjust for auditable vote count size (or precinct size) variation we need:

VoteCountSize = (n0, n1,…,) the total number of ballots cast in each precinct in which ballots could be cast for a race.

VoteCountRunnerVotes = (m0, m1, …) an array with the number of votes counted in each corresponding precinct or auditable vote count to the VoteCountsize array, for the losing candidate with the most votes

VoteCountWinnerVotes = (w0, w1,…) an array with the number of votes counted for the overall winner with the least number of votes in each vote count.

Maxshift = the maximum percentage of undetectable corruption expected on any one voting machine

[53] Dopp had widely released her and Stenger's work on email discussion lists months before their paper was published, but without exact numerical details. Upon Stanislevic's request, Dopp-Stenger mistakenly stated, in their "Election Integrity Audit" paper, that "Stanislevic and Dopp had independently derived equivalent audit calculation methods, both taking precinct size into account, using slightly different assumed maximum miscount rates per vote count." After careful evaluation, Dopp discovered that Stanislevic's algorithm for calculations was imprecise, often resulting in small errors in sample size and due to the fact that Stanislevic insisted on basing calculations on the number of votes counted, rather than on the number of ballots cast, he assumes that undervotes and overvotes are nonexistent or are recorded correctly. Stanislevic insisted in emails to Dopp, in using the number of votes counted, rather than the number of ballots cast, as a base for calculations.

Note that the number of ballots cast should be used as the metric for vote count size, so that election audit sample sizes are adequate to detect miscounted undervotes or overvotes. The algorithm for calculating the minimum number of miscounted vote counts (machine counts, batches, or precincts) which could alter election outcomes, accounting for vote count-sizes, in words, is: Calculate the margin error upper bounds for each vote count. Sort these error bounds in descending order (largest to smallest). Sum the margin error bounds beginning with the largest first, then adding the second largest until the sum of the error bounds is more than the margin (in votes) between the winner and the runnerup. Use the number of vote counts required for the minimum number of vote counts with error sufficient to reverse the outcome to calculate the election audit size.

This corrected algorithm uses the maximum possible miscount within each precinct that could alter an overall race outcome = the number of ballots plus the margin (in votes) in each vote count between the overall lowest-scoring winner and the highest-scoring runner up in the race:

```
% "%" precede comments.
% These input values are supplied by the user
% N = total number of precincts (or other auditable machine vote counts) in the race - optional
% Maxshift = the assumed maximum rate of undetectable margin miscount in each auditable vote count (or
precinct)
%      Suggest this be set by experts. I recommend using 0.50 = 50%
%       margin swing from the true vote – as the maximum amount a fraudster would dare to switch.
% Margin = margin between the highest runner-up and lowest winning candidates
%      in number of votes counted for each candidate in unofficial results
%
% VoteCountSize = an array with the total number of ballots cast in each precinct.
% VoteCountRunnerVotes = an array with the number of votes counted for the overall leading runnerup in
% each  race with precincts listed in the same order as in the Precinctsize array
% VoteCountWinnerVotes = an array with the number of votes counted for the overall winner with the least
% number of votes in the race
% PrecinctErrorBounds = an array with error bounds for each precinct
%
% Compute the number of precinct (or other auditable vote counts) required to alter the election outcome
%

Initialize Variables
for i=1 to i=N
          PrecinctErrorBound(i) = Maxshift*(VoteCountSize(i) +VoteCountWinnerVotes(i) –
     VoteCountRunnerVotes(i));
         TotBall = TotBall + VoteCountSize(i);
end for
Sort PrecinctErrorBound(i) in descending order (largest to smallest values)
     i = 1; % assumes array index begins with 1
     TotalError = PrecinctErrorBound(1);
while (TotalError < Margin)
          i = i +1;
          TotalError = TotalError + PrecinctErrorBound(i);
 end while
c = i;
Return (c);
```

Note that the maximum amount of vote miscount in any one precinct that could alter an election outcome is equal to the number of ballots plus the margin in votes between the overall winner and the overall loser with the highest number of votes.

Appendix C: Certain Audit Calculations Must be Based on "Ballots Cast", Not "Votes Counted"

This appendix shows why it is true that the election audit sample sizes recommended by following the calculation methods recommended by Verified Voting, the Brennan Center, and some other authors are insufficient to detect outcome-altering miscount whenever undervotes or overvotes have been the cause of overturning an election outcome because the margins and the number of miscounted precincts or auditable vote counts required to wrongly alter an election outcome, must be calculated with the number of ballots cast, not the number of votes, as a base.

An algebraic proof is not possible for the same reason that Stenger and Dopp used a numerical, rather than an algebraic, solution to the problem of determining exactly (to the correct integer value of auditable vote count units) the minimum election audit sample size required to achieve a desired probability for detecting outcome-altering vote miscount. It is the same reason that Aslam, Popa, and Rivest found a formula to estimate the exact solution, rather than solving the problem algebraically. I.e. there is probably not an algebraic proof available due to the complexity of the probability equations, so we settle for a logic proof.

PROOF:

Assertion: Using the total number of ballots cast in a race as a base automatically includes the number of under-votes/over-votes in the calculations - so under-votes are handled accurately because if

B = # cast ballots,
U = # undervotes and overvotes
W = #votes for winner,
R = #votes for the runner-up
M=margin %

for simplicity sake, we assume only two candidates.

B = U+W+R

$M = \frac{W - R}{U + W + R}$ i.e. difference between leading candidates out of total ballots cast.

Note that the above calculations include the number of undervotes and overvotes in the denominator, making the margins slightly smaller if there are undervotes and overvotes. The under-votes and over-votes must be included for calculating margins and for calculating the minimum number of auditable vote counts (machine counts, batches, or precincts) which could alter an election outcome.

Another illustration for why the Dopp method of using ballots cast works for calculating under-votes is seen by comparing Dopp's method of using ballots cast to the method recommended by Verified Voting and the Brennan Center of using votes counted.

On the other hand, margins calculated as the difference between leading candidates out of total votes counted would be:

$$M = \frac{W - R}{W + R}$$

Compare the two calculation methods for the rate of miscount that could alter an election outcome (overall) where V is the constant assumed max level of undetectability (= m/2V). Dopp's result is on the left side of the inequality below and the calculations recommended by Verified Voting and The Brennan Center is on the right side of the inequality:

$$\frac{W - R}{(U + W + R) * 2V} < \frac{W - R}{(W + R) * 2V}$$

Notice that whenever U, the number of undervotes and overvotes, is nonzero, the denominator of the expression on the left is larger, and hence the value of the expression on the left is smaller than the value of the expression on the right side.

In words, because the expression on the right side fails to include the number of undervotes/overvotes in its denominator, the result is an allegedly larger minimum number of miscounted precincts required to alter an election outcome resulting in a smaller election audit sample size than is actually necessary to detect a situation when miscounted undervotes or overvotes alter an election outcome.

Roy Saltman, in 1975 recognized and acknowledged this forgotten problem.

Basing auditing calculations on using the total number of ballots, as recommended by Dopp-Stenger in September 2006, "Election Integrity Audit", is vital in order for election audit samples to be sufficient for detecting situations when overvotes and undervotes wrongly alter election outcomes.

Dopp-Stenger (September 2006) recommend using the **the minimum of the county margin and the state margin in that race** as the margin for audit samples calculated on the county-level. See Appendix F.

Appendix D: 100% Hand-Counts Are Rarely Needed To Confirm Election Results

Simple arithmetic shows that 100% hand counts are rarely necessary to achieve 100% certainty of accurate election outcomes. We can stop counting ballots when it is clear that the winner-so-far won't change even if all of the remaining uncounted ballots have votes cast for the leading runner-up-so-far. When you stop in this way, the election outcome is known with 100% certainty, without a 100% hand count.

In one-winner races can stop counting ballots with 100% certainty of the outcome when:

1. The leading candidate has more than 50% of the total votes cast, because a candidate with over 50% of the votes is the clear winner, or when
2. The number of uncounted ballots is less than the margin between the two leading candidates because no candidate can catch up to the leader. If
   N1 is the count-so-far for the winner-so-far, and
   N2 is the count-so-far for the leading runner-up-so-far, and
   U is the number of uncounted votes,
   then if N1 > N2 and you can stop if N1 > N2 + U (i.e. if N1 - N2 > U)

With an elementary level arithmetic background, anyone can understand why manual counts of less than 100% of ballots effectively verify electronic election outcomes.

Partial hand counts can effectively detect cases when election outcomes are inaccurate to any desired probability.

For a spreadsheet expert, it is also fairly simple to correctly calculate sample sizes for 99% confidence-level election audits using the method described in Appendix B, along with the Aslam, Popa, & Rivest estimate, which at most overestimates the minimum necessary audit sample size by one auditable vote count.

Appendix E: Discrepancy Conditions Requiring Expanded Audits or Full Counts

The following discrepancies between voter-verified paper ballots and machine counts would put an election outcome in doubt:

1. if even one audited vote count has discrepancy changes favoring the leading runner-up and the discrepancy rate in that one vote count is equal to the assumed maximum level of undetectability (weighting vote switches by a factor of 1 and uncounted votes by ½)
2. if two times the net rate of switched vote discrepancies favoring a runner-up in the sample plus the net rate of uncounted vote discrepancies favoring the same runner-up in the sample is equal to the margin between the same runner-up and a winning candidate
3. if a losing candidate's votes are misrecorded as undervotes in the sample and the number of total undervotes in the race is equal to or greater than the margin between this losing candidate and a winning candidate times the proportion of such misrecorded undervotes for that candidate in the sample
4. if votes are consistently switched between any losing candidate and any winning candidate in the sample

Missing or unreadable paper ballots must be counted as "switched vote" discrepancies because otherwise anyone wanting to fraudulently alter outcomes could simply destroy or lose paper ballot records corresponding to fraudulent counts.

Philip Stark's October 2007 paper presents another method for evaluating election audit discrepancies in order to calculate a confidence level that the election outcome is correct, given the discrepancies found in the sample. Stark's method uses the margin between the winning and losing candidate and the maximum value of the sum of the absolute value of the differences between the manual and machine counts in every precinct for all candidates, treating overvotes and undervotes as one candidate, out of all sampled precincts.

It makes sense to use both Dopp and Stark's methods to evaluate election audit discrepancies because once the audit discrepancy data is available, it is easy to perform as many types of evaluations as desired, using a spreadsheet or other tools.

Further study is needed to precisely specify all the conditions that should trigger expanded election audits or full manual counts; and to specify exactly how to determine the amounts of such additional audits. A manual and toolkit for laypersons, created collaboratively with election officials, professional manual writers, and others, is needed to simply and fully explain all discrepancy conditions and to help auditors easily evaluate discrepancies between hand counts and machine counts.

Appendix F: County-Level Election Auditing

Dopp-Stenger (Sept. 2006) recommend calculating election audit sample sizes for each county by using the margin in each county that is the minimum of the county and the state race margins, using each county's total number of auditable vote counts for each race. This approach does more than verifying the statewide outcome. It

- verifies that the county-level outcome is correct in each county, and
- verifies that the miscount rate in each county is less than the average miscount that could incorrectly alter the statewide election outcome.

The benefits of using audit sample sizes designed to give confidence of county outcomes include:

- county level audits of even low design probability provide high confidence statewide, and
- if miscount in one county puts that county's outcome or the outcome of the entire race into question, then the expansion of an election audit may only need to occur in that one county and avoid inconveniencing other counties, whereas in state-level audits any significant discrepancy found in any one county requires expanding the audit in all counties.

If we believe that discrepancies requiring expanded audits would rarely, if ever, occur, then a state-level election audit will be more efficient (require smaller sample sizes). However, if discrepancies occur regularly and tend to be restricted to one or a few counties, then county-level election audits are more efficient because expanding audits in counties where no miscount is discovered in the initial audit may be unnecessary.

Using county-level audits, what is the chance for detecting at least one miscounted vote count state-wide if an election outcome is wrongly altered?

Given sufficient miscount sufficient to alter any county's outcome or sufficient miscount to alter an election outcome if it were to occur in all counties, the probability of finding one or more sufficiently corrupt vote counts state-wide is one (100%) minus the probability of finding no counties with such corrupt counts. If x = the probability for counties having such miscount to have *zero* miscounted vote counts in its sample (= 1 – the probability of having one or more corrupt counts), and c = the number of counties, then the probability finding at least one miscounted vote count statewide is $1-x^c$. If the audit design probability for each county is a relatively low 80% to detect one or more miscounted vote counts, and there are 50 counties, then the probability for the election audit to detect one or more miscounted vote counts state-wide would be $1-(0.20)^{50}$ = virtually 1 (100%).

The more counties a race occurs in, the lower the initial county confidence levels need to be to achieve a high state-wide confidence level for detecting any outcome-altering miscount. The above calculation does not find exactly the probability some experts may be looking for. Using the minimum of the county or state-wide margin for any race to calculate sample sizes, it finds the probability for detecting that at least one county outcome is incorrect if vote miscount level would change the statewide outcome if it occurs in all counties or if vote miscount level could change the county outcome. We have an exact probability for a somewhat complex situation that verifies both each county's and the state's outcome.

The state-wide total of audit sample sizes for county-level audits would be roughly the number of counties times the state-level audit sample size for a given race margin and audit design probability for detecting at least one corrupt count.

Appendix G: The Debate over Which Value to Use for Max Miscount per Vote Count

In order to calculate the size of an election audit sample that would ensure that outcome-reversing vote miscount is detected, one must first calculate the minimum number of miscounted vote counts (or precincts) that could wrongly alter (or reverse) an election outcome. The smaller the number of miscounted vote counts that could reverse an election outcome, the larger the audit sample size must be. So the larger the upper bound used for the amount of maximum vote miscount within any vote count, the larger the election audit sample size. There is a debate about which upper bound for the amount of miscount or error in each vote count (or precinct) should be used for calculations, and several proposals exist:

1. Saltman (1975) first proposed assuming a maximum level of vote miscount within any one vote count beyond (a maximum level of undetectability or MLU) beyond which any vote miscount would be noticed. If one assumes an MLU of 15% for vote switching, then the change in margins between winning and losing candidates would be altered 30% because 15% is subtracted from one candidate's total and 15% is added to the other's total.
2. Verified Voting proposes assuming 0.20 (20%) times the vote count as the maximum rate of vote switched miscount that would occur within any precinct.[54]
3. Dopp-Stenger (September 2006) initially proposed using 15% times the number of ballots as the maximum. Dopp later suggested (August 2007) using the minimum of 20% times the number of ballots, and the maximum possible bound for error which could alter or reverse the election outcome for the runnerup (However, Dopp's first attempt to calculate this amount worked only in situations of at most two candidates. i.e. the total number of ballots minus the number of votes for the overall loser with the most votes).
4. Stark (2007) proposes using an upper bound for vote miscount within each vote count or precinct that is the total amount of error possible within any vote count, i.e. the number of total ballots minus the minimum of the number of votes for the candidate with the least number of votes, or the number of total undervotes and overvotes. Starks's upper bound for the amount of total error possible is higher than the amount of error that can contribute to reversing an election outcome because some errors cancel out other errors for the outcome.
5. Dopp proposes (Dec 2007) as shown in Appendix B above, using the maximum possible bound for margin error within each vote count that could reverse the outcome of the overall runner-up and multiplying it time a factor such as 0.50 to reflect that fact that 100% of the available votes cannot be shifted without raising suspicion[55]: $\left(r_{i1} - r_{i2} + b_i\right)$

[54] Verified Voting also uses the number of votes counted as its base for calculations rather than ballots cast, which means their sample sizes are smaller than those Dopp or Stark recommend whenever there are significant undervotes or overvotes.
[55] Note Aslam, Popa, & Rivest also mention this same error bound in an intermediate step on p. 11 of their paper *On Auditing Elections When Precincts Have Different Sizes*. Dopp also recommends Dopp also recommends doing calculations to detect any vote counts (or precincts) which appear to have margin shifts from expected values of more than the 30% or 40% value assumed for designing the audit sample sizes; and auditing those vote counts as well and doing county-level audits which increase the overall state-wide sample size. See appendix F in this paper.

Appendix H: Derivation of the Probability Proportional to Margin Error Bound (PPMEB) Method

This attempts to provide an easier-to-understand and more complete description of the derivation for the number of selection rounds t for the PPMEB method which derivations were originally provided in the Aslam, Rivest, Popa, and Stark papers.

> *On Auditing Elections When Precincts Have Different Sizes* Javed A. Aslam, College of Computer and Information Science, Northeastern University, Raluca A. Popa and Ronald L. Rivest, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology (Draft of January 17, 2008)
> http://people.csail.mit.edu/rivest/AslamPopaRivest-OnAuditingElectionsWhenPrecinctsHaveDifferentSizes.pdf
>
> *Election audits by sampling with probability proportional to an error bound: dealing with discrepancies* Philip B. Stark, Department of Statistics, University of California, Berkeley, CA 94720-3860, (DRAFT 20 February 2008)
> http://statistics.berkeley.edu/~stark/Preprints/ppebwrwd08.pdf

Let:
M= the difference in the number of votes counted for the winner and the closest runnerup
N= the total number of precincts in an election contest
e_i = the discrepancy between the manual count and the electronic count for precinct i, i=1,…,N.
T = a set of precincts with discrepancies for which the sum of their discrepancy could wipe out the margin $\sum_{i \in T} e_i \geq M$ (1)

If audits find strong statistical evidence that no such T exists, then the election result can be certified with high confidence.

Let $\{u\}$ = a vector, the set of upper bounds, u_i on the amount of discrepancy possible within each precinct. For example, $u_i = b_i + w_i - r_i$, the number of ballots cast plus the margin in votes between the winner and runner up is the upper bound for error which could change the outcome between the winner and the closest runnerup. Then for any set of precincts, the sum of the upper error bounds is greater or equal to the sum of the actual errors $\sum_i u_i \geq \sum_i e_i$ and if T exists, then $\sum_{i \in T} u_i \geq \sum_{i \in T} e_i \geq M$

For simplicity in our notation, let $E = \sum_i u_i$ the sum of upper bounds in all precincts.

If we draw precinct i according to its proportional share of possible margin error
Then the probability of precinct i being selected is $p_i = \frac{u_i}{E}$ (2)
The probability that any single draw yields an element of T would be

$$p_T = \sum_{j \in T} \frac{u_j}{E} = \frac{1}{E} \sum_{j \in T} u_j \qquad (3)$$

If T exists, then the chance that we get no elements of T in one draw is $1 - \frac{1}{E} \sum_{j \in T} u_j$ (4)

The chance β that we get *no* elements of T, a set of precincts with sufficient error to wipe out the margin, in t draws is $\beta = \left(1 - \frac{1}{E} \sum_{j \in T} u_j\right)^t \le \left(1 - \frac{M}{E}\right)^t$ or $\beta \le \left(1 - \frac{M}{E}\right)^t$ (5)

The probability $1 - \beta$ that we get at least one element of T in t draws is

$$1 - \beta = 1 - \left(1 - \frac{1}{E} \sum_{j \in T} u_j\right)^t \ge 1 - \left(1 - \frac{M}{E}\right)^t \qquad (6)$$

Next our goal is to set an upper bound α for the chance β of drawing no elements of T if T exists.

$$\beta \le \left(1 - \frac{M}{E}\right)^t \le \alpha \qquad (7)$$

In other words, if T exists we want the chance β of drawing *no* elements of T to be less than an α that we set. For instance if $\alpha = 0.05$ then the chance $1 - \alpha$ that we get at least one element of T in t draws is greater or equal than $1 - \alpha = 95\%$

In other words, we want a lower bound for the chance $1 - \beta$ that we draw at least one element of T in t selection rounds $1 - \alpha \ge 1 - \left(1 - \frac{M}{E}\right)^t \ge 1 - \beta$ (8)

So, solving for t in (7) or (8) we obtain a lower bound for t, the number of selection rounds of

$$t \ge \frac{\ln \alpha}{\ln(1 - \frac{M}{E})} \qquad (9)$$

So, if our desired confidence is 95% then $\alpha = 1 - 95\%$

And then t the number of selection rounds must be at least, $t \ge \frac{\ln(0.05)}{\ln(1 - \frac{M}{E})}$ (10)

For a simpler description of the steps in this method see *Mandatory Post-Election Vote Count Audits --Probability Proportional to Error Bound (PPEB) Method* Dopp (DRAFT February 21, 2008) http://electionarchive.org/ucvAnalysis/US/paper-audits/UT/VoteCountAudits.pdf