



How Can Independent Paper Audits Ensure Election Integrity?

Routine independent audits of vote counts in all elections in a properly calculated percentage of randomly selected precincts¹ could detect virtually any vote count errors that might wrongfully alter an election outcome so that they could be corrected.

Why Independently Audit Vote Counts?

A majority of U.S. jurisdictions tally votes electronically and perform no routine independent audits to ensure accuracy. Independent audits of banks, churches and businesses are performed to detect and prevent errors and embezzlement. Vote counts determine who controls budgets worth millions to trillions of dollars, yet little is done to ensure that vote counts are accurate. This is especially alarming in light of the fact that U.S. jurisdictions publicly report their vote counts aggregated in a way that hides evidence of vote count tampering.²

Requiring voter-verifiable paper ballot records is insufficient. The paper ballot records must be audited by persons other than the voting machine vendor, elections officials, or other insiders.

How to Independently Audit Vote Counts?

If voter-verifiable paper ballots are counted for a sufficient percentage of randomly selected precincts or machines there is a high probability of detecting any errors in vote counts that could wrongly alter an election outcome.³ If discrepancies are found, an investigation, an expanded audit or a county-wide recount could be triggered. Additional funding may need to be allocated in order to perform routine independent audits of vote counts.

What is the Probability of Finding a Miscalculated Precinct with a 2% or 5% Audit?

To calculate the probability of detecting *at least one* miscalculated precinct in any county or township, we need to know:

1. How many precincts are there in the county?
2. What percentage of corrupt precincts do we want to be able to detect?⁴
3. What percentage of precincts or machines will be randomly selected for audits?

In the tables on the next page, probabilities for finding at least one corrupted precinct are calculated, based on 1000 precincts with varying percentages of corrupted precincts, using 2% and 5% independent audits.⁵

¹ "Randomly selected" is a specifically designed process. An expert needs to be hired to design a truly random process for selecting precincts. Having election officials select the precincts does not qualify as "random" selection.

² See "What Election Data Can We Collect and Publicly Release to Monitor our Elections for Accuracy?" http://uscountvotes.org/ucvAnalysis/US/election_officials/ElectionArchive_advice.pdf

³ A random audit, in combination with a mathematical analysis of detailed election results, would ensure election integrity because if we assume that the proportion of precincts with vote miscounts is small, then the amount of error in the miscalculated precincts would be high enough to be detectable via mathematical analysis of election results. i.e. The analysis of election results would be likely to catch errors that were under the design threshold of the paper audit.

⁴ A good rule of thumb for determining the percentage of corruption is to divide the smallest margin between two candidates by 0.60 because this gives an approximation of the smallest percentage of corrupted machines needed to alter the outcome of any race thus ensuring that any outcome-altering vote miscounts would be detected.

⁵ Details of the calculation are in Appendix A. USCountVotes has made a spreadsheet available to find the probability for detecting miscalculated precincts for a particular county. <http://ElectionArchive.org/ucvAnalysis/US/paper-audits/>
Copyright 2006 Kathy Dopp Anyone may freely use, copy, or disseminate under the sole condition that it be properly attributed to the author. June 30, 2005 updated 7/25/2006 ElectionArchive.org

Table 1: Chance of Finding a Corrupted Precinct with a 2% Audit of 1000 Precincts

Total Number of Precincts	Hypothetical Number of Corrupted Precincts	Percent Corrupted Precincts	Number of Audited Precincts for 2% Audit	Expected Value	Standard Deviation	Chance of Finding a Corrupted Precinct
1000	500	50.00%	20	10.00	2.24	100.00%
1000	450	45.00%	20	9.00	2.22	100.00%
1000	400	40.00%	20	8.00	2.19	100.00%
1000	350	35.00%	20	7.00	2.13	99.98%
1000	300	30.00%	20	6.00	2.05	99.93%
1000	250	25.00%	20	5.00	1.94	99.70%
1000	200	20.00%	20	4.00	1.79	98.90%
1000	150	15.00%	20	3.00	1.60	96.25%
1000	100	10.00%	20	2.00	1.34	88.10%
1000	50	5.00%	20	1.00	0.97	64.51%
1000	40	4.00%	20	0.80	0.88	56.15%
1000	30	3.00%	20	0.60	0.76	45.94%
1000	10	1.00%	20	0.20	0.44	18.37%
1000	5	0.50%	20	0.10	0.32	9.63%
1000	1	0.10%	20	0.02	0.14	2.00%
1000	0.5	0.05%	20	0.01	0.10	0.00%

Table 1 (above) shows that a 2% audit has a probability of over 95% of finding a corrupted precinct if at least 15% or 150 out of 1000 precincts are corrupted.

Table 2: Chance of Finding a Corrupted Precinct with a 5% Audit of 1000 Precincts

Total Number of Precincts	Hypothetical Number of Corrupted Precincts	Percent Corrupted Precincts	Number of Audited Precincts for 5% Audit	Expected Value	Standard Deviation	Chance of Finding a Corrupted Precinct
1000	500	50.00%	50	25.00	3.54	100.00%
1000	450	45.00%	50	22.50	3.52	100.00%
1000	400	40.00%	50	20.00	3.46	100.00%
1000	350	35.00%	50	17.50	3.37	100.00%
1000	300	30.00%	50	15.00	3.24	100.00%
1000	250	25.00%	50	12.50	3.06	100.00%
1000	200	20.00%	50	10.00	2.83	100.00%
1000	150	15.00%	50	7.50	2.52	99.98%
1000	100	10.00%	50	5.00	2.12	99.55%
1000	50	5.00%	50	2.50	1.54	92.80%
1000	40	4.00%	50	2.00	1.39	87.68%
1000	30	3.00%	50	1.50	1.21	79.03%
1000	10	1.00%	50	0.50	0.70	40.27%
1000	5	0.50%	50	0.25	0.50	22.66%
1000	1	0.10%	50	0.05	0.22	5.00%
1000	0.5	0.05%	50	0.03	0.16	0.00%

Table 2 shows that a 5% audit has a probability of over 90% (the lower end of customary “statistical significance”) of finding a corrupted precinct if at least 5% or 50 out of the 1000 are corrupted. See <http://ElectionArchive.org/ucvAnalysis/US/paper-audits/> for an Audit_Calculator.xls spreadsheet.

The purpose for auditing election results is to ensure that **no outcome is altered by vote count errors**. Therefore, the calculation to determine the number of vote counts to audit incorporates an estimate for the minimum number of vote counts that would need to be corrupted to alter an election outcome. (More information is available in NEDA's AuditCalculator.xls spreadsheet.)

The following chart shows how the election results margins between candidates affect the likelihood of a 3% audit percentage detecting a sufficient level of corruption to alter the outcome of a race with margins given in the left-most column. A 3% audit is clearly insufficient to detect whether close races have correct outcomes in a small county with 155 total machine or precinct-level vote counts to audit.

Exact Probabilities of Finding One or More Corrupt Counts						
CONSTANTS	Max Vote Switch per Machine = 15%		Max Margin Shift per machine = 30%		Note: The HYPGEOMDIST(A,B,C,D) formula is the chance of finding A target elts s.t. B = #elts in the sample, C= #target elts in sample, D=#elts in total pop	
INPUTS	# Total County-Wide Vote Counts = 155		# County Vote Counts to Audit = 4			
Margin between Candidates	Percent of Corrupt Counts Needed to Alter Race Outcome	Hypothetical Number of Corrupt Counts to Alter a Race	Expected # of Corrupt Counts in Sample w/ Outcome-Altering Corruption	Standard Deviation	Chance of Finding one or more Corrupt Counts	% Audit
1%	3.33%	6	0.1	0.36	14.74%	3%
2%	6.67%	11	0.3	0.50	25.73%	3%
3%	10.00%	16	0.4	0.60	35.62%	3%
4%	13.33%	21	0.5	0.68	44.48%	3%
5%	16.67%	26	0.7	0.75	52.40%	3%
6%	20.00%	32	0.8	0.80	60.75%	3%
7%	23.33%	37	0.9	0.85	66.82%	3%
8%	26.67%	42	1.1	0.88	72.16%	3%
9%	30.00%	47	1.2	0.92	76.83%	3%
10%	33.33%	52	1.3	0.94	80.89%	3%
11%	36.67%	57	1.5	0.96	84.38%	3%
12%	40.00%	63	1.6	0.98	87.92%	3%
13%	43.33%	68	1.7	0.99	90.38%	3%
14%	46.67%	73	1.9	1.00	92.44%	3%
15%	50.00%	78	2.0	1.00	94.15%	3%
16%	53.33%	83	2.1	1.00	95.55%	3%
17%	56.67%	88	2.3	0.99	96.69%	3%
18%	60.00%	94	2.4	0.98	97.74%	3%
19%	63.33%	99	2.5	0.96	98.41%	3%
20%	66.67%	104	2.7	0.94	98.92%	3%
21%	70.00%	109	2.8	0.92	99.29%	3%
22%	73.33%	114	2.9	0.88	99.56%	3%
23%	76.67%	119	3.1	0.85	99.75%	3%
24%	80.00%	125	3.2	0.80	99.88%	3%
25%	83.33%	130	3.3	0.75	99.95%	3%

Appendix A:

The probability estimates are based on a “Hypergeometric” distribution which determines the probability of finding:

- a) x (target corrupted precincts) We let $x = 0$ to find the probability of detecting no corrupted precincts.
- b) in an overall sample of n (Number of Audited Precincts) which in this case is 20 (for 2% audit) or 50 (for 5% audit),
- c) when there are X (Hypothetical Number of Corrupted Precincts) which in two tables above range from 50% to 0.05% of 1000. However, the best way to calculate X is N times [the smallest margin between any candidates in any race] divided by [two times the maximum percentage of corruption expected on any one voting machine] to give the number of corrupt counts that could alter an election outcome.
- d) out of N (Total Number of Precincts) which in this case is 1000.

This distribution is calculated using the Excel Function:

$$\text{HYPGEOMDIST}(x, n, X, N) = \frac{\binom{X}{x} \binom{N-X}{n-x}}{\binom{N}{n}}$$

The hypergeometric function assumes all individual “picks” are random but adjusts this random probability for each pick. The first row in Table 1 for example, assumes a 50% probability that the first of 20 picks will be corrupted, a 499/999 probability that the second precinct chosen will be corrupted, and so on for all 20 picks.

The probability that one or more of the 20 precincts will be a corrupted precinct is 1 or 100% minus the probability that none of the 20 precincts will be corrupted. So, the probability that *at least one* of the 20 is corrupted which equals:

$$P = 1 - \frac{\binom{X}{0} \binom{N-X}{n}}{\binom{N}{n}}$$

This equation can be solved using numerical methods for the sample size n .

For any county, the probability distribution will depend on the exact values of n , X , and N , so that it is best to recreate tables that are specific to the number of precincts of the county for which the audit probability estimates are being made.

A free spreadsheet calculator for determining the probabilities of detecting one or more miscounted precincts is available on the USCV web site:

<http://ElectionArchive.org/ucvAnalysis/US/paper-audits/>

Prepared by

National Election Data Archive Project (NEDA) - ElectionArchive.org

501-3c status approved

Contact: Kathy Dopp kathy@uscountvotes.org 435-608-1382

The National Election Data Archive is a scientific research project whose mission is to help ensure the accuracy of elections in America through development of best audit practices and the creation and analysis of a database containing precinct-level election data broken out by vote type for the entire United States. To volunteer please email volunteers@uscountvotes.org

Copyright 2006 Kathy Dopp Anyone may freely use, copy, or disseminate under the sole condition that it be properly attributed to the author. June 30, 2005 updated 7/25/2006 ElectionArchive.org