VoTeR Center

UConn Voting Technology Research Center

PI : A. Shvartsman, Ph.D.
Co-PIs: A. Kiayias Ph.D., L. Michel Ph.D., A. Russell Ph.D.
Staff: T. Antonyan, S. Davtyan, R. Jancewicz, E. Kovalev, J. Newmann

Statistical Analysis of the Post-Election Audit Data
2011 November Election

June 5, 2012

Version 1.1

Abstract

This report presents the analysis of the post-election audit returns performed in the State of Connecticut following the November 2011 election. The audit involved the randomly selected 10% of the districts. The total of 877 audit records were submitted for analysis.

The initial review of audit reports prepared by the towns contained 45 records with differences between hand and machine count higher than 5. There were 7 records with discrepancies higher than 9, and only one record with a discrepancy higher than 15 votes between hand and machine counts.

Follow up investigations were conducted by the Office of the Secretary of the State (SOTS) to determine the cause of discrepancies. This resulted in a revision of the audit data for the districts that were the subject of the follow up.

This report presents the analysis of the audit returns in two parts: (i) the analysis of 887 records (100%) among which 45 records (5.1%) were revised based on the follow up investigation conducted by the SOTS Office, (ii) the analysis of the original 887 records (100%) prior to the follow up investigation. The second part is provided for completeness and comparison.

This analysis was performed on request of the Office of the Secretary of the State.

Table of Contents

Summary

The University of Connecticut Center for Voting Technology Research (VoTeR Center) received the data gathered in the post-election audit performed in the State of Connecticut following the November 2011 election. The audits involved the randomly selected 10% of the districts and the audit returns were conveyed by the Office of the Secretary of the State (SOTS) to the VoTeR Center on December 22nd of 2011. The audit data received by the Center contains 887 records, where each record represents information about a given candidate: date, district, machine seal number, office, candidate, machine counted total, hand counted total of the votes considered unquestionable by the auditors, hand counted total of the votes considered questionable by the auditors, and the hand counted total, that is, the sum of undisputed and questionable ballots. This report contains several statistical analyses of the audit returns and recommendations.

The VoTeR Center's initial review of audit reports prepared by the towns identified a number of returns with unexplained differences between hand and machine counts. Audit returns included 45 records with discrepancies higher than 5, with the highest reported discrepancy of 40. It is worth noting that 75% (30 out of 45) of the records that were subject to the follow up investigation already contained information indicating that the discrepancies were due to the human error. Following this initial review the SOTS Office performed additional information gathering and investigation of those 45 records. The final information was conveyed to the Center on May 18th of 2012. The rest of the records (842 out of 887) discussed in this audit report are the original records reported by the towns.

This report presents the analysis of 887 records (100%) in two parts.

- First, we present analysis of 887 records (100%), among which 842 records (94.9% of 887) are original records, as reported by the districts, and 45 records (5.1% of 887) were revised based on the follow up conducted by the SOTS Office.

- Second, we present analysis of the original 887 records (100%) prior to the SOTS Office follow up investigation. This is provided for completeness and comparison.

For the revised records SOTS Office confirmed with the districts that the discrepancies were due to human counting errors. After the revised records were incorporated, the analysis shows that among 887 (100%) records there are 400 (45.1%) records showing no discrepancy, 244 records (27.5%) showing discrepancy of 1 vote, 107 records (12.1%) showing discrepancy of 2 votes, and 136 records (15.3%) showing discrepancy of 3 to 5 votes.

The data presented in this analysis show that the average reported discrepancy is lower than the number of questionable votes on the ballots (1.1 versus 3.5).

This analysis was performed on request of the Office of the Secretary of the State.

Preface

The University Of Connecticut Center for Voting Technology Research (VoTeR Center) received the data gathered in the post-election audit performed in the State of Connecticut following the November 2011 election. The audits of the randomly selected 10% of the districts were conducted in November of 2011, and the returns were conveyed by the Office of the Secretary of the State (SOTS) to the VoTeR Center on December 22nd of 2011. The initial analysis feedback was subsequently communicated to SOTS, and on May 18th, the final revised data was received by the Center.

For the definition of the audit see Connecticut Public Act 07-194 AN ACT CONCERNING THE INTEGRITY AND SECURITY OF THE VOTING PROCESS, approved July 5, 2007. For the instructions on conducting the audit, see Audit Procedures Optical Scan Voting Equipment, Office of the Secretary of the State, November 2007.

In accordance with the Act, the SOTS office conveys the report documenting hand audit returns to the VoTeR Center, and the Center is in turn required to report on its analysis to the SOTS Office:

"(d)….Such report shall be filed with the Secretary of the State who shall immediately forward such report to The University of Connecticut for analysis. The University of Connecticut shall file a written report with the Secretary of the State regarding such analysis that describes any discrepancies identified. After receipt of such report, the Secretary of the State shall file such report with the State Elections Enforcement Commission."

The following subsections of the audit law are also highly relevant:

"(i) If the audit officials are unable to reconcile the manual count with the electronic vote tabulation and discrepancies, the Secretary of the State shall conduct such further investigation of the voting machine or tabulator malfunction as may be necessary for the purpose of reviewing whether or not to decertify the voting machine or machines in question or to order the voting machine to be examined and recertified….

(o) As used in this section, "discrepancy" means any difference in vote totals between machine and manual counts in a voting district that exceeds one-half of one percent of the lesser amount of the vote totals between machine and manual counts where such differences cannot be resolved through an accounting of ballots…"

This analysis was performed on request of the Office of the Secretary of the State.

1   Overview of the Analysis

This report contains several statistical analyses of the audit returns. The VoTeR Center received 887 records on December 22nd of 2011. On May 18th of 2011 VoTeR Center received the final revised 45 records that originally contained discrepancies higher than 5 between the machine counts and audit hand counts.

The statistical analysis in this report deals with the 887 records (100%). We present our analysis in three parts. First, we present analysis of 887 records (100%), among which 842 records (94.9% of 887) are original records, as reported by the districts, and 45 records (5.1% of 887) were revised based on the follow up conducted by the SOTS Office. Second, we present analysis of the original 887 records (100%) prior to the follow up investigation.

The conclusion of the SOTS Office follow up is that for all cases where non-trivial discrepancies were originally reported, it was determined that hand counting errors or vote misallocation were the causes. No discrepancies in these cases were reported to be attributable to machine tabulation. For the original audit returns for which no follow up investigation was performed, the discrepancies are relatively small.

For completeness we present our analysis of the audit returns in two parts (all records prior to the revisions to the 45 records, and all, unrevised records). The main part of the analysis deals with 887 (100%) records, among which 45 records (5.1%) were revised based on the follow up investigation conducted by the Secretary of the State Office. Among 887 (100%) records there are 400 (45.1%) records showing no discrepancy, 244 records (27.5%) showing discrepancy of 1 vote, 107 records (12.1%) showing discrepancy of 2 votes, and 136 records (15.3%) showing discrepancy of 3 to 5 votes.

The average number of votes recorded for the candidates is 372. The overall average number of questionable votes per district is 3.47. The marked ballots are determined to be "questionable" by the human auditors: a ballot is questionable if the auditors believe that it is marked in such a way that the machine will likely not be able to read it properly. Note that this does not mean that the machine absolutely would not read it. Given that this assessment is based on human judgment call, it is predictable that in many cases hand counts would not match machine counts.

The average absolute discrepancy between the machine count and the hand count performed in the audit is 1.09. This number is computed by taking the sum of the absolute (positive) values of the discrepancies in all records and dividing this sum by the number of records. Thus, on the average reported discrepancy (1.09) is smaller than the average number of reported questionable votes (3.47). Overall this is a good indication, suggesting that, on average, despite the presence of questionably marked ballots, the machine count is very close to the hand count.

One conclusion is that hand counting of the ballots during the audit is an error-prone process. In all cases where large discrepancies were investigated, it was reported that hand counting was not performed correctly, or that the correctly counted votes were misallocated as is common with the cross-party endorsements.

The detailed analysis of the audit returns are given in two parts in Sections 3 and 4.

2   Introduction and Notation

Throughout this document we use the following notation:

- M is used to denote the machine counted ballots

- U is used to denote the number of undisputed hand counted ballots

- Q is used to denote the number of questionable hand counted ballots

- H is the sum of undisputed and questionable ballots, that is, $H = U + Q$

- D is the discrepancy between the hand counted total and machine total, that is, $D = H - M$

  Thus for a given candidate, we define discrepancy D as the difference between H (the sum of the undisputed ballots U and the questionable ballots Q) and M (the machine count).

  If the discrepancy D is positive then we say that we observe a machine undercount relative to the hand count H, i.e., the machine counted fewer ballots than the auditors.

  If the discrepancy D is negative then we say that we observe a machine overcount relative to the hand count H, i.e., the machine counted more ballots than the auditors.

- |D| is the absolute value of the discrepancy (or the positive value of D)

  This means that if D is positive, then $|D| = D$, and if D is negative, then $|D| = -D$.

Note that this presupposes that the hand count does not contain (human counting) errors. This is not necessarily so in actuality. However, since in general it is not possible to ascertain whether the hand

counted data contain errors, we assume that the hand counted data is reported correctly, unless a follow up investigation determined otherwise.

3   Statistical Analysis of 887 Records (includes 45 revised records)

This section deals with 887 records (100%), among which 842 records (94.9%) are from the original data and 45 records (5.1%) were revised based on the follow up conducted by the SOTS office.

Figure 1 is the graphical representation of the discrepancy distribution. Later in this section we analyze the absolute value of discrepancy, the pattern of undercounts and overcounts, and the percentage of the votes reported as questionable.
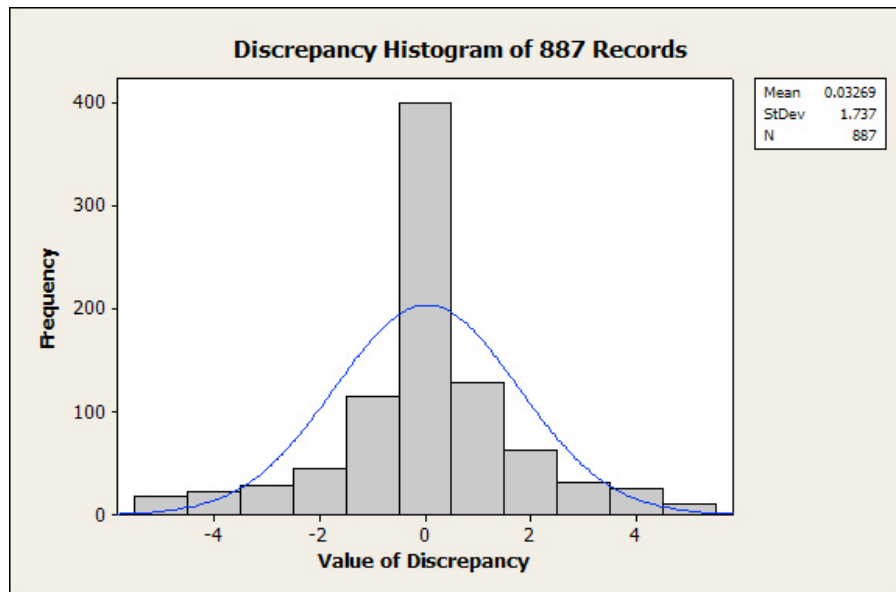

Figure 1. Final discrepancy histogram (including 45 revised records)

3.1   Absolute Value of Discrepancy

Here we give the analysis for the absolute number of discrepancies, |D|. We include discrepancies for all records for which both the machine count M and the total hand count H is given. For the 887 records considered here, the average absolute discrepancy is 1.09, and the standard deviation is 1.35, suggesting that the occurrences of discrepancies are clustered in the vicinity of the average. Table 1 presents tiered view of the absolute discrepancies.

Table 1: Absolute value of discrepancy.

Description	Counts	% of Counts
Records with discrepancy |D| of 0	400	45.1%
Records with discrepancy |D| of 1	244	27.5%

Records with discrepancy |D| of 2 to 3	168	18.9%
Records with discrepancy |D| of 4 to 5	75	8.5%
Totals:	**887**	**100%**

Table 2 presents tiered view of the absolute discrepancies by the percentage of discrepancy.

Table 2: By Percentage of Discrepancy

Description	Counts	% of Counts
Records with discrepancy 0% to 0.5%	665	75%
Records with discrepancy > 0.5% to 1%	112	12.6%
Records with discrepancy > 1% to 2%	45	5.1%
Records with discrepancy > 2% to 5%	42	4.8%
Records with discrepancy > 5% to 10%	10	1.1%
Records with discrepancy > 10% to 20%	6	0.7%
Records with discrepancy > 20% to 30%	3	0.3%
Records with discrepancy > 30% to 40%	2	0.2%
Records with discrepancy > 40% to 50%	0	0%
Records with discrepancy > 50%	2	0.2%
Totals:	**887**	**100%**

3.2   Undercount and Overcount Discrepancies

When considering negative discrepancies (overcounts) and positive discrepancies (undercounts) for the 887 records, the average discrepancy is 0.03, and the standard deviation is 1.74.

Table 3 presents discrepancies for the records that indicate overcounts.

**Table 3: Records indicating overcounting: 229 (25.8% of 887) records
with negative values of discrepancy.**

Description	Counts	% of Counts
Records with discrepancy D of -1	115	50.2%
Records with discrepancy D of -2 to -3	74	32.3%
Records with discrepancy D of -4 to -5	40	17.5%
Totals:	**229**	**100%**

Table 4 presents discrepancies for the records that indicate undercounts.

**Table 4: Records indicating undercounting: 258 (29.1% of 887) records
with positive values of discrepancy.**

Description	Counts	% of Counts
Records with Discrepancy D of 1	129	50%
Records with Discrepancy D of 2 to 3	94	36.4%
Records with Discrepancy D of 4 to 5	35	13.6%
Totals:	**258**	**100%**

3.3   Statistics for Questionable Ballot Counts

The average number of questionable votes per record is 3.47.

Table 5 presents statistics with respect to the questionable ballots per candidate.

Table 5: Questionable Ballot Counts.

Description	Counts	% of Counts
Records with questionable count Q of 0	445	50.2%
Records with questionable count Q > 0 to 2%	313	35.3%
Records with questionable count Q > 2% to 5%	102	11.5%
Records with questionable count Q > 5% to 10%	14	1.6%

Records with questionable count Q > 10%	13	1.4%
Totals:	**887**	**100%**

4 Statistical Analysis of 887 Original Records

This section deals with 887 original records (100%). Figure 2 is the graphical representation of the discrepancy distribution. Later in this section we analyze the absolute value of discrepancy, the pattern of undercounts and overcounts, and the percentage of the votes reported as questionable.
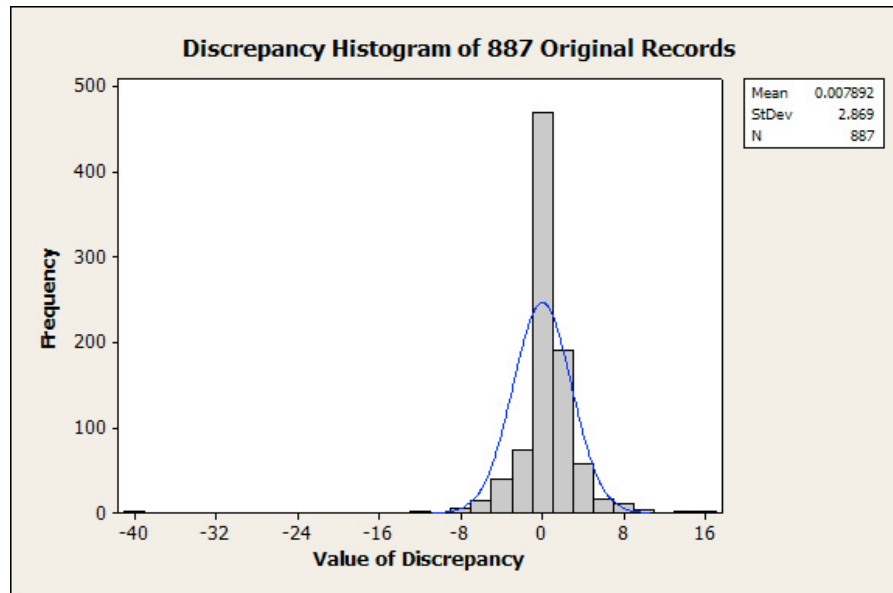

Figure 2. Discrepancy histogram of 887 Original Records

4.1 Absolute Value of Discrepancy

Here we give the analysis for the absolute number of discrepancies, |D|. We include discrepancies for all records for which both the machine count M and the total hand count H is given. For the 887 records considered here, the average absolute discrepancy is 1.53, and the standard deviation is 2.43, suggesting that the occurrences of discrepancies are clustered in the vicinity of the average. Table 6 presents tiered view of the absolute discrepancies.

Table 6: Absolute value of discrepancy.

Description	Counts	% of Counts
Records with discrepancy |D| of 0	355	40%
Records with discrepancy |D| of 1	244	27.5%

Records with discrepancy |D| of 2 to 3	168	18.9%
Records with discrepancy |D| of 4 to 5	75	8.5%
Records with discrepancy |D| of 6 to 9	37	4.2%
Records with discrepancy |D| of 10 to 15	7	0.8%
Records with discrepancy |D| of 40	1	0.1%
Totals:	**887**	**100%**

Table 7 presents tiered view of the absolute discrepancies by the percentage of discrepancy.

Table 7: By Percentage of Discrepancy

Description	Counts	% of Counts
Records with discrepancy 0% to 0.5%	625	70.5%
Records with discrepancy > 0.5% to 1%	124	14%
Records with discrepancy > 1% to 2%	59	6.7%
Records with discrepancy > 2% to 5%	49	5.5%
Records with discrepancy > 5% to 10%	11	1.2%
Records with discrepancy > 10% to 20%	9	1%
Records with discrepancy > 20% to 30%	6	0.7%
Records with discrepancy > 30% to 40%	2	0.2%
Records with discrepancy > 40% to 50%	0	0%
Records with discrepancy > 50%	2	0.2%
Totals:	**887**	**100%**

4.2   Undercount and Overcount Discrepancies

When considering negative discrepancies (overcounts) and positive discrepancies (undercounts) for the 887 records, the average discrepancy is 0.008, and the standard deviation is 2.89.

Table 8 presents discrepancies for the records that indicate overcounts.

Table 8: Records indicating overcounting: 252 (28.4% of 887) records with negative values of discrepancy.

Description	Counts	% of Counts
Records with discrepancy D of -1	115	45.6%
Records with discrepancy D of -2 to -3	74	29.4%
Records with discrepancy D of -4 to -6	49	19.4%
Records with discrepancy D of -7 to -9	11	4.4%
Records with discrepancy D of -10 to -15	2	0.8%
Records with discrepancy D of -40	1	0.4%
Totals:	**252**	**100%**

Table 9 presents discrepancies for the records that indicate undercounts.

Table 9: Records indicating undercounting: 280 (31.6% of 887) records with positive values of discrepancy.

Description	Counts	% of Counts
Records with Discrepancy D of 1	129	46.1%
Records with Discrepancy D of 2 to 3	94	33.6%
Records with Discrepancy D of 4 to 6	41	14.6%
Records with Discrepancy D of 7 to 9	11	3.9%
Records with Discrepancy D of 10 to 15	5	1.8%
Totals:	**280**	**100%**

11

4.3   Statistics for Questionable Ballot Counts

The average number of questionable votes per record is 3.47.

Table 10 presents statistics with respect to the questionable ballots per candidate.

Table 10: Questionable Ballot Counts.

Description	Counts	% of Counts
Records with questionable count Q of 0	445	50.2%
Records with questionable count Q > 0 to 2%	313	35.3%
Records with questionable count Q > 2% to 5%	102	11.5%
Records with questionable count Q > 5% to 10%	14	1.6%
Records with questionable count Q > 10%	13	1.4%
Totals:	**887**	**100%**

5   Conclusions

The analysis observes that on the average the absolute number of reported discrepancies (for complete audit records) is smaller than the average of the reported questionable votes. This is consistent with prior audits. Here one may conclude that the machines counts are mostly not affected by the questionable ballots, resulting in small differences between the machine reported totals and the hand counted totals.

Future improvements should address the reporting of the analysis and the analysis itself. A major change planned for future analysis is to assess the impact of the perceived discrepancies on the election outcomes. This is going to be exceedingly important for the cases where a race may be very close, but where the difference between candidates is over 0.5% (thus not triggering an automatic recount).

We also note that there are improvements in the hand count process. In most cases great attention was paid to the discovered discrepancies. As noted by the auditors, in most of the cases when a discrepancy between the hand and machine counts was observed it was due to the fact that the bubbles were not marked or filled in correctly by the voters. The examples of such improper marking of the ballots, as given by the auditors, are: bubbles with light coloring, having 'X' mark or check mark in the bubble or circling the bubble instead of filling it in.

Lastly, the analysis of and the follow up on the hand-counted audit returns in this and prior years show that hand-counting of ballots is a laborious process that is prone to human counting errors. The SOTS Office in collaboration with the VoTeR Center recently won a U.S. Election Assistance Commission grant to improve the audit methodology and to develop computer-assistive technology that will make audits more effective and easier to conduct.

[end]