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, E. Kovalev, D. Lawrence

Statistical Analysis of the Post-Election Audit Data
2010 November Election

June 28, 2011

Version 1.2

Abstract

This report presents the analysis of the post-election audit returns performed in the State of Connecticut following the November 2010 election. The audit involved the randomly selected 10% of the districts. The initial review of audit reports prepared by the towns revealed a number of returns with substantially high unexplained differences between hand and machine counts. The majority of the records with discrepancies higher than 20 votes were concentrated in the following three districts: East Haven (Deer Run School) - the highest absolute discrepancy reported was 180, Hartford (Burns School) - the highest absolute discrepancy reported was 170, Preston (Town Hall) - the highest absolute discrepancy reported was 55. In addition, in spite of the fact that the highest discrepancy reported in Orange (Mary L Tracy School) was 14, the differences were still unexplained as no questionable ballots were reported.

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.

The audit data received by the VoTeR Center contains 867 records. Among the 867 records received by the Center, 20 records (2.3%) were incomplete. This report deals with 847 records (97.7%) among which 799 records (94.3%) are from the original data and 48 records (5.7%) were revised based on the follow up conducted by the SOTS office.

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 2010 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 2010. The audit data received by the VoTeR Center contains 867 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 statistical analysis in this report deals with the 847 records that are sufficiently complete to perform the analysis.

The VoTeR Center's initial review of audit reports prepared by the towns revealed a number of returns with unexplained differences between hand and machine counts. The vast majority of records with high discrepancies were concentrated in the following three districts: East Haven (Deer Run School) with the highest reported discrepancy of 180, Hartford (Burns School) with the highest reported discrepancy of 170, and Preston (Town Hall) with the highest reported discrepancy of 55. Additionally, one or more discrepancies were reported in all but one district for the town of Orange; here the highest reported discrepancy was 14, however this could not be explained as no questionable ballots were reported. Following this initial review the SOTS Office performed additional information gathering and investigation and, in some cases, conducted independent hand-counting of ballots in the four districts mentioned above. The final information was conveyed to the VoTeR Center on June 17th of 2011 for the 48 records pertaining to those districts. The rest of the records (799 out of 847) discussed in this audit report are the original records reported by the towns.

This report presents the analysis of 847 records (97.7%), among which 799 records (94.3%) are from the original data and 48 records (5.7%) were revised based on the follow up conducted by the SOTS office. For the revised records the discrepancy was reduced to 0 for 38 records (79.2%), the remaining 10 records indicate a discrepancy of either 1 vote (10.4% of the revised records) or 2 votes (10.4% of the revised records). Among 847 (100%) records there are 485 (57.3%) records showing no discrepancy, 132 records (15.6%) showing discrepancy of 1 vote, 187 records (22.1%) showing discrepancy of 2 to 5 votes, 42 records (4.9%) showing discrepancy of 6 to 13 votes (for this group, although no manual review of the discrepancies was conducted, the SOTS Office affirmed that the discrepancies were due to hand counting errors), and 1 record (0.1%) showing discrepancy of 25 votes. The last discrepancy of 25 votes belongs to a cross-party endorsed candidate. The SOTS Office confirmed that this discrepancy was due to misallocation of correctly counted votes to different party totals.

The data presented in this analysis show that the average reported discrepancy is lower than the number of questionable ballots (1.26 versus 1.54).

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 2010 election. The audits of the randomly selected 10% of the districts were conducted in November of 2010, and the returns were conveyed by the Office of the Secretary of the State (SOTS) to the VoTeR Center on December 22nd of 2010.

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 867 records on December 22nd of 2010. Among 867 records received by the VoTeR Center 20 records were incomplete and are not part of the analysis performed in this report. On June 17th of 2010 VoTeR Center received revised records for four towns that originally reported unexplained differences between the machine counts and audit hand counts.

The statistical analysis in this report deals with the 847 records (97.7%) that were complete and contained no obvious errors. Among those, 799 records (94.3%) are from the original data and 48 records (5.7%) were revised based on the follow up conducted by the SOTS office. For completeness the report includes a discrepancy histogram for of the original (unrevised) 847 complete records in Figure 1 of Section 3.

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.

Among 847 (100%) records there are 485 (57.3%) records showing no discrepancy, 132 records (15.6%) showing discrepancy of 1 vote, 187 records (22.1%) showing discrepancy of 2 to 5 votes, 42 records (4.9%) showing discrepancy of 6 to 13 votes (for this group, although no manual review of the discrepancies was conducted, the SOTS Office affirmed that the discrepancies were due to hand counting errors), and 1 record (0.1%) showing discrepancy of 25 votes between the machine counts and audit hand counts. The last discrepancy of 25 votes belongs to a cross-party endorsed candidate. The SOTS Office confirmed that this discrepancy was due to misallocation of correctly counted votes to different party totals.

The average number of votes recorded for the candidates is 329. The overall average number of questionable votes per district is 1.54. The 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 1.26. 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 is smaller than the average number of reported questionable votes. 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.

It is also noted that 20 records (2.3% of 867) were found to be incomplete. The statistical analysis does not include these records. While some problematic records are clearly due to human error (e.g., errors in addition), in other cases it appears that auditors either did not follow the audit instructions precisely, or found the instructions to be unclear.

The detailed analysis is in Section 3.

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 847 Complete Records

This section deals with 847 complete records (97.7%), among which 799 records (94.3%) are from the original data and 48 records (5.7%) were revised based on the follow up conducted by the SOTS office.

For completeness we first present the data distribution for the original 847 records received by the VoTeR Center that do not include any revised records. Figure 1 is the graphical representation of the data distribution for discrepancies found in 847 original records.
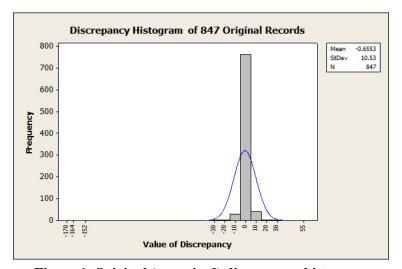

Figure 1. Original (unrevised) discrepancy histogram

As Figure 1 shows most of the originally reported discrepancies for discrepancies for the 778 (91.8%) complete records fall between -5 and 5. There are 57 records (6.7%) with absolute discrepancies in range between 5 and 15. And 12 records (1.5%) have discrepancies with the absolute value higher than 15. In particular, there were 4 records with discrepancies 55, -152, -164, and -170. The follow up investigation showed that the highest discrepancies were due to the hand counting error.

The rest of this section will deal with the analysis of the 847 records containing 48 revised records. 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.
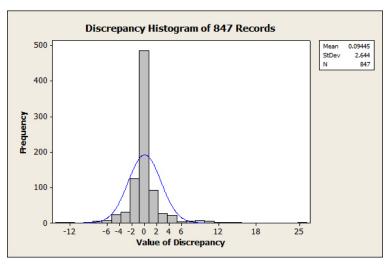
Figure 2. Final discrepancy histogram (including 48 revised records)

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 847 records considered here, the average absolute discrepancy is 1.25, and the standard deviation is 2.33, 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	485	57.3%
Records with discrepancy |D| of 1-3	274	32.3%
Records with discrepancy |D| of 4-5	45	5.3%
Records with discrepancy |D| of 6-10	33	3.9%
Records with discrepancy |D| of 11-13	9	1.1%
Records with discrepancy |D| of 25	1	0.1%
Totals:	**847**	**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%	628	74.1%
Records with discrepancy > 0.5% to 1%	47	5.5%
Records with discrepancy > 1% to 2%	18	2.1%
Records with discrepancy > 2% to 5%	53	6.3%
Records with discrepancy > 5% to 10%	33	3.9%
Records with discrepancy > 10% to 20%	26	3.1%
Records with discrepancy > 20% to 30%	16	1.9%
Records with discrepancy > 30% to 40%	10	1.2%
Records with discrepancy > 40% to 50%	4	0.5%
Records with discrepancy > 50%	12	1.4%
Totals:	**847**	**100%**

3.2 Undercount and Overcount Discrepancies

When considering negative discrepancies (overcounts) and positive discrepancies (undercounts) for the 847 records, the average discrepancy is 0.09, and the standard deviation is 2.64.

Table 3 presents discrepancies for the records that indicate overcounts.

**Table 3: Records indicating overcounting: 195 (23% of 847) records
with negative values of discrepancy.**

Description	Counts	% of Counts
Records with discrepancy D of -1 to -3	155	79.5%
Records with discrepancy D of -4 to -6	31	15.9%
Records with discrepancy D of -7 to -9	7	3.6%
Records with discrepancy D of -12	1	0.5%
Records with discrepancy D of -14	1	0.5%
Totals:	**195**	**100%**

Table 4 presents discrepancies for the records that indicate undercounts.

Table 4: Records indicating undercounting: 167 (19.7% of 847) records with positive values of discrepancy.

Description	Counts	% of Counts
Records with Discrepancy D of 1-5	140	83.8%
Records with Discrepancy D of 6-10	19	11.4%
Records with Discrepancy D of 11-15	7	4.2%
Records with Discrepancy D of 25	1	0.6%
Totals:	**167**	**100%**

3.3   Statistics for Questionable Ballot Counts

The average number of questionable votes per record is 1.54.

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	590	69.7%
Records with questionable count Q > 0 to 2%	160	18.9%
Records with questionable count Q > 2% to 5%	62	7.3%
Records with questionable count Q > 5% to 10%	16	1.9%
Records with questionable count Q > 10%	19	2.2%
Totals:	**847**	**100%**

4   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).

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.

[This space is left intentionally blank.]

[end]