VoTeR Center

The Center for Voting Technology Research

PI : A. Schwarzmann, Ph.D.
Co-PIs: L. Michel Ph.D., A. Russell Ph.D.
Senior Staff: D. Farney
Staff: Th. Hadjistasi, O. Kisielius

Statistical Analysis of the Post-Election Audit Data for the November 2016 Presidential Elections

November 27th, 2017

Version 1.1

Abstract

This report presents the analysis of the post-election returns from the audit performed in the State of Connecticut following the November 08, 2016 presidential elections. The audit involved the 10% of the districts randomly selected by the Office of the Secretary of the State. The total of 654 audit records were submitted for analysis.

In this report we consider an audit record to be unusable if either (a) it reports the machine count or the undisputed hand count or the overall hand count as zero; or (b) the record contains non-integer data; or (c) the record was improperly filled out. In this election, 39 (5.9%) of the 654 audit records were unusable and all of them fell into category (a).

Out of these, 615 records (94.0% of 654 records) contained sufficient data for analysis, and this report focuses on these records.

Among the 615 usable records the breakdown of discrepancies is as follows: 445 records (72.4%) show no discrepancy, 85 records (13.8%) show discrepancy of 1 vote. There are 58 records (9.4%) showing a discrepancy of 2 or 3 votes; 14 records (2.3%) showing a discrepancy of 4 to 6 votes; 6 records (1%) showing a discrepancy of 7 to 10 votes; 3 records (0.5%) showing a discrepancy of 11 to 20 votes and 4 records (0.6%) showing a discrepancy of more than 20 votes.

The Center followed up with the towns that had records showing a discrepancy of 6 or more votes. There were 15 records with discrepancies of 6 or higher. Based on the follow up and explanations of discrepancies given in hand count audit reports, all the 15 records with discrepancies 6 or higher are considered as explained discrepancies in this report.

This analysis of the 615 records revealed no indication suggesting inaccuracy in the tabulator counts in the audited districts.

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

Table of Contents

Summary

The Center for Voting Technology Research (VoTeR Center) at the School of Engineering of the University of Connecticut received the data gathered in the post-election audit performed in the State of Connecticut following the November 8th, 2016 election. The audit 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 February 23, 2017. The audit data received by the Center contains 654 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.

This report presents the analysis of 615 records. This is the total number of records less the 39 unusable records where all of them reported the machine count, the undisputed hand count, or the overall hand count as zero.

Among the 615 usable records the breakdown of discrepancies is as follows: 445 records (72.4%) show no discrepancy, 85 records (13.8%) show discrepancy of 1 vote. There are 58 records (9.4%) showing a discrepancy of 2 or 3 votes; 14 records (2.3%) showing a discrepancy of 4 to 6 votes; 6 records (1%) showing a discrepancy of 7 to 10 votes; 3 records (0.5%) showing a discrepancy of 11 to 20 votes and 4 records (0.6%) showing a discrepancy of more than 20 votes.

The Center followed up with the towns that had records showing a discrepancy of 6 or more votes. There were 15 records with discrepancies of 6 or higher. Based on the follow up and explanations of discrepancies given in hand count audit reports, all the 15 records with discrepancies 6 or higher are considered as explained discrepancies in this report.

The main cause for discrepancies between the hand and machine counts appears to be human error in counting as reported by the auditors.

The data presented in this analysis show that the average reported discrepancy is lower than the average number of questionable votes on the ballots (0.75 versus 1.39). This analysis revealed no indication suggesting inaccuracy in the tabulator counts in the audited districts.

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

Preface

The Center for Voting Technology Research (VoTeR Center) at the School of Engineering of the University of Connecticut received the data gathered in the post-election audit performed in the State of Connecticut following the November 8th, 2016 election. The audits of the randomly selected 10% of the districts were conducted in November of 2016, and the returns were conveyed by the Office of the Secretary of the State (SOTS) to the VoTeR Center on February 23, 2017.

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.

Notation

Throughout this document we use the following notation to discuss the audit records:

- M is used to denote the number of machine counted votes.

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

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

- H is the sum of undisputed and questionable votes, 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 votes 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 votes 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. 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 with a district determined otherwise.

1   Overview of the Analysis

This report contains several statistical analyses of the audit returns. The VoTeR Center received 654 records on February 23, 2017. Of these 654 records, a total of 39 were improperly filled out and considered unusable for the report. All the unusable records reported the machine count, the undisputed hand count, or the overall hand count as zero. Thus, the total number of records considered in this report is 615.

The statistical analysis in this report deals with the 615 records. Among the 615 usable records the breakdown of discrepancies is as follows: 445 records (72.4%) show no discrepancy, 85 records (13.8%) show discrepancy of 1 vote. There are 58 records (9.4%) showing a discrepancy of 2 or 3 votes; 14 records (2.3%) showing a discrepancy of 4 to 6 votes; 6 records (1%) showing a discrepancy of 7 to 10 votes; 3 records (0.5%) showing a discrepancy of 11 to 20 votes and 4 records (0.6%) showing a discrepancy of more than 20 votes.

The auditors provided explanations for many cases of such differences, with the typical explanations quoted below:

- "Ballots were counted only once thus probably human hand-count error".

- "Human Error."

- "Mis-attribution of votes across the cross-endorsed candidates due to human error".

The average number of votes recorded per candidate is 430.2. The overall average number of questionable votes is 1.39. A marked ballot is determined to be "questionable" by the auditors 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, 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 0.75. This number is computed by taking the sum of the absolute (positive) values of the discrepancies (464) in all records and dividing this sum by the number of records (615). Thus, on the average reported discrepancy (0.75) is lower than the average number of reported questionable votes (1.39). 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.

There are 39 records that were not usable for this report. These records merit a follow up. Out of the 39 records, 2 (5.2%) of them appear to show no discrepancy between the hand count and machine count, however the audit return shoes that the tabulator count is equal to the questionable vote count. There are 14 records (35.9%) out of the 39 that show discrepancies between the hand count and machine count. This is because of the improperly filled out and unreadable data. Table 1 lists districts that have these records.

Table 1: List of districts that have not properly filled out records.

Town and District	# of Records
Meriden District 10	1
Hartford District 8	2
Colchester District 3	4
Glastonbury District 4	1
Montville District 3	1
Enfield District 1	5
Total	14

Lastly, there are 23 records (58.9%) out of 39 that were provided containing empty data. These records were clearly not usable for the audits and are not considered for this report.

The detailed analyses of the audit returns are given in Sections 3.

Discrepancies of 6 and More Votes

There are 15 out of 615 records (2.4%) showing a discrepancy of 6 and more votes. The VoTeR Center contacted registrars of voters in several districts to discuss discrepancies that were higher than usual. In particular, attention was paid to records where the machine total was either less than the undisputed hand count (M < U) or greater than the total hand count (M > O).

In all cases when the registrars were contacted, they stated that they did not believe the tabulators were responsible for the high discrepancies. In light of those conversations, Table 2 presents the 600 records that remain if the districts with unusual discrepancies are not considered.

Table 2: Percentage absolute discrepancy for records, omitting records from districts contacted by the VoTeR Center.

Description	Counts	% of Counts
Records with discrepancy from 0% to 0.1%	591	98.5
Records with discrepancy from 0.1% to 0.2%	6	1
Records with discrepancy from 0.2% to 0.3%	3	0.5
Totals:	600	100

Based on the follow up and explanations of discrepancies given in hand count audit reports, all the 15 records with discrepancies 6 or higher are considered as "explained" discrepancies in this report.

Table 3 lists districts that have these records and the sizes of discrepancies given in the records.

Table 3: List of districts that have "unexplained" records.

Town and District	# of Records	Discrepancies
Hartford District 8	2	6 and 7
Hamden District 7	1	8
East Windsor District 1	4	8, 21, 22 and 26

Stamford District 6	1	24
Groton District 6	1	8
Enfield District 1	1	10
Wolcott District 2	5	6, 7, 13, 11 and 15

A summary of explanations follows.

- Hartford District 8 had 2 records with 6 and 7 discrepancies. The explanation of these differences is mis-attribution of votes across the cross-endorsed candidates due to human error.
- Hamden District 1 had 1 record with 8 discrepancies. A follow up with the district explains this to be a "hand-counting error."
- East Windsor District 1 had 4 records with 8, 21, 22, and 26 discrepancies. The explanation of these differences is mis-attribution of votes across the cross-endorsed candidates due to human error.
- Stamford District 6 had 1 record with 24 discrepancies. The explanation given was "human error".
- Groton District 6 had 1 record with 8 discrepancies. The explanation given was that "ballots were counted only once thus this is probably a hand-count error".
- Enfield District 1 had 1 record with 10 discrepancies. The explanation of these differences is mis-attribution of votes across the cross-endorsed candidate" due to human error.
- Wolcott District 2 had 5 records with 6, 7, 13, 11, and 15 discrepancies. The explanation given in the record was "in handcount-audit the unknowns may have been misallocated". This is due to "human error".

2   Statistical Analysis of 615 Records

This section deals with 615 records. 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.
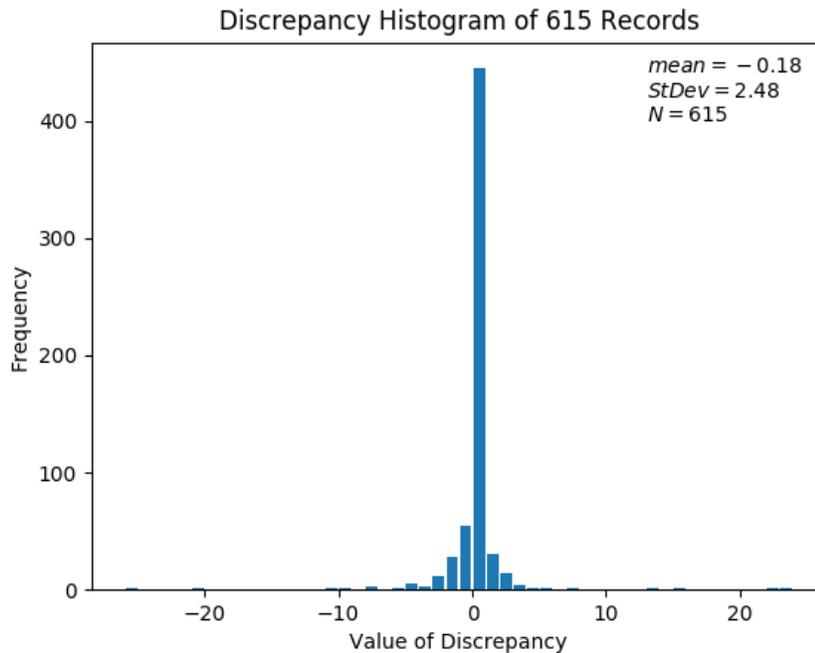
Figure 1. Discrepancy Histogram of 615 Records.

2.1 Absolute Value of Discrepancy

Here we give the analysis for the absolute number of discrepancies, |D|. We include discrepancies for all 615 records for which both the machine count M and the total hand count H is given. For the 615 records considered here, the average absolute discrepancy is 0.75, and the standard deviation is 2.48, suggesting that the occurrences of discrepancies are clustered in the vicinity of the average. Table 4 presents tiered view of the absolute value of discrepancies.

Table 4: Absolute value of discrepancy.

Description	Counts	% of Counts
Records with a discrepancy |D| of 0	445	72.4
Records with a discrepancy |D| of 1	85	13.8
Records with a discrepancy |D| of 2 to 3	58	9.4
Records with a discrepancy |D| of 4 to 6	14	2.3
Records with a discrepancy |D| of 7 to 10	6	1
Records with a discrepancy |D| of 11 to 20	3	0.5
Records with a discrepancy |D| above 20	4	0.6
Totals:	**615**	**100**

Table 5 presents a tiered view of the absolute discrepancies as a percentage of the total hand-counted votes (H) for each record.

Table 5: Absolute value of discrepancy as percentage of records.

Description	Counts	% of Counts
Records with discrepancy from 0% to 0.1%	604	98.2
Records with discrepancy from 0.1% to 0.2%	8	1.3
Records with discrepancy from 0.2% to 0.3%	3	0.5
Totals:	**615**	**100**

The percentage of discrepancy is calculated by dividing the absolute discrepancy |D| by the machine total M and multiplying by 100%, that is, the percentage $(|D| / M) \cdot 100\%$ (the machine total M is assumed to be non-zero).

2.2   Undercount and Overcount Discrepancies

When considering negative discrepancies (overcounts) and positive discrepancies (undercounts) for the 615 records, the average discrepancy is -0.18 (-112/615) votes, and the standard deviation is 2.48 votes. 112 records (18.2%) show negative discrepancies, 445 records (72.4%) show no discrepancies, and 58 records (9.4%) show positive discrepancies.

Table 6 presents a tiered view of discrepancies for the records that indicate overcounts.

Table 6: Records with negative discrepancy D, indicating overcounting.

Description	Counts	% of Counts
Records with a discrepancy D of -1	55	8.9
Records with a discrepancy D of -2 to -3	40	6.5
Records with a discrepancy D of -4 to -6	10	1.6
Records with a discrepancy D of -7 to -10	4	0.7
Records with a discrepancy D greater than -11	3	0.5
Totals:	**112**	**18.2**

Table 7 presents a tiered view of discrepancies for the records that indicate undercounts.

Table 7: Records with positive discrepancy D, indicating undercounting.

Description	Counts	% of Counts
Records with a discrepancy D of 1	30	4.8
Records with a discrepancy D of 2 to 3	18	2.9
Records with a discrepancy D of 4 to 6	4	0.7
Records with a discrepancy D of 7 to 10	2	0.3
Records with a discrepancy D greater than 11	4	0.7
Totals:	**58**	**9.4**

2.3   Statistics for Questionable Ballot Counts

The average number of questionable votes Q per record is 1.39, and the standard deviation is 2.96 votes.

Table 8 presents a tiered view of questionable vote counts Q.

Table 8: Number of questionable vote counts Q.

Description	Counts	% of Counts
Records with questionable votes Q of 0	401	65.2
Records with questionable votes Q of 1 to 2	107	17.4
Records with questionable votes Q of 3 to 4	37	6
Records with questionable votes Q of 5 to 6	28	4.6
Records with questionable votes Q of 7 to 9	20	3.3
Records with questionable votes Q of 10 to 20	21	3.4
Records with questionable votes Q greater than 20	1	0.1
Totals:	**615**	**100**

Table 9 presents a tiered view of questionable vote counts Q as a percentage of the total hand-counted votes (H) for each record.

Table 9: Questionable vote counts Q as percentage.

Description	Counts	% of Counts
Records with 0% questionable votes	401	65.2
Records with questionable votes from 0% to 0.1%	214	34.8
Records with questionable votes greater than 0.1%	0	0
Totals:	**615**	**100**

3   Conclusions

The analysis shows that on average the absolute number of reported discrepancies for complete audit reports (0.75 votes) is smaller than the absolute value of the reported questionable votes (1.39 votes). This is consistent with prior audits. This suggests that on the average tabulation imprecision affects the counts less than the discrepancies that are possibly due to questionable votes. Overall, based on the reported audits, there is no indication of errors that would be due to tabulator issues/errors.

We note that reporting of questionable votes is likely inconsistent due to the differing interpretation of bubbles by some auditors.

While the overall audit process may be improving, there remain a number of unexplained differences for which in all cases districts concluded that "human error" was the cause. Hand-counted audit is a laborious process that is prone to human counting errors as many auditors reported for this and previous elections. Computer-assisted audits, when used appropriately, have the potential to make audits more effective, more accurate, as well as faster and easier to conduct. Such a system was piloted successfully, and it is anticipated that it will be deployed in the near future.

There were 39 records (5.9%) that contained inaccurate hand counted audit data. These records reported the machine count, the undisputed hand count, or the overall hand count as zero. Clearly these audit reports are not completed carefully, and they are unusable for this analysis. In a number of cases the audit reports are poorly scanned or even faxed, and in many cases hand-written audit reports are hard to read and interpret, and it is suggested that in the future such reports are produced with the help of a word processor. The auditors should carefully complete the audit reports in the future.

[end]