VoTeR Center

UConn Center for Voting Technology Research

PI : A. Shvartsman, Ph.D.
Co-PIs: A. Kiayias Ph.D., L. Michel Ph.D., A. Russell Ph.D.
Staff: T. Antonyan, S. Davtyan, D. Lawrence, N. Nicolaou, T. Smith

Statistical Analysis of the Post-Election Audit Data
2010 August Primary Election

October 26, 2010

Version 1.1

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 August 2010 primary 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 from September 21st (almost all returns) to October 26th of 2010. The audit data received by the VoTeR Center contains 465 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.

The review of the audit reports prepared by the towns did not reveal any returns with high unexplained differences between hand and machine counts. The largest discrepancy is a single case of 6 votes (4.5%), the second largest is a single case of 4 votes (0.99%). The rest show discrepancies of one or less, with 95.7% of the returns showing no discrepancies. There are in total twenty records with discrepancies. We note that in seven cases out of twenty the auditors provided an explanation for the observed discrepancies. Majority of the explicated cases refer to the improper markings of the ballots by voters. One of the examples of such an improper marking is having 'X' mark in the bubble instead of filling in the bubble. The optical scan machine used in Connecticut, the AccuVote optical scanner, is able to recognize partially marked bubbles in some cases, but this depends on the bubble fill pattern.

This report presents the analysis of the submitted audit returns. No records were found to be incomplete, unusable, or obviously incorrect; this is a substantial improvement. Among 465 (100%) records there are 445 (95.7%) records showing no discrepancy, 18 records (3.9%) show discrepancy of 1 vote, and 2 records (0.4%) show discrepancy of 4 and 6 votes between the machine counts and audit hand counts. The maximum number of discrepancies is 6 votes.

The data presented in this analysis show that the average reported discrepancy is much lower than the number of questionable ballots (0.06 versus 1). Moreover, for the majority of the records (70%) showing discrepancy, the number of questionably marked ballots (as determined by the human auditors) is greater than the reported discrepancy.

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

Table of Contents

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 August 2010 election. The audits of the randomly selected 10% of the districts were conducted in August and September of 2010, and the returns were conveyed by the Office of the Secretary of the State (SOTS) to the VoTeR Center between September 21st (almost all) and October 26th 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, updated in 2010 and issued before the primary.

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 465 records on September 21st and October 26th of 2010. All 465 original records are sufficiently complete to perform the analysis. The maximum discrepancy is a single case of 6 votes.

Among the 465 records discussed in this report there are 445 records (95.7%) showing no discrepancy, 18 records (3.9%) showing discrepancy of 1 vote, 1 record (0.2%) showing discrepancy of 4 votes, and 1 record (0.2%) showing discrepancy of 6 votes between the machine counts and audit hand counts. The largest discrepancy is 6.

The average number of votes recorded for the candidates is 81. The overall average number of questionable votes per district is 1. 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 0.06. 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 much smaller than the average number of reported questionable votes. Overall this is a good indication that suggests that, on average, despite the presence of questionably marked ballots, the machine count is very close to the hand count.

The detailed analysis is in Section 3.

It is worth mentioning that this is the first post-election audit where no records were found to be incomplete, unusable, or obviously incorrect (one exception is a single audit return with unclear handwriting; this was resolved after a follow up). This indicates a substantial improvement in the overall hand count audit process.

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 465 Records

This section deals with all 465 original audit records – all are complete and contain no obvious audit errors.

Figure 1 is the graphical representation of the data distribution for discrepancies found in these 465 records. We then analyze the absolute value of discrepancy, the pattern of undercounts and overcounts, and the percentage of the votes reported as questionable.
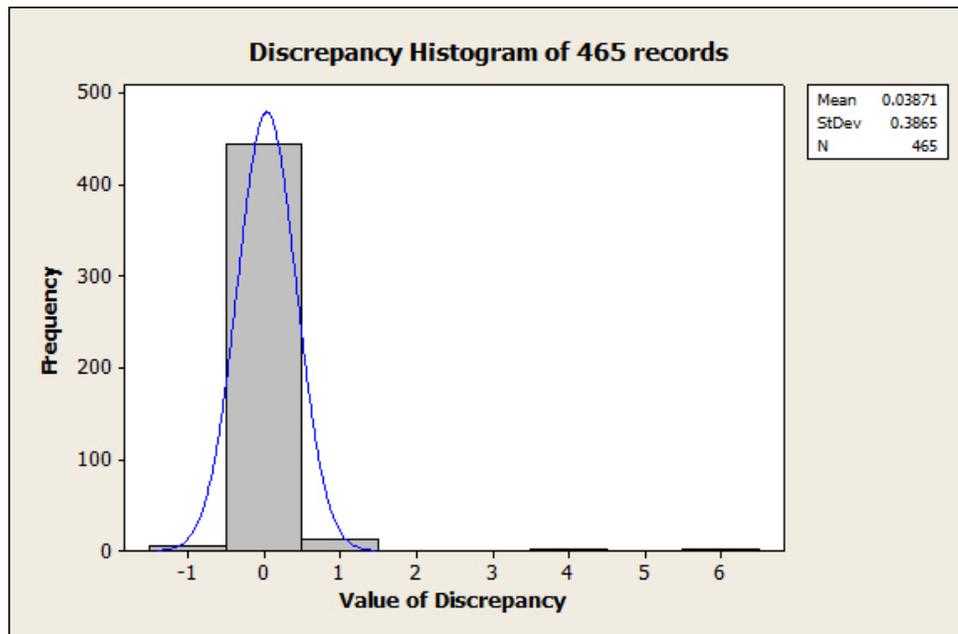
Figure 1

3.1 Absolute Value of Discrepancy

Here we give the analysis considering 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 465 records considered here, the average absolute discrepancy is 0.06, and the standard deviation is 0.38, 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	445	95.7%
Records with discrepancy |D| of 1	18	3.87%
Records with discrepancy |D| of 4	1	0.215%
Records with discrepancy |D| of 6	1	0.215%
Totals:	**465**	**100%**

We note that for the two cases where the discrepancies were above 2, the situation is as follows.

- Discrepancy of 4 is reported in Windham (Windham High School), where the difference is between the machine count of 400 differs from the hand count of 404, which represents less than 1% undercount.

- Discrepancy of 6 is reported in Easton (Samuel Staples School), where the difference is between the machine count of 127 differs from the hand count of 133, which represents 4.5% undercount. Here the handwriting on the audit report is very unclear, and the results reported here is the best effort estimate based on the discussion with Easton.

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 less than 0.5%	451	96.99%
Records with discrepancy 0.5% to 1%	6	1.29%
Records with discrepancy 1% to 2%	4	0.86%
Records with discrepancy 2% to 5%	2	0.43%
Records with discrepancy 5% to 10%	1	0.215%
Records with discrepancy 10% to 20%	0	0%
Records with discrepancy over 20%	1	0.215%
Totals:	**465**	**100%**

We note that for the four cases where the discrepancies were above 2%, the situation is as follows.

- New Britain (Graham Apartments) reports 25 machine counted votes vs. 26 hand counted votes. So while the discrepancy is 3.8%, this is due to a single vote difference.

- Easton (Samuel Staples School) reports a discrepancy of 4.5% (this is discussed earlier).

- East Haven (East Farm Village) reports 11 machine votes vs. 12 hand counted votes. While the discrepancy is 8.33%, this is due to a single vote difference.

- Bridgeport (Hallen School) reports 2 machine votes vs. 3 hand counted votes. While the discrepancy is 833.3%, this is due to a single vote difference.

3.2   Undercount and Overcount Discrepancies

When considering negative discrepancies (overcounts) and positive discrepancies (undercounts) for the 465 records, the average discrepancy is 0.04, and the standard deviation is 0.39, again suggesting that the occurrences of discrepancies are clustered in the vicinity of the average.

Table 3 presents discrepancies for the records that indicate overcounts.

Table 3: Records indicating overcounting: 5 (1.07% of 465) records with negative values of discrepancy.

Description	Counts	% of Counts
Records with discrepancy D of -1	5	100%
Totals:	**5**	**100%**

Table 4 presents discrepancies for the records that indicate undercounts.

Table 4: Records indicating undercounting: 15 (3.23% of 432) records with positive values of discrepancy.

Description	Counts	% of Counts
Records with Discrepancy D of 1	13	86.66%
Records with Discrepancy D of 4	1	6.67%
Records with Discrepancy D of 6	1	6.67%
Totals:	**15**	**100%**

3.3   Statistics for Questionable Ballot Counts

The average number of questionable votes per record is 0.97.

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	316	67.96%
Records with questionable count Q > 0 to 2%	72	15.48%
Records with questionable count Q > 2% to 5%	41	8.82%
Records with questionable count Q > 5% to 10%	26	5.59%
Records with questionable count Q > 10%	10	2.15%
Totals:	**465**	**100%**

4   Conclusions

We note that the quality of the submitted audit reports improved substantially as compared to previous audits. Only one audit report required follow up due to unclear handwriting.

The analysis did not reveal any audit returns with high differences between the machine counts and hand counts. In the very few cases where high percentage of discrepancies is reported, the actual differences are one vote for a very small number of ballots cast. Thus these cases are not statistically significant.

One important additional observation is that on the average the absolute number of reported discrepancies (for complete, unrevised, original audit records) is substantially smaller than the average of the reported questionable votes. This is consistent with prior audits. Here one may conclude that the machines are still able to correctly tabulate most questionable ballots, resulting in small differences between the machine reported totals and the hand counted totals.

[This space is left intentionally blank.]

[end]