VoTeR Center
UConn Voting Technology Research Center
http://voter.engr.uconn.edu

PI : A. Shvartsman
Co-PIs : A. Kiayias, L. Michel, A. Russell

Analysis of the Audit Returns for the November 2006 Election

Kishori Konwar Laurent Michel Alexander A. Shvartsman
with the assistance of Karpoor Shashidhar, Andrew See, and Seda Davtyan

September 12, 2007

Abstract

The University of Connecticut Voting Technology Research (VoTeR) Center received the data gathered in the random audits performed in the State of Connecticut following the November 2006 election. The data was submitted to the Secretary of the State Office on November 28, 2006, and preliminary analysis was published on December 6, 2006. One district was re-audited in March 2007 to examine a substantial discrepancy; it was determined that the original audit results contained errors. The analysis presented here addresses the following questions:

1. Is the mean of the discrepancies between the machine-counted results and the results hand-counted in the audit close to 0?
2. What is the impact of the individual machines at distinct districts on the differences in counts?
3. Are the overall machine counts and hand counts proportional? (More specifically, is the proportionality constant 1?)
4. What is the impact of the accuracy of the voting machines on election results?

For the questions above, the statistical analysis of the audit data indicates that:

1. The overall discrepancies between the machine counts and hand counts are not statistically significant.
2. The effect of individual machines is not significant (a single machine was used in each district).
3. The machine counts and hand counts are proportional with the proportionality constant 1, although the machines have a tendency to overcount, on the average by 1/2 vote.
4. The accuracy of the voting machines did not affect the election results (although no audited races were particularly close). One needs to be careful, however, in relying on machine counts when the total number of ballots is small, especially in multi-opening races.

1 Preface

The University of Connecticut Voting Technology Research (VoTeR) Center received the data gathered in the random audits performed in the State of Connecticut following the November 2006 election. The data was submitted to the Secretary of the State Office on November 28, 2006 and conveyed to the VoTeR Center. The preliminary analysis was published on December 6, 2006. Subsequently a substantial discrepancy was found in the audit returns from the Monroe District #1. The ballots from that district were re-audited during March of 2007, and it was determined that the discrepancy was caused by a human counting error. This document reports the overall analysis results using the corrected data. This report is submitted to the Office of the State of Connecticut Secretary of the State (SOTS) and Staff.

2 Overview of the Conclusions

In the November 2006 election in the State of Connecticut the AccuVote Optical Scan tabulators (firmware version 1.96.6) were used in 125 districts. In each instant, a single machine was used. The tabulators are manufactured by Diebold Election Systems (currently Premier Election Solutions), and provided to the State by LHS Associates. After the election randomly-selected 13 districts were audited. The audit consisted of manual counting of the ballots.

This report presents the statistical analysis of the audit returns. The following (informally stated) conclusions have been reached based on the analysis:

1. The average discrepancy between the reported machine counted votes and the hand counted votes is not significant.

2. There is no significant variation in the above discrepancies based on the specific audited voting machines.

3. Machine counted totals and hand counted totals are directly proportional, however the machines have a tendency to overcount, on the average by half a vote.

4. The disagreements between the machine counted and the hand counted results are small, but not negligible in one case. Specifically, one has to be careful when relying on machine counts in multi-opening races when the total number of votes is small (presumably in such cases a recount will be triggered anyway). However, in the 2006 election there was no impact on the outcome.

We also note that in several cases the audit reports submitted to the SOTS office contained minor addition errors. In these cases, we used the reported raw values, disregarding the reported sums.

3 Introduction to the Analysis

We compute several basic *statistics* in order to draw conclusions about the accuracy of the AccuVote Optical Scan voting terminals. In this report we document our findings. The page numbers mentioned in this report refer to the 54-page fax transmission received by the VoTeR Center on

November 28, 2006, from the Secretary of the State Office [1]. The returns for Monroe District #1 were corrected following the re-audit in March of 2007.

In Table 1 we provide some descriptive information of the data.

Attribute	Size
Voting counts: total number of rows in the reports	482
Districts: total number of voting machines	13
Offices: electoral positions	27
Candidates[3]	117

Table 1: Description of the audit data (note: in two rows, among 482, the candidate fields are missing).

The data collection process is described in the initial pages of the audit report [1], and we do not repeat it here. Based on the reported audit data, we define the following symbols for the numeric fields of a record (beginning with the "Machine Totals" column):

- X – the machine totals;
- U – undisputed ballot hand count totals;
- Q – questionable ballot hand count totals; and
- Y – overall hand count totals (undisputed + questionable).

Also, for each record we define an additional quantity, $\Delta := X - Y$, as the difference between the "total machine counts" and the "overall hand counted totals".

In an ideal situation all Δ quantities are 0, signifying the machines' ability to provide the accuracy of a careful analysis done by an expert.

As a first step, we study whether there is a significant variation of the Δ quantities with respect to the districts. In order to draw conclusions of the possible effect of the voting locations (i.e., districts) we test for an ANOVA model, detailed in a later section. Based on these conclusions we carry out further analysis and aim to fit a linear-model. We present the results of the analysis in the following section, dealing the following questions:

- Is the mean of the Δ quantities really close to 0?
- Does the location (district) have any effect on Δ?
- Do the overall machine counts and hand counts agree linearly with the proportionality constant 1?
- What is the impact of the accuracy of the voting machines on election results?

4 The Results of the Analysis

4.1 Is the mean of the Δ quantities really close to 0?

Goal: The Δ quantities are essentially the discrepancy between the machine counted votes and the hand counted votes. A natural question is:

Are the machine counts above or below the hand counted results, and how significant are the differences?

Analysis. Following are some descriptive statistics of the vote counting data.

Statistics	Machine	Undisputed	Questionable	Hand	Δ
Mean	371.49	362.54	8.79	371.38	0.0290
Minimum	0	0	0	0	-29
Maximum	1750	1729	161	1747	29
Standard Dev.	438.89	430.43	13.65	440.07	5.30

Table 2: A summary of the votes counting over all the 482 counts.

In Fig. 1 we provide a histogram for Δ over all 482 rows of observation. Note that although the mean of the Δ is around 0, the histogram gives some minor indication of the left skewness of error, Δ. This indicates that the machines have a small tendency of over-counting.

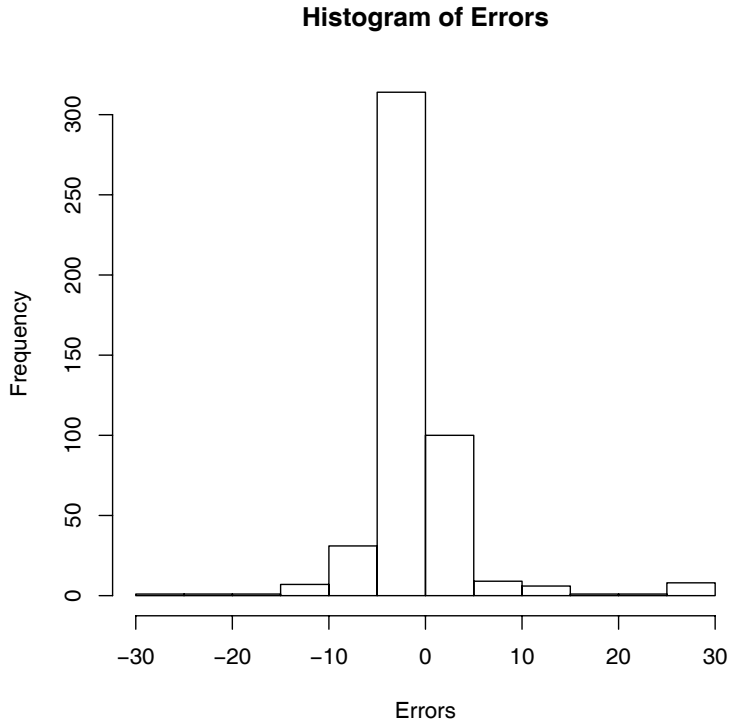

Figure 1: Histogram of the Δs from the sampled voting machines. The X-axis is the value of the Δ and the Y-axis is the frequency.

Also a $t-$statistic calculated from date for Δ indicates that the mean Δ, $\overline{\Delta} = 0$ with 95% level of confidence. We test for the following hypothesis:

$$H_0 : \bar{\Delta} = 0 \text{ vs. } H_1 : \bar{\Delta} \neq 0$$

The $t-$test statistic is calculated as:

$$t = \frac{\bar{\Delta} - \mu}{s/\sqrt{n}}$$

where $\bar{\Delta}$ is the mean of Δ, μ the estimated population mean, s is the population standard deviation, and n is the sample size. The $(1-\alpha)100\%$ confidence interval is given by

$$(\bar{\Delta} - t_{\alpha/2}s/\sqrt{n}, \bar{\Delta} + t_{\alpha/2}s/\sqrt{n})$$

where $t_{\alpha/2}$ is the value from the t-table corresponding to the specified confidence, $1-\alpha$ and $n-1$ degrees of freedom.

Variable	N	Mean	Std Dev	SE Mean	95.0 % C.I.
Δ	480	0.02905	5.29771	0.241806	(-0.4497, 0.5079)

Table 3: The results for the $t-$test for the hypothesis that the mean Δ, $\bar{\Delta} = 0$.

Conclusion 1: ***The mean*** Δ ***is not significantly different from*** 0.

4.2 Does the location (district) have any effect on Δ?

Goal: To study the variation of the Δ quantities with respect to the district where the voting machine was deployed.

Model and Analysis.

Boxplot. In Figure 2 we provide a boxplot for the Δ quantities, categorized according to the machines used in individual district locations. For each category, i.e., each district, a boxplot is provided as a means for graphical representation of the variation within a category. A boxplot contains a box, whiskers, and outliers (the starred points). A line across the box shows the median. The bottom of the box is at the first quartile (Q_1) (25 percentile) and the top is at the third quartile (Q_3) (75 percentile). The whiskers are the lines that extend from the top and bottom of the box to the adjacent values, the lowest and highest observations still inside the region defined by the lower limit $Q_1 - 1.5(Q_3 - Q_1)$ and the upper limit $Q_1 + 1.5(Q_3 - Q_1)$. Outliers are points outside the lower and upper limits, plotted with asterisks (*).

The boxplots shown in Fig. 2 does not indicate any significant difference in the spread of Δ within a category. Observe that that there are a few outliers, which we propose to ignore as noise.

One-way fixed-effects ANOVA model. To study this we used a one-way fixed-effects ANOVA model as follows:

$$\Delta_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \text{for } j = 1, \cdots, n_i; \quad i = 1, \ldots, a$$

where i denotes the district; j refers to the jth record for that district; and τ_i refers to the additive term in the model owing to the district i. There the ϵ's are assumed as $\epsilon \overset{i.i.d.}{\sim} N(0, \sigma^2)$, independent

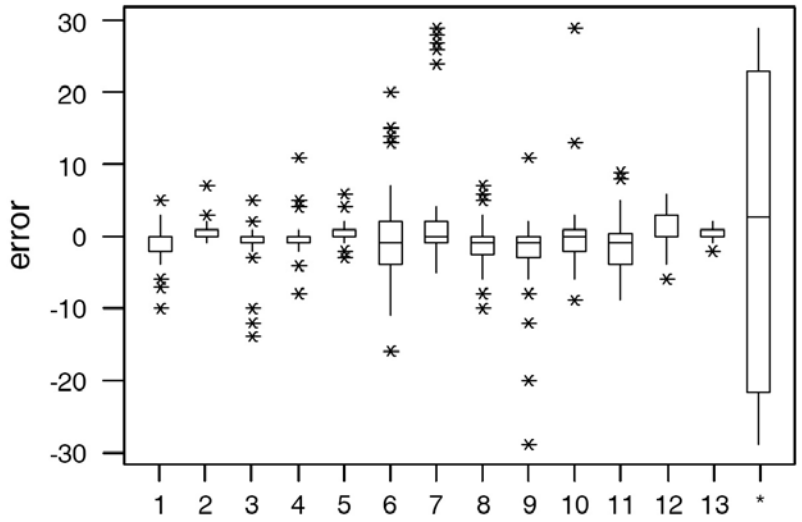

Figure 2: Boxplot of the error Δ categorized with respect to the voting machines/districts (shown as numbers 1 through 13). From the plot the variability exhibited by different categories does not seem to be significant.

and identically distributed according to the normal distribution with mean 0 and variance σ^2. It is also assumed that $\sum_i \tau_i = 0$.

We look at the hypotheses:

$$H_0 : \text{no effect of districts on } \Delta, \quad \text{vs.} \quad H_1 : \text{at least a pair of means are different}$$

From the ANOVA we get the F statistic value and we accept hypothesis H_0 with 95% level of confidence (actually fail to reject H_0).

Conclusion 2: ***There is no significant variation of the Δs due to the district (i.e., specific machine).***

4.3 Do X and Y agree linearly, with a proportionality constant of 1?

Goal: We want to know if there is a linear relationship between the Y and X, i.e., $Y_j = \beta_0 + X_j\beta_1 + \epsilon$, with $\epsilon \sim N(0, \sigma^2)$. In other words, does Y and X vary linearly? is β_0 very close to 0? and is β_1 very close to 1?

Model and Analysis. Given that earlier we concluded that there is no apparent variation in the difference $Y_j - X_j$ (i.e., Δ_i) we put the data for all voting machines (district locations) together in order to see a possible dependence model. Let N be the total number of pairs of (X_j, Y_j). We model this with a linear regression model as

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon \quad \text{where} \quad \epsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I})$$

or

$$Y_j = \beta_0 + X_j\beta_1 + \epsilon_j \quad \text{where} \quad \epsilon_j \overset{i.i.d}{\sim} N(0, \sigma^2)$$

where $\mathbf{Y} = (Y_1, \cdots Y_N)'$ is the set of responses over the entire data set, $\beta = (\beta_0, \beta_1)'$ is a vector of unknown parameters, $\mathbf{X}$ is an $N \times 3$ matrix of rank r and $\epsilon = (\epsilon_1, \cdots, \epsilon_N)'$ is an N-dimensional random vector of unobserved errors distributed as $N(\mathbf{0}, \sigma^2\mathbf{I})$, a multivariate normal distribution with the covariance matrix $\sigma^2\mathbf{I}$. The matrix $\mathbf{X}$ can be written as

$$\mathbf{X} = \begin{pmatrix} 1 & X_{11} \\ 1 & X_{21} \\ \vdots & \vdots \\ 1 & X_{N1} \end{pmatrix} \tag{1}$$

From the least squares method of estimation of β_0 and β_1 we find the following estimates:

$\beta_0 = -0.772$

$\beta_1 = 1.00$ with more than 99.99% level of confidence.

Conclusion 3: ***The offset added by the machine count is not significant, however the machine can over-count by one half vote on the average. The machine count is proportional to the hand count with a proportionality constant of unity.***

4.4 Accuracy of the voting machines and impact on election results

Goal: We want to see the effects of accuracy of the voting machines on the election results as compared to hand counting. *We say that the results due to hand counting agree with the results obtained by a voting machine, only if, the ordering of the candidates, for a particular electoral position, according to the number of votes they received, are alike.*

Model. We considered the individual decisions of each of the sampled voting machines for every distinct electoral position that appears. For example, an audit report may provide the voting counts for three electoral candidacies, say, Governor, Treasurer and State Senators, then consider them as three results by that machine.

- Total of 45 results were observed;
- The machine counted and hand counted ranking differs in 1 case;
- Hand counted had ties in ranking in 2 cases;
- The winner decided by the machine counted results and hand counted results agreed in all the 156 cases;

Table 4 below shows the case where the rankings due to machine counted and hand counted differ.

Note that given the small number of votes cast for the last two candidates the accuracy of the machine count would have impacted the outcome if this was a multi-opening race. However, a recount would be triggered if there are a very few votes that separate the last winner and the first looser of a multi-opening race.

At the moment we do not have any information on the distribution of the agreement/disagreement (machine/hand counted) results. A simple and quick way would be to assume that the agreement/disagreement (machine/hand counted) result of each electoral position is independent and

Office	Candidate	Machine	Unambiguous	Ambiguous	Manual
Comptroller	Nancy Wyman	1051	1051	6	1057
Comptroller	Cathy Cook	204	201	1	202
Comptroller	Colin Daniel Bennett	21	21	0	21
Comptroller	Richard C. Connelly, Jr	24	18	0	18

Table 4: East Hartford #3 Mayberry School district location: The case where the manual and machine counted results differs (the last two rows).

identical. A reasonable estimate (unbiased and maximum–likelihood) for the probability of disagreement, $p = \frac{2}{45} = 0.0444$. Thus, the reliability can be deduced from the binomial distribution

$$\mathbf{Pr}(D \geq x) = \sum_{i=x}^{n} \binom{n}{i} (0.044)^i 0.956^{n-i}$$

where D is the number of disagreements between the machine and hand counted results.

Conclusion 4: ***The disagreement between the machine counted and the hand counted results are small, but not negligible in one case. However, in the current election there is no impact on the outcome.***

5 Additional Notes

The reported audit returns contained several errors, most common of which were the errors in additions of hand-counted undisputed and questionable ballots. Although such errors were small, it would be of value if the audit workers were assisted in their work by a computerized tool. Note that we do not recommend that such tool is used as the "official audit system." Rather it should be used to provide sanity checks to the audit workers. This tool can also alert audit workers to the situation where the machine results differ substantially from the audit results, as was the case with Monroe #1.

References

[1] Audit Reports. Received via a facsimile transmission. 54 pages. November 28, 2006.

[2] Applied Linear Statistical Models, 4th ed., John Neter, Michael H. Kutner, Chirstopher J. Nachtsheim and William Wasserman, WCB McGraw Hill Company (1996).

[3] Graphical Methods for Data Analysis, John M. Chambers, William S. Cleveland, Beat Kleiner and Paul A. Tukey, Chapman & Hall/CRC (1983).

[end of report]