VoTeR Center

UConn Center for Voting Technology Research

PI : A. Russell, Ph.D.
Co-PIs : L. Michel, Ph.D., B Fuller, Ph.D.
Research Associates: J. Wohl, G. Johnson

Statistical Analysis of Post-Election Audit Data for the November 8, 2022 State Election

February 28, 2023

Summary

This report presents an analysis of the returns from the post-election audit performed in the State of Connecticut following the November 8, 2022 state election. The audit involved a random selection of 5% of the precincts in which ballots were cast. In each precinct so selected, the ballots cast on election day were either hand-counted or counted with electronic assistance.[1]

A total of 36 precinct audits were submitted for analysis, comprising 512 audit records. Precincts in two towns selected for audit, Greenwich and Waterbury, were later exempted as a result of a town-wide recount.

The specific goal of the analysis was to use statistical methods to detect instances of tabulator malfunction. The 512 precinct audit results evaluated show discrepancies between tabulated and audited totals that are consistent with anticipated human error in counting or ballot handling. The analysis revealed no conclusive signs of tabulator malfunction.

[1] Auditors are allowed to use an electronic audit assistance tool, provided that they visually examine each ballot to confirm bubble interpretations.

Town	District
Bristol	West Bristol School
Cheshire	District 3-2 Artsplace
Danbury	War Memorial
East Hartford	Goodwin School
East Hartford	Hockanum School
Enfield	Enfield Street School
Guilford	Guilford Fire Quarters
Hartford	A. Burr Middle School
Hartford	Mary Shepard Community Room
Harwinton	Assembly Hall
Killingly	Board of Education Central Office
Killingly	Killingly High School
Lebanon	The Fire Safety Complex - District 2 Republican
Ledyard	Juliet Long School
Manchester	Highland Park School
Meriden	Israel Putnam School
Milford	John F. Kennedy School
Montville	District 3
Naugatuck	Cross Street School - Ward 1, District, 1
New Britain	Generale Ameglio Hall
New Britain	St. John Paul II School
New Haven	Truman School - Gym
New Haven	Wexler Grant School – Cafeteria
Norwalk	Kendall Elementary School - 140A
Norwalk	Marvin Elementary School - 137A
Norwalk	Tracey Elementary - 137C
Putnam	Municipal Complex - Town Hall
Rocky Hill	Griswold Middle School
Scotland	Scotland Volunteer Community Hall
Southington	Hatton School
Stonington	District 2 - Former Pawcatuck Middle School
Stratford	Wilcoxson School 7^{th} -Dem
Stratford	Wooster Middle School 6^{th} - Dem
West Haven	Seth Haley School
Wolcott	Wakelee Elementary School
Woodbridge	Center Gym (D)

Table 1: Audit precincts analyzed

1 Analysis Description

1.1 Audit Records

The audit returns are presented in a result report in which auditors record information about the precinct under audit, the result of their count, and the corresponding count value from the tabulator. This analysis considers the vote totals for each candidate as a separate record. Each record consists of three items: the total votes as reported by the tabulator, the number of bubbles containing an "undisputed mark," and the number of bubbles containing a "questionable mark." An "undisputed mark" is a mark that covers the majority of the bubble and is dark enough that all auditors agree that it should have been read as a mark by a working tabulator. A "questionable mark" is a mark that is not large or dark enough to convince all of the auditors that a working tabulator would have recorded it as a mark.

1.2 Expected Vote Ranges

For each record, the undisputed hand-counted mark total and questionable hand-counted mark total are used to define an *expected tabulator total range*. The range is defined as having a minimum that is equal to the undisputed mark count and a maximum that is equal to the sum of the undisputed mark count and questionable mark count. If the total as reported by the tabulator is at least the undisputed mark count and no more than the sum of the undisputed and questionable mark counts, the tabulated results are consistent with the hand-counted results. In this case, the tabulator is considered to be functioning properly.

1.3 Discrepancies

Total Ballot Count Discrepancies. If the tabulator total falls outside of this expected range then it is considered an unexplained discrepancy. In general we measure unexplained discrepancies in different ways for different purposes. First, we define the *raw discrepancy* to be the signed distance between the tabulator count to the expected vote range: Specifically, if the tabulator count lies in the expected range, the raw discrepancy is defined to be zero; if the tabulator count is k larger than the maximum of the range, the raw discrepancy is k; if the tabulator count is k below the minimum of the range, the raw discrepancy is $-k$. The *relative discrepancy* is defined to be the raw discrepancy divided by the total number of ballots cast in the precinct under audit. Finally, we define *discrepancy* to be the absolute value of this relative discrepancy.

If the total ballot count is different from the total number of ballots counted during the audit, and the raw discrepancy value falls somewhere between zero and the ballot count difference, then the source of the discrepancy is potentially attributable to the difference in the ballot count. For this reason, it is important that auditors reconcile the tabulator ballot count and the audit ballot count. As mentioned above, we treat discrepancy as a percentage of the total number of cast ballots. However, in some circumstances, the number of ballots reported by the hand-count audit disagrees with the number of ballots reported by the tabulator. To be conservative, we evaluate relative discrepancy as a percentage of the minimum of these two quantities. (Observe that this convention can only increase the reported discrepancy in comparison with use of either of the individual numbers.) We call this method **Known Ballots Cast**.

In the unusual situation that the total number of tabulated or hand-counted ballots was not recorded on the audit report we instead adopt the total number of votes cast in the largest single-choice race in the district under consideration. This may lead to an underestimate of the total number of cast ballots, and hence can only increase the reported discrepancy as it is treated as a percentage of cast ballots. When this method has been used for either (or both) the total number of tabulated or hand-counted ballots, we say that discrepancies are determined by **Inferred Ballots Cast**. Note that this alternate convention is only relevant for records with nonzero discrepancy.

Anticipated Human Error. We anticipate that a small amount of error will be present in a hand count. This error presumably depends on a wide variety of factors, including the complexity of the race to be audited, the operational details of the hand counting procedure, and the physical details of the ballots themselves. The study of Goggin, Byrna, and Gilbert [GBG12] observed an empirical error rate of 1.87% (with a standard error of .678%) for Optical Scan ballots; the study adopted simple two-candidate races and averaged over several counting methods. The study also measured human miscounting of the total ballot population, observing an empirical error rate of 0.95% (with a standard error of 0.328%).

With this as a guide, we treat discrepancies of approximately 1% (or less) of the audit ballot count as consistent with errors arising from human hand counting; in particular, such error rates are not a conclusive indicator of tabulator malfunction. Historically, the majority of our observed individual discrepancies are less than 1% of the total number of cast ballots, though discrepancies tend to be higher on complicated races where voters can specify multiple candidates.

Records of Interest. We treat discrepancies exceeding 1.5% as records of special interest, and include in the report any additional information we have that may put the errors in context.

1.4 Statistical conclusions concerning the entire election

Statistical inference of significant tabulator failures. The probability of observing zero significant tabulator failures among s (independent, uniform) samples from a population of M tabulators

that in fact has fM significant failures (for a value $f \in [0,1]$) is no more than

$$(1-f)^s. \tag{1}$$

Thus, after observation of s tabulators without significant failures, the total fraction of the population of voting tabulators with significant failures is less than f with confidence $1-(1-f)^s$.

Statistical inference of global discrepancy. While these audits were not designed to offer high-confidence estimates of global discrepancy itself, some rough conclusions can be drawn. In general, we consider a setting consisting of m precincts, each using one tabulator, with ballot counts $N_1, \ldots, N_m$ and (single candidate) raw discrepancies $D_1, \ldots, D_m$. We define $N = \sum_i N_i$, $D = \sum_i D_i$, and define the *global relative discrepancy* to be D/N. We wish to estimate the global discrepancy as a result of randomly sampling the discrepancy of k precincts. For simplicity we consider sampling the precincts with replacement (in fact, our samples are drawn without replacement which slightly improves the results): this leads to k independent measured discrepancies $X_1, \ldots, X_k$, where each X_j is the raw discrepancy of a precinct selected independently and uniformly among all precincts. Then the expectation of X_j, written $\mathbb{E}[X_j]$, is equal to $(\sum_i D_i)/m$. In order to estimate global discrepancy, we consider the scaled random variables $Z_j = mX_j/(Nk)$ and the sum $Z = \sum_{i=1}^{k} Z_i$; note that the expected value $\mathbb{E}[Z]$ of this quantity is D/N, the target value. We observe that, for each j, $|Z_j| \leq m(\max_i N_i)/(Nk)$ with certainty as the discrepancy in any one precinct can be no more than the total number of ballots in that precinct; we let M denote this absolute maximum value. The classical Hoeffding bound guarantees that

$$\Pr\left[\frac{D}{N} - Z \geq \lambda\right] \leq \exp\left(\frac{-2\lambda^2}{kM^2}\right) = \exp\left(-2k\lambda^2\left(\frac{N}{m(\max_i N_i)}\right)^2\right) = \exp\left(-2k\lambda^2\left(\frac{\operatorname{avg}_i N_i}{\max_i N_i}\right)^2\right).$$

Here we use the notation $\operatorname{avg}_i N_i = N/m$ to be the average size of a precinct. In particular, for a observed value of Z we find that with confidence $1-\exp(-2\lambda^2/(kM))$ the quantity D/N is no more than $Z+\lambda$.

2 Analysis Results

The State Senate, Treasurer, and Comptroller races were selected as the focus of this audit.

Of the 512 submitted records, 318 (62.11%) of the audit records exactly confirmed the tabulator count and an additional 74 (14.45%) records represented counts within the expected range for a total of 392 (76.56%) records with 0 discrepancy. The remaining 120 audit records reported a tabulator count that differed from the audit count. Of these 120 records that differed, 119 showed a discrepancy less than or equal to 1%. The remaining record showed a discrepancy of 1.05%. We note that all of the audit reports received include both the total ballot count and the total number of ballots hand-counted at the audit. Therefore, we use the Known Ballots Cast method, outlined above, to evaluate all discrepancies in the submitted audit records.

One audit report contained a record indicating a tabulator count of 30 and an audit count of 11 for one candidate in a particular race. Although a relatively small discrepancy compared to the total number of ballots cast in that precinct, a difference of 19 votes was of concern. Upon following up with the registrars of voters in the corresponding town, the discrepancy was attributed to a transcription error in recording the tabulator count for that candidate and the audit report was updated and resubmitted.

Table 2 shows the audit record categories as well as the number of audit records that fall into that category.

Category	Record count
Records within expected range	392
Records outside expected range but with $\leq 1\%$ discrepancy:	119
Records with discrepancy approximately 1.05%	1
Total	512

Table 2: Categorization of audit records

Further discussion of discrepancies. Of the 120 records showing a discrepancy between the audit count and the tabulator count, 119 are within 1% of the audit ballot count; the remaining record is only slightly above this threshold at 1.05%. The results are therefore considered within the range of anticipated human error.

Remarks on rate of questionable marks. Considering the role played by questionable marks in the definition of discrepancy, we note the rate of questionable marks in the election. We observe 595 questionable marks over a population of 35522 ballots counted by hand and $135,258$ total cast (non-questionable) marks. The total number of questionable marks as a fraction of the total number of votes cast in the election is 0.5%.

2.1 Statistical inference concerning the election

Statistical inference of significant tabulator failures. 36 tabulators were sampled, resulting in no indication of significant tabulator error. Thus, we find that the total fraction of voting tabulators with significant errors is less than 8% with 95% confidence.

Statistical inference of global discrepancy. The 36 sampled precinct generated a sample mean raw discrepancy equal to $d = 0.86$; the estimator Z (as described above) is equal to $dm/N = 0.86 \cdot 759/1130289 = 0.0578\%$.

We calcuate the ratio $\max_i N_i / \operatorname{avg}_i N_i$ to be no more than 4.5. It follows that with 50% confidence the global relative discrepancy is no more than $0.0578\% + 45\%$.

One can obtain stronger estimates under additional assumptions on the maximum discrepancy per precinct. For example, under the assumption that a tabulator error leading to a discrepancy of over $1,000$ votes in any one precinct would have been detected by other means, with 95% confidence we conclude that the global relative discrepancy is no more than $0.0578\% + 15\%$. We remark that such weak guarantees are to be expected considering the relatively small number (36) of samples.

3 Conclusion

The University of Connecticut Center for Voting Technology Research (VoTeR Center) received data gathered in the post-election audit performed in the State of Connecticut following the November 8, 2022 state election. The audit involved the 5% of the precincts at which ballots were cast randomly selected for audits; the audit returns were conveyed by the Office of the Secretary of the State (SotS) to the VoTeR Center. The audit data analyzed by the Center contains 512 records, where each record represents information about a given candidate: date, district, office, candidate, tabulator 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 votes.

While one always wishes for no discrepancies, the magnitude of the numbers for precincts participating in the audit is consistent with anticipated human error. To conclude, the analyzed audits offer no conclusive evidence of tabulator malfunction in the 2022 state election.

References

[GBG12] Stephen N. Goggin, Michael D. Byrne and Juan E. Gilbert. Post-Election Auditing: Effects of Procedure and Ballot Type on Manual Counting Accuracy, Efficiency, and Auditor Satisfaction and Confidence. Election Law Journal: Rules, Politics, and Policy. 11(1): 36–51. March, 2012.