Considering Vote Count Distribution in Designing Election Audits

By
Jerry Lobdill
October 9, 2006
Revision 2
November 26, 2006

Following the Willie Sutton theory of bank robbery, an election bandit will always want to attack the large precincts that favor his desired loser. Whether she attacks all precincts or uses a more sophisticated algorithm that attacks only those larger than a certain size, large precincts are more attractive than small ones because that's where the votes are. This is a conclusion based on the discussion of the attacker's limits on page 22 of the Brennan report of June 2006[1].

This paper assumes that an audit is required in every county of a state in state-wide or federal elections. The county audits are designed according to the county-wide margin for each race to be audited, and the sample size is determined from the Poisson formula given in th Appendix. In designing a county audit one must determine how many precincts of a county to audit to achieve a probability of 0.99 that corruption will be detected if present. There is a formula for calculating the sample size, s, which is called the Poisson formula or election integrity formula. It requires as inputs the number of precincts involved in the county in which the audit is being conducted, N, and the minimum number of corrupt precincts, C, required to reverse the outcome of the election results for that county.

The graph in Figure 1 shows the general character of all Poisson formula[2] curves of sample size vs fraction of corrupt precincts in the total county population for a given probability of detecting corruption. The specific values given in this example are not important. Suffice it to say that the curve is always of the form $y = ax^{-b}$.

[1] The Machinery of Democracy: Protecting Elections in an Electronic World, by Brennan Center Task Force on Voting System Security, Lawrence Norden, Chair, June 28, 2006.
[2] See Appendix

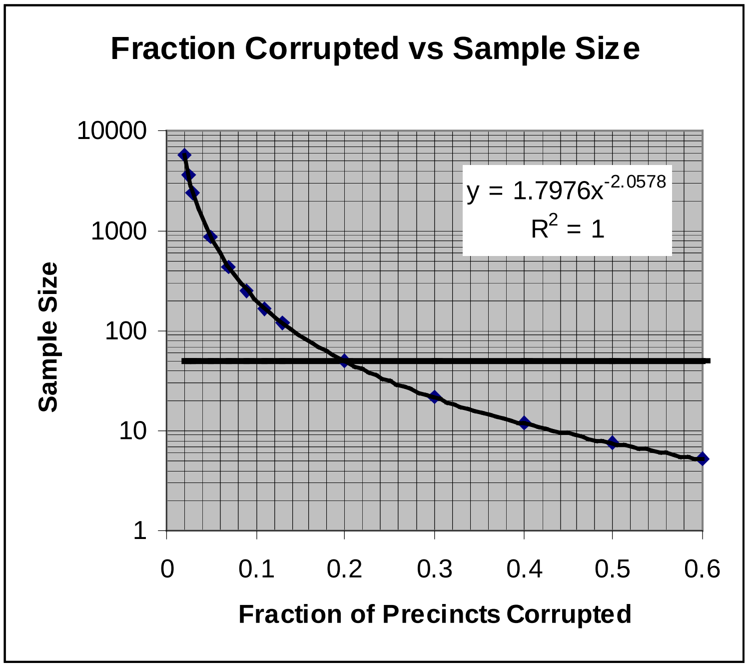

Figure 1 – Example of the Curve of Sample Size vs Fraction of Precincts Corrupted Required to Detect a Corrupt Precinct in a County of N Precincts With a Specified Probability, N = 500, Pd = 0.99

When the fraction of corrupted precincts is small, the sample size, s, is large and vice versa.

Our problem in designing an audit is to detect at least one corrupted precinct with probability 0.99 if an attack sufficiently potent to reverse the county tally has been launched in the race in question. How do we do this? Our first task is to answer the question, “What is the smallest number of precincts that can be corrupted that will produce a reversal of the county tally?” In the scheme for estimating the minimum number of corrupt precincts recently independently proposed by Lobdill[3] and Stanislevic[4], the vote counts of all precincts are sorted in descending order. Then the postulated maximum percent of the total vote count of each precinct that may be switched from the desired loser to the desired winner is calculated. The running sum of switched votes is entered in another column. The number of precincts required to make the running sum equal or exceed half the race margin plus one is the minimum number of corrupt precincts that could produce a win for the desired winner. This quantity, C, is dependent on the precinct size frequency distribution, the distribution of vote count by precinct for the two candidates, the population size, N, and the total vote margin size, M.

[3] “Designing Mandatory Election Audits”, by Jerry Lobdill 8/15/06
[4] “Random Auditing of E-Voting Systems: How Much Is Enough?”, by Howard Stanislevic, 8/16/06, p 6

The horizontal line at a sample size of 50 in Figure 1 illustrates another point about sampling. Some proposed or existing laws regarding mandatory audits specify a specific sample size of 2%, 5% or even 10% of the total number of precincts of a county involved in the race. The horizontal line happens to be a 10% requirement for this 500 precinct county. Where the exponential curve lies above the 10% line, a 10% sample is inadequate to provide a 0.99 Pd. Where the exponential curve lies below the 10% line, a 10% audit would give more than a 0.99 Pd. Figure 1 makes clear that specifying a constant percentage audit does not make mathematical sense. Even if the law says that where 10% of the precincts is larger than the calculated sample size 10% shall be audited, such a law results in needless auditing expense. The recommended law would eliminate any reference to a constant percentage to be mathematically sensible and to avoid needless expense.

In the above-cited paper by Lobdill an example from a Texas election was given. In that example there was a very skewed frequency distribution of polling place sizes. The right hand tail of the distribution was quite long and completely dominated by the incumbent (winner) (See Figure 2). The margin for the incumbent was 11.79%.

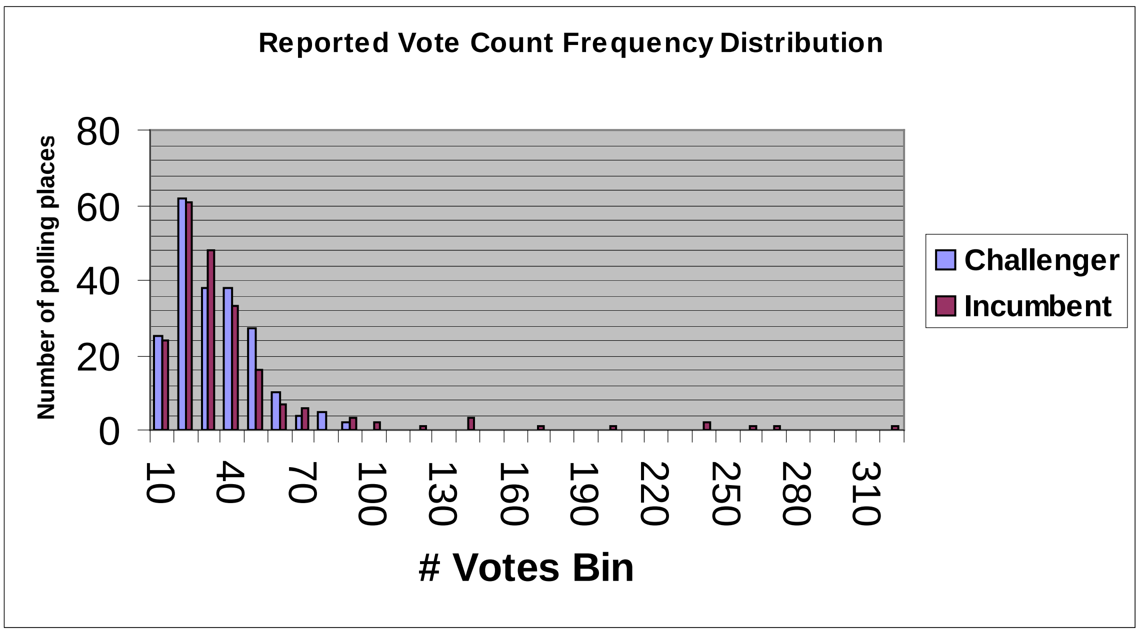

Figure 2—Distribution of Vote Counts, Texas Election 2006

As a result of the skewed distribution, a reversal of the race could be obtained by corrupting the 9 largest polling places in spite of the 11.79% margin. When this number was introduced into the Poisson formula for 0.99 probability of detecting corruption (211 polling places) the sample size required was 83. A more efficient audit plan was devised by stratifying the polling places into two strata with the largest 9 polling places in the top stratum and auditing the top stratum in descending size order, stopping the audit if and when a polling place was found to be corrupt. If corruption was not found in the top stratum the sampled lower stratum would be audited. This procedure was shown to be more efficient than auditing the 83 precincts as prescribed by the standard procedure. As it turned out, the skewed vote count distribution and the margin were instrumental in the achievement of improved efficiency.

Mark Lindeman provided a counterexample (OH CD 15, 2004 general election), Figure 3, in which the number of assumed corrupted precincts (polling places) was larger than the calculated sample size. This example demonstrated that the procedure used with the Texas election would not work for this data set.

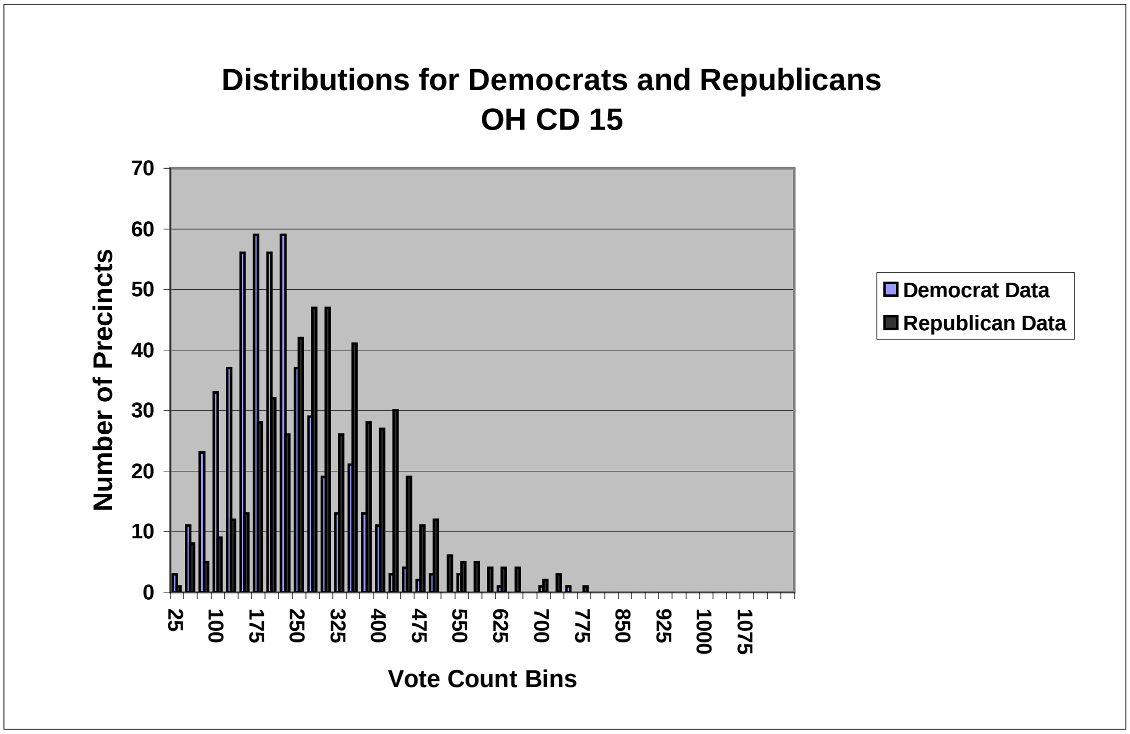

Figure 3—OH CD 15 Vote Count Distribution Data, General Election 2004

The 2004 Ohio election involved three counties, Union, Franklin, and Madison. The data combines all precincts from these counties. There were 498 precincts total. The Democrat got 100,381 votes, and the Republican got 150,918 votes. If one switches 10.06% of the total vote from the Republican to the Democrat using the PCOS 41 Trojan Horse attack of the Brennan report[5], the Democrat just wins and the resulting distribution of vote counts is given in Figure 4.

[5] as modified and described in "Myth and Reason in Designing Election Audits", by Jerry Lobdill, 9/10/06

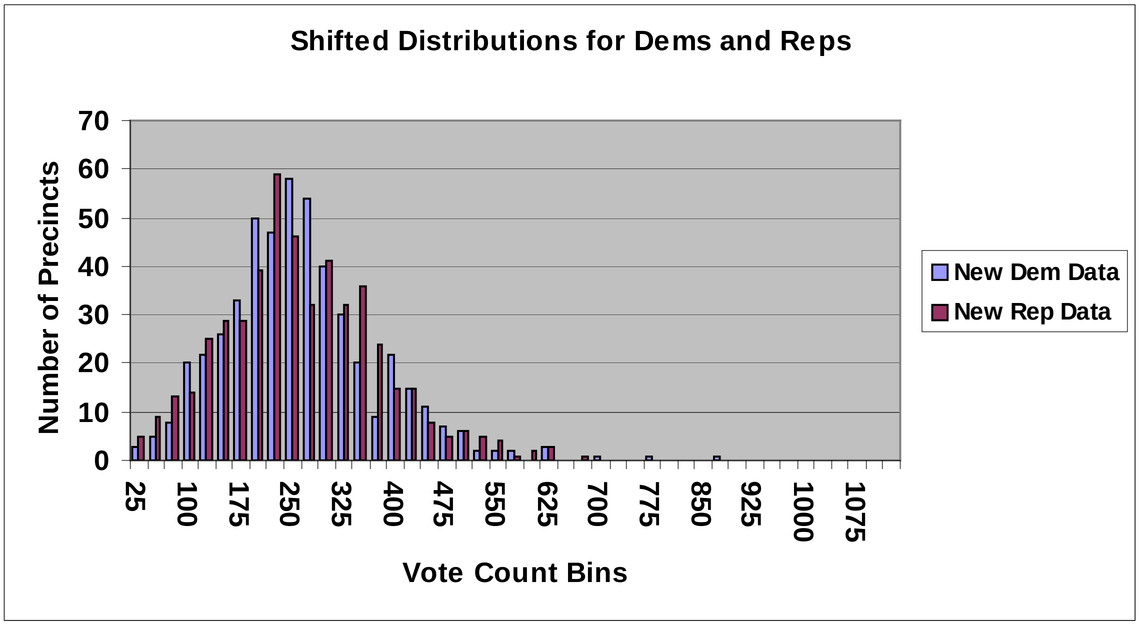

Figure 4—OH CD 15, Distribution of Vote Counts After Shifting 10.06% of the Votes to Reverse the Election

Obviously Figures 3 and 4 have distinctly different distributions. The means of the two distributions (Rep and Dem) have been visibly shifted. If we sort the precinct data by total precinct vote count, descending, and look at the running sum of switched votes we find that to effect the vote change necessary to reverse the election we would have to corrupt 496 out of 498 precincts. If we use the Poisson formula to compute the number of precincts to audit under these circumstances we find that the sample size for 0.99 probability of detecting corruption is 1.

If instead of switching 10.06% of the vote, we switch 20% of the precinct vote counts, as suggested by Stanislevic, we would need to corrupt 189 precincts out of the 498 precincts to change the winner, and the Poisson formula under that scenario yields a sample size of 10 precincts. In this case the attacker, trying to minimize the number of attacked precincts while winning by fraud, would probably set a lower threshold for vote switching high enough that most of the remaining 309 precincts would not be corrupted.

This Ohio election actually had a 20.11% margin in favor of the Republican. Such a margin might or might not be surprising depending on the historical distribution of voting preferences for the district. It is difficult to see how such a result could have been produced fraudulently if the election were held using hand counted paper ballots, because wholesale attacks on election results are not feasible with hand counted paper ballots. With E-voting, however, this is not out of the realm of feasibility.

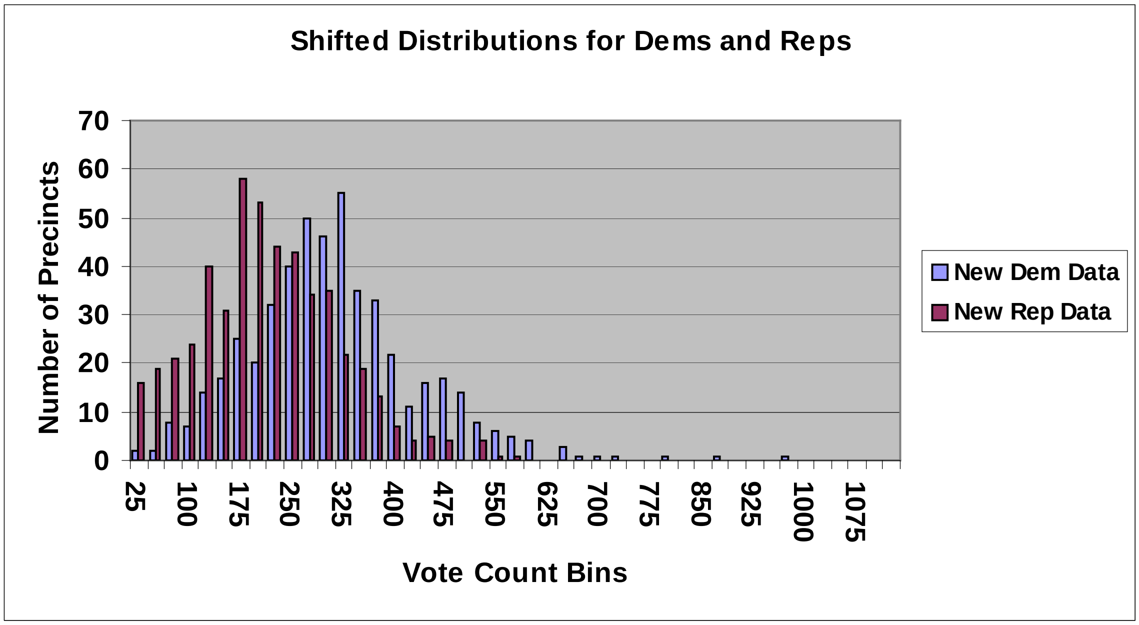

Figure 5-- OH CD 15, Distribution of Vote Counts After Shifting 20% of the Votes to Reverse the Election

Compare Figure 5 with Figure 3. The Democratic vote count distribution has been shifted to the right in Figure 5 by a huge amount. Though it is beyond the scope of this paper to suggest an upper limit on the vote switching percentage to avoid suspicion of tampering, this comparison shows a very marked difference in the distributions by Party and would probably be highly suspicious if
OH CD 15 was known to be historically heavily Republican.

Suppose an estimate of the voting preferences of the voters in this district were available by precinct. This could be obtained from county Party organization historical data or, in the case of states that require Party registration to vote in primary elections, from that data. Such estimates would provide expected voter preference distributions for comparison with the election results in Figure 3. The general shapes of the distributions should not change much from one election to another unless an area experiences high growth or loss of population.

Unfortunately, I do not have access to any such voter preference data for the 2004 OH CD 15 election. However, in the following example such data is available.

For the Multnomah County, OR 2004 presidential election we have two such estimates available, one from the 2004 election cycle (Figure 6), and one from the 2006 election cycle (Figure 7). Notice how little change there is from 2004 to 2006 in these figures. Notice the humped distribution for Republicans with a peak around 600 registered partisans and the much broader, almost uniform distribution of Democratic voters in both the 2004 and 2006 data. This is probably a slowly varying signature of county demographics in most cases except where rapid development is occurring. It is also worth noting that Figures 6 and 7 include data only on those voters who elected to register as a member of a Party for purposes of voting in the primary election.

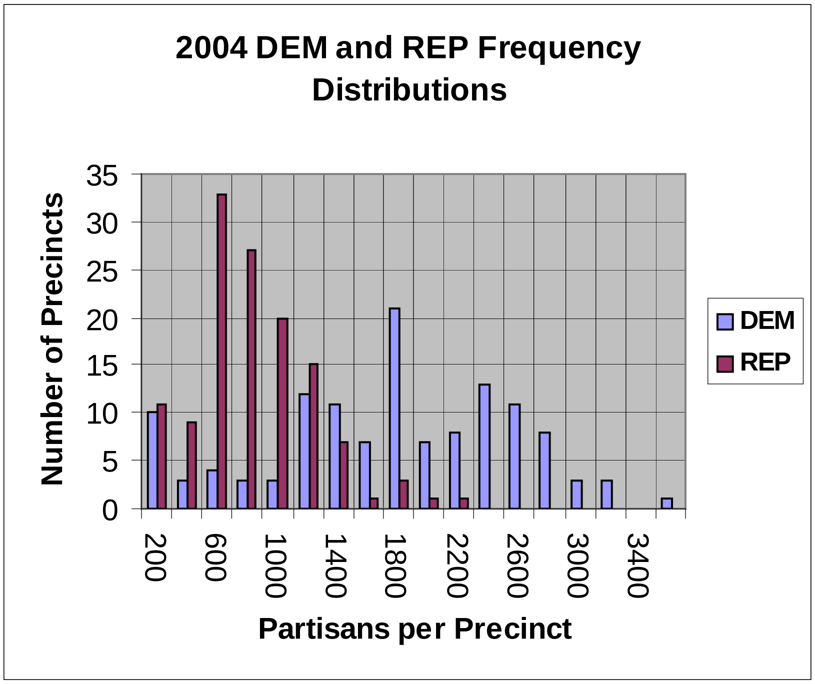

Figure 6—Multnomah County OR Party Registration Precinct Distribution 2004

Figure 7 gives the distributions of registered voters in Multnomah County in 2006.

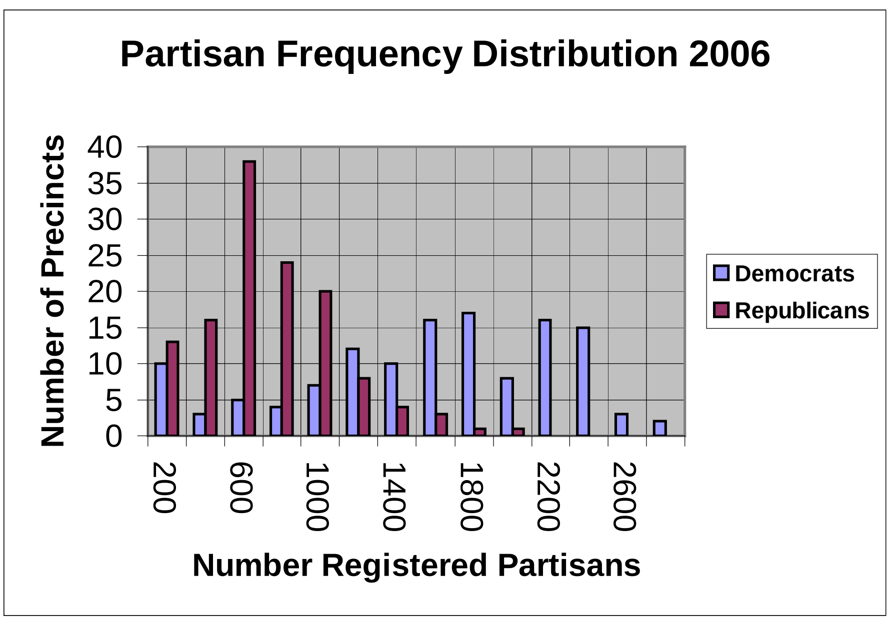

Figure 7--Multnomah County OR Party Registration Precinct Distribution 2006

Figures 6 and 7 show how little change occurred over the period from 2004 to 2006. From these charts one would expect the same general shape in vote count distributions in the 2004 and 2006 elections for this county.

Figure 8 shows the actual reported vote count distribution for Multnomah County in the 2004 presidential election. Notice the similarity between the distributions of Republicans and Democrats in comparison with Figure 6. This suggests qualitatively that the election was probably honest in this county. The total vote count was 259,585 for Kerry, and 98,439 for Bush, a margin of 45% for Kerry. It would be necessary to corrupt 126 precincts out of 128 at a switch value of 22.49% of the total precinct vote to produce a win for Bush. Of course, this amount of switching is unlikely, but if it occurred, the shifted vote count distribution would be as shown in Figure 9. The Poisson formula yields a sample size of 2 precincts for the audit in this case. Compare Figures 6 and 7 with Figure 9. The corruption is obvious just from the qualitative comparison of the distributions.

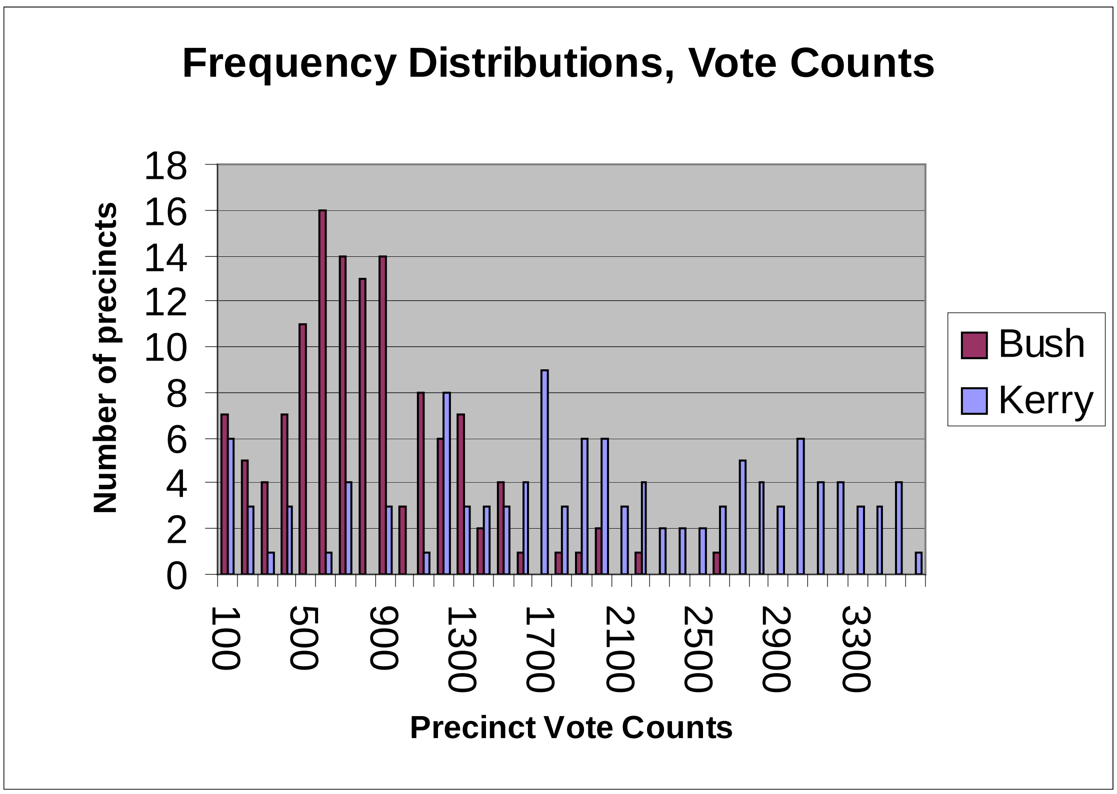

Figure 8—Multnomah County Presidential election 2004 Precinct Vote Count Distribution.

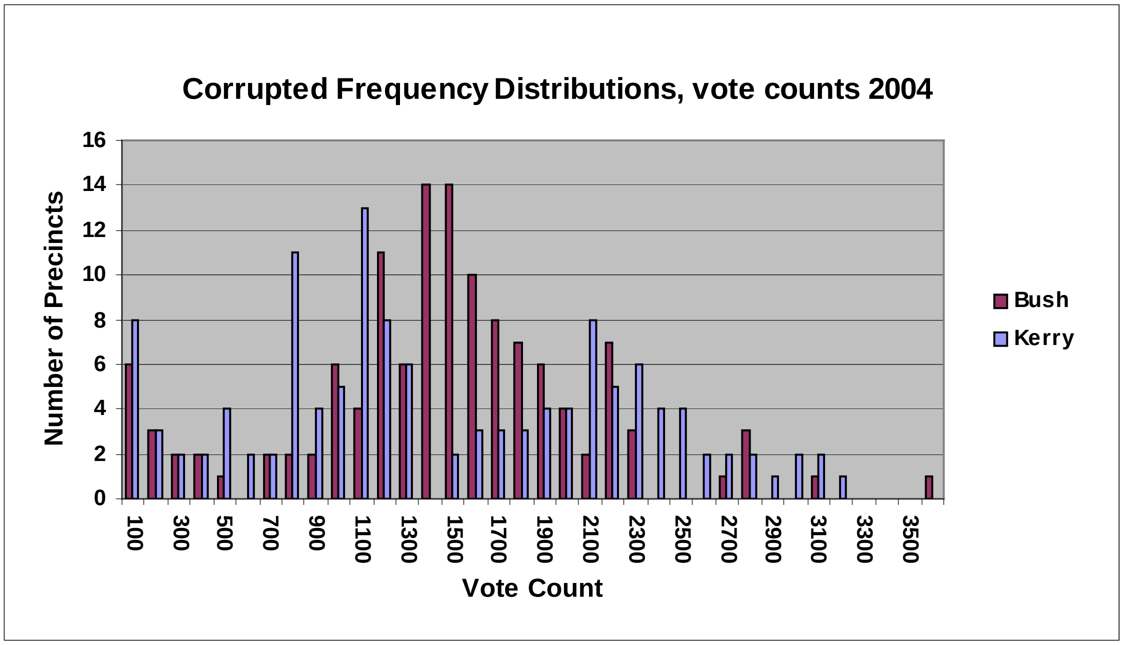

Figure 9—Multnomah County Corrupted Vote Count Distribution, 22.49% Vote Switch per Precinct.

Given the realities of the Multnomah County, OR voter preferences and distributions by precinct and the result of this election, it might seem that an audit of this election would be unnecessary and a waste of effort. However, the state-wide election for President does not turn on a single county's tally. If a mandatory audit of this election had been required, every county would be audited regardless of its margin. Though it is unlikely that an attacker would launch such an attack given the historic circumstances of a county like Multnomah, it would be easily detected. Only a few precincts (in this case 2) would need to be audited at random to precipitate a full hand recount. The OH CD 15 election is a similar situation.

The Texas 2006 election, however, shows what happens when the vote count distribution is highly skewed to the right, and the large precincts go heavily in favor of the ostensible winner. In cases of that sort, the suspicious precincts are few, but the Poisson formula suggests a random sample size that is much larger than the minimum number of corrupted precincts, C. Clearly, in such situations the use of more of the information available about attacker motivations will provide a much more efficient audit plan than simple use of the Poisson formula would provide. This topic is explored in a companion paper.[6]

Type II errors

To this point we have concerned ourselves with Type I errors; i.e., errors that result in deciding not to call a complete recount when significant fraud actually exists. But so far

[6] "Election Audit Sampling Plan Design— It's Not Just About Sampling Without Replacement", by Jerry Lobdill, October 9, 2006.

we have said nothing about Type II errors, errors where we call for a complete recount only to find that the county results were substantially correct.

Decisions about the relative importance of Type I and Type II errors are always based on risk. For example, in a submarine warfare context, the captain of a submarine may tell the sonar operator to be sure he has a real contact before reporting it to the conn. Or he may tell him to report anything that he thinks might possibly be a contact. It depends on the tactical situation and the mission. In the case of potential election fraud, at least at this time, many would say that it is most important that we not leave any doubt about the propriety of close elections. However, some, especially those who stand to benefit from the absence of recounts, argue that we must not only have a high probability of detecting fraud, but also a low probability of false alarm, (PFA), i.e., a low probability of recounting when there is actually no fraud (a Type II error).

In our audits we call a detection when we find the first precinct for which we cannot reject the hypothesis that vote switching to favor the announced winner has occurred. How many votes have to be switched in a precinct before we can make such a decision—1?, 5?, 10? Well, what is the probability that an unintended ballot definition programming error would produce such a result in only one precinct for the particular race we audit? What about any other error that results in such a miscount in only one race in one precinct?

The kinds of errors wherein a paper ballot is misread because of scanner calibration problems or the ballot is destroyed by the machine, or some other kind of calamity that causes an inaccurate ballot record are required by Federal Election Commission Standards to occur only once in a million ballots.[7] As of 2001 no voting machine system met this requirement. In Johnson County, IA, it is reported that in recounts the results are rarely off by more than 1 in 10,000, and in Dade County, FL punched card hanging chad errors occurred at a rate of 1 per 6,000.

Let's assume the worst case. If the probability of a ballot error is 1/6000 then if we have found a precinct in our sample that has K votes switched in favor of the ostensible race winner should we call for a full recount? Here we ought to consider the conditional probability that given that we have found one such error (one ballot) in a precinct what is the probability that there is another identical error involving the same race in that same precinct due to normally occurring machine failure?

I submit that the probability of false alarm is so small that we need not consider it further.

[7] http://www.cs.uiowa.edu/~jones/voting/congress.html, p 9

Summary of Conclusions

This paper explores how the shape of the vote count distribution affects computed sample size.

We have discussed the nature of the variation of computed sample size (using the Poisson hypergeometric formula) with percent corruption. This functional dependence shows that a given sample size provides better than the target probability of detection for a percent corruption greater than the value for which the sample size was calculated. This fact guarantees that if we find the smallest number of corrupt precincts that could produce a change of winner in the county, then the sample size corresponding to that number of corrupt precincts provides a 0.99 probability (or better) of finding at least one corrupt precinct in the sample for all possible distributions of corruption that would produce the change of winner. We have explored the effects of the shape of precinct distributions and the county-wide margin on the number of corrupt precincts needed to reverse the county-wide outcome. We have seen how small margins result in larger sample sizes and vice versa (Figure 1). We have observed that the distribution of Democrats and Republicans among the precincts of a county is a demographic signature that is slowly varying except in high population growth situations and that this information can be used as a qualitative indicator of possible election fraud.

Finally we have made an argument that the kinds of errors we expect to find in a corrupt precinct where votes have been switched from the reported loser to the reported winner of the county-wide vote are incompatible with errors caused by machine errors.

The author wishes to thank Mark Lindeman and Howard Stanislevic for many lengthy email interchanges in which the issues addressed here were discussed at length.

Appendix

Probability Formulas

Consider a set of n independent events, A_1, A_2, A_3, … A_n whose probabilities of occurrence are $P(A_1)$, $P(A_1)$, $P(A_2)$, … $P(A_n)$.

The probability of A_1 or A_2 or both occurring is given by $P(A_1) + P(A_2)$. (1)

The probability of A_1 and A_2 both occurring is given by $P(A_1)xP(A_2)$.. (2)

The probability that A_1 does not occur is given by 1- $P(A_1)$. (3)

The number of combinations of N things taken r at a time is given by

$$\binom{N}{r} \equiv \frac{N!}{r!(N-r)!}$$, where $N!$ is N factorial or Nx(N-1)x(N-2)...3 x 2 (4)

Poisson's Hypergeometric Formula

The famous mathematician Siméon-Denis Poisson (1781-1840) published the hypergeometric formula that is central to the election audit problem in an 1837 monograph. He had used this formula in connection with elections in France. Today we are faced with designing election audits in the United States, and we naturally turn to Poisson's work for the necessary math.

Given a total of N marbles that have one of two characteristics (say white or black marbles), if there are B black marbles and N-B white marbles, then if a random sample of S marbles is drawn from the total, the probability, P, that there are exactly b black marbles in the sample is given by Poisson's Hypergeometric Formula:

$$P(b,S,B,N) = \frac{\binom{B}{b}\binom{N-B}{S}}{\binom{N}{S}} \quad (5)$$

The probability that there are exactly zero black marbles in the sample of S marbles is computed by setting b = 0. Then the probability that there is at least one black marble in the sample, S is given by:

$$1-P(0,S,B,N) = 1-\frac{\binom{B}{0}\binom{N-B}{S}}{\binom{N}{S}} \quad (6)$$

This equation is used in election audit design to determine the sample size, S, required to assure a probability of P(0,S,B,N) that the sample contains at least 1 corrupt polling place when there are B corrupt polling places in the population, N.

Professor Ronald Rivest has published a simple approximate formula that gives sample size, S in a paper available at the following URL:

http://theory.csail.mit.edu/~rivest/Rivest-OnEstimatingTheSizeOfAStatisticalAudit.pdf

For P = 0.99, Rivest's sample size is given by

$$S = N(1 - \exp(-4.6/B)), \qquad (7)$$

where the terms are as previously defined.