Republican Primary Election 2012 Results: Amazing Statistical Anomalies

Francois Choquette, James Johnson

August 13, 2012

I. Introduction

Vote counting is usually considered impartial and accurate. Laws to protect election integrity and public servants committed to fair elections usually protect our democratic process from those who would win an election by illicit means. Election data can be analyzed using statistics to detect fraud. This report summarizes the methodology for detecting statistical data anomalies in election results, as well as providing statistical inferences for specific datasets. The method analyzes 100% of the published results, not just a sample.

When election vote count fraud is committed, data anomalies are introduced. These can be detected statistically. This report describes the statistical analysis used to detect election fraud. The results and conclusions are offered as a springboard for further analysis and investigation.

Soon after the elections, state or county Election Divisions release detailed preliminary reports with precinct-by-precinct vote counts. Once the election is certified, official “Statements of Vote” are published. For the purposes of this analysis, it is these reports that were collected and analyzed to find statistical anomalies.

We plot this data on a new type of chart that helps visualize the particular anomaly we wish to expose. These charts are named “Cumulative Vote Tally” charts and relate the candidate % success as a function of the summation of votes from small to large precincts.

The ensuing analysis was performed using official state-published data sets from the Republican Presidential Primary Elections 2012 in all 50 states. The audiences for this paper are political science and election statistics researchers that have an interest in election integrity and have a general knowledge of statistics.

II. Hypothesis Description

First, let's summarize the theory of election fraud and how it impacts the results. Altering election results leaves statistical traces (or anomalies) in the datasets. Regardless of whether the fraud is committed electronically or physically, these anomalies will appear, and are virtually impossible to mask. These are multiple ways of committing election fraud:

1. **Vote Injection** (sometimes referred to as "ballot stuffing")
2. **Vote Removal** (including damaging of unfavorable ballots)
3. **Precinct Dropping** (removal of entire unfavorable precincts)
4. **Vote Flipping** (vote exchange or swap between candidates)
5. **Result Aggregation** Aggregating results into bigger categories during reporting for the purpose of hiding election fraud impact.

Analyzing statistics of election result data sets

1. **Vote Injection** The votes for a particular candidate can be illegally injected in a particular precinct. Most likely, such an injection will not lead the vote tally (the total number of votes that was cast for all candidates) to exceed the total number of registered voters in this precinct[1]. Otherwise, fraud detection would be trivial. However, such an injection causes both the vote tally and the number of votes in favor of this particular candidate to increase at the same time. In other words, statistically speaking, the correlation between these two counts increases. Maintaining such correlation at its original level across all precincts would be virtually impossible, especially considering the fact that such activity is illegal, and thus cannot be managed perfectly.

   Along the same lines, another way of testing election data for fraud is to compute correlation between turnout and vote percent for each candidate. Moreover, these correlations can be computed not only for the overall turnout, but also for the non-rejected turnout. These methods are beyond the scope of this paper.

2. **Vote Removal** This is another method through which an election can be affected. Unfavorable ballots are simply thrown out by the perpetrators. Vote removal has recently been confirmed and prosecuted (California, Los Angeles County, city of Cudahy[2])

[1] 'Unofficial results' on county website shows turnout exceeds registered voters in precinct D001. http://www.bradblog.com/?p=5632&print=1
[2] http://www.bbvforums.org/forums/messages/132/82172.html?1342639169

a. Officials admitted to opening absentee ballots and looking at them before deciding whether to count them
b. They admit to throwing away the ballots with votes for candidates they did not want
c. They kept the ballots for candidates they did want

Again statistical techniques can be used to detect such fraud, provided that the vote removal was gross and a substantial portion of the total precinct vote tally. In this case, the statistical properties of the election results will be distorted, and the benefiting candidate will get outliers in the frequency distribution of its results. This method is also beyond the scope of this paper.

3. **Precinct Dropping** Election results may be announced before all votes are counted. This has happened in Maine in 2012[3]. There may be prejudice about announcing results prematurely, especially if there is close race with "unfavorable" predictable results in the dropped precincts. Detecting this event is relatively trivial, as long as the precinct by precinct data is available.

4. **Vote Flipping** Votes for a particular candidate can be exchanged (flipped) from another candidate in a particular precinct. This can best be done electronically. There are two advantages to perform vote flipping in precincts with larger vote tallies. First, this risky operation will have bigger impact on the final result if conducted in "large" precincts (measured by the vote tally), since more votes can be flipped per precinct, and, therefore, the number of flipped precincts can be reduced, while keeping the target total vote for this candidate fixed. Second, detection in larger precincts is more difficult. If voters detect fraud in a small precinct, they can easily get together, sign affidavits, and file a lawsuit. For example, if a precinct records only two votes for a candidate, while five friends from this precinct voted together for that candidate, the fraud becomes trivially detectable.

   Incidentally, similar line of arguments can be used to explain why the primary election results may significantly differ from the results of straw polls and conventions with the smaller number of participants. Vote flipping is very risky in the election system with open ballots. For example, West Virginia allows optional open ballots (see WV Constitution[4]).

5. **Result Aggregation** If any type of fraud is committed during the election, then aggregation of results presumably can hide this illegal act. In reality, this is not true. This aggregated set of precincts can be statistically considered as a

[3] http://online.wsj.com/article/SB10001424052970204062704577221583378430246.html
[4] http://www.legis.state.wv.us/WVCODE/WV_CON.cfm#articleIV

single precinct with large vote tally. If it happens to have higher than average percent of votes for a particular candidate, this fact is easily traceable.

Our primary observation in this paper is that a particular candidate almost always gains a higher percentage of votes in precincts with higher vote counts. See Figure 1. Note Mitt Romney's results are at 18% in small precincts and gains to nearly 25% in the final averaged results. Please note that the horizontal (X-Axis) of this chart represents the *cumulative* vote count; the rightmost points on the chart include all votes. In the larger Iowa precincts Mitt Romney's results are much higher than average, reaching as high as 54% (Polk County - West Des Moines #314.)

Other candidates, the victims of this effect exhibit the opposite slopes:

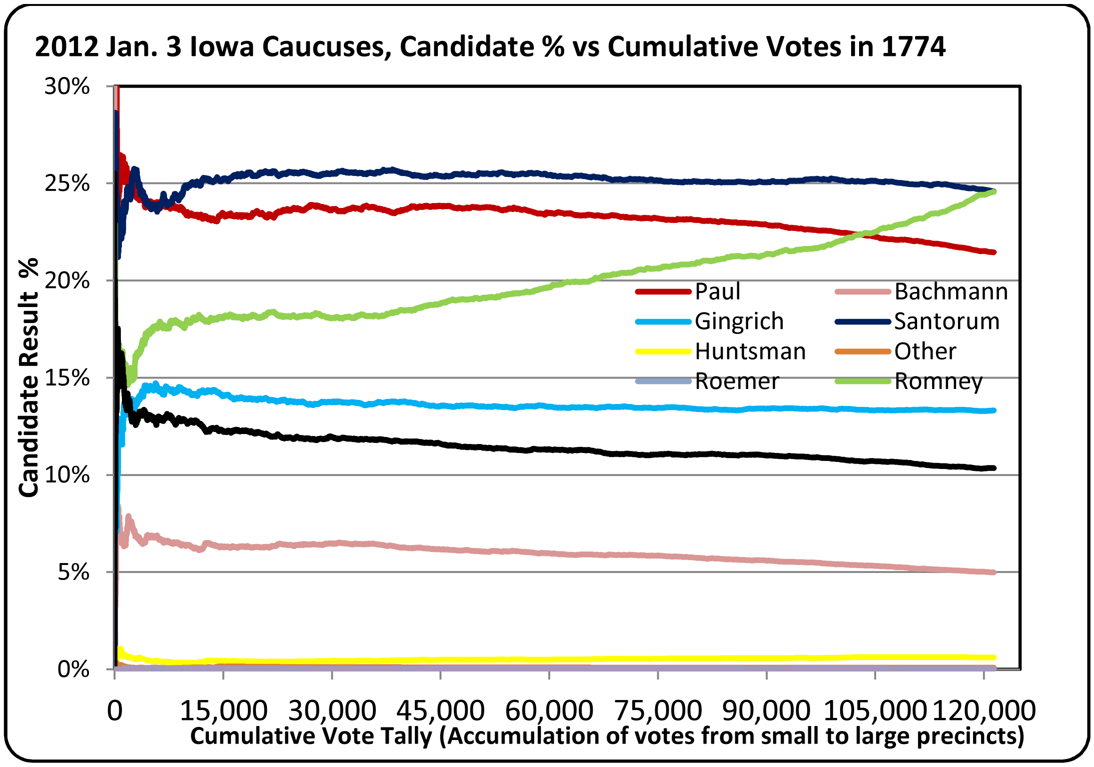

Figure 1 2012 Jan. 3 Iowa Caucuses, Candidate % vs Cumulative Votes in 1774 Precincts

At this analysis stage, this strange vote gain is only a hypothesis, but it can be rigorously tested using statistical methodology and tools.

Some argue that more liberal candidates are more popular in urban than in the rural precincts. To verify this hypothesis, we drew geographically random samples of precincts and computed partial correlations to filter out the population density factor. We will demonstrate in this paper that this factor has no impact on our conclusions.

In general, the size of the precinct (defined as the number of counted votes cast and approved) is relatively independent from the above factor by precincts' design, since more densely populated areas have more precincts, although we note that urban areas tend to have somewhat larger precincts in the number of registered voters.

Another counter-argument is that a more libertarian candidate, like Ron Paul, actually gets less “official” vote percentages in larger precincts, but again, we have seen no evidence for this argument.

This precinct-size dependent slope anomaly becomes especially compelling and conclusive because it is observed across many states and counties and favors the same candidate, regardless of its rank in the official vote count in that state.

We have sought, but have not yet found concluding factors such as demographics that would justify a particular candidate’s higher success in precincts with a higher vote tally. To verify this claim, pure demographics are charted as a function of precinct size, and the resulting chart is flat (Figure 2), indicating that demographics don’t vary appreciably as a function of precinct size.

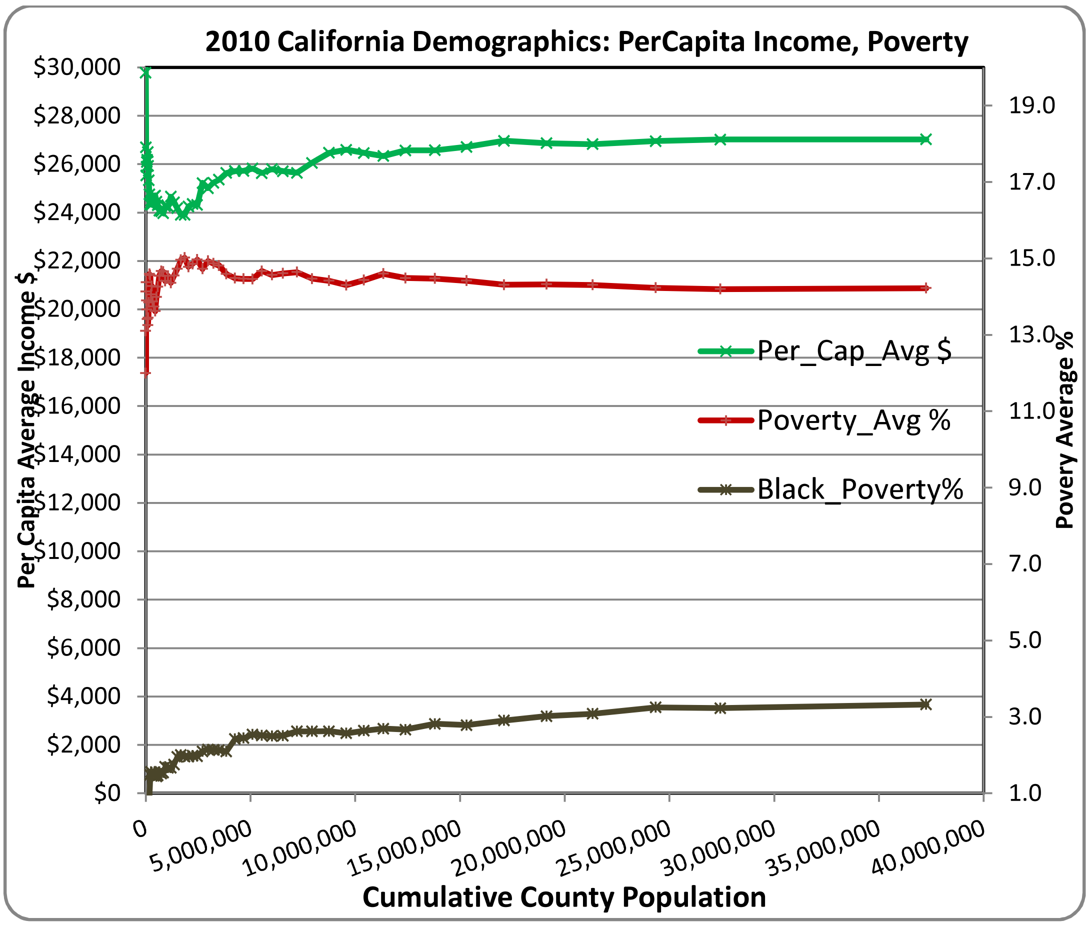

Figure 2 2010 California Demographics: Per Capita Income, Poverty5

[5] http://swdb.berkeley.edu/data.html

III Three Statistical Rules:

1. **Sampling bias.** If a hypothesis is tested 100 times with 95% confidence level, then it can be falsely rejected in roughly 5 cases out of 100. Obviously, if it is rejected 60 times out of 100, then most of these rejections are true. This rule is applicable when we need to decide whether to test election fraud hypothesis on the state level once or test it many times for each county.

2. **Power of the test.** When a hypothesis is tested on a large data sample, the false hypothesis is more likely to be correctly rejected. That's why the first way to address this issue is to test all precinct data at the state level (analyze all precincts together for the entire state), rather than county level results (county final results at the state level). However, if the magnitude of the anomaly is large, then even relatively smaller samples (like counties) will provide strong statistical evidence for its existence. These county level statistics can be computed from the precinct level data.

   Extremely abnormal test results at the county level can spur an audit in that county. Aggregating statistics of counties may lead interesting inferences on the state level. The second way to increase the power of the test is to widen the breadth of the analysis, and include all states and multiple elections with roughly the same candidates. Preferably, detailed precinct-level data should be used, and multiple statistical tests be combined. An additional method to increase detection power is to use larger significance criterion (0.05 versus 0.01). However, this increases the probability of rejecting the true hypothesis, and thus it should be used only as a last resort to improve power of the test, if necessary.

3. **Random Independent Identically Distributed Sample.** One can argue that using sorted sequences of precincts violates a random sample. Intuitively, we can expect that "small" precincts (i.e. the ones that have small vote tally) are uniformly distributed across the entire state. In other words, if one randomly selected county has 100 precincts, while another randomly selected county has 50 precincts, than at any stage of our precincts selection we will have approximately twice as many precincts from the first county as from the second one. However, this intuition is refuted by formal non-parametric tests for randomness. Without quantified randomness, we can still run hypothesis test on correlations, but we cannot use these samples to infer conclusions about the entire population of votes and then observe that such samples are very unlikely to come from the population with such "official" vote count. In spite of this caveat about randomness, this paper describes the method on how to generate the random sample from the data, meet all randomness tests, and at the same time observe the data anomaly in this random sample.

IV. Statistical Methods

The data from election result sources are converted in to tabular format for each precinct as rows with the following columns:

1. Congressional district or county or parish unique name (optional).
2. Precinct/Ward/Municipality unique name (optional).
3. Geographical location of a precinct or county: latitude and longitude (optional[6])
4. Population Density of a precinct or county: number of residents per square mile (optional[7]).
5. Total number of registered voters in the precinct (optional).
6. Number of rejected ballots (optional).
7. Total number of vote tally for each candidate in the precinct (mandatory).
8. Total number of "over votes" and "under votes" in the precinct (optional).

Step-by-Step Computations

After the data preparation stage, where the data set is defined and prepared, the following statistics need to be computed and the following tests run at either the state level (county summary results) or county level (precinct-by-precinct results):

1. Compute and compare statewide vote percentage for each candidate with the average (and median) for all counties (precincts).

2. Compute the correlation between vote tally and vote percentage for each candidate in each county using precinct-level data.

3. Compute the point estimate average of these correlations, its confidence interval, and median.

4. Hypothesis test the Pearson correlation coefficient[8] between vote tally (the sum of votes for all candidates) and votes cast for each candidate in each precinct. The null hypothesis is zero correlation for the entire state. Alternatively, the confidence intervals around correlation point estimate can be used, and the inference is made based on whether zero is in this interval or not.

[6] US Census Bureau: Census 2000 U.S. Gazetteer Files. http://www.census.gov/geo/www/gazetteer/places2k.html

[7] US Census Bureau: Population, Housing Units, Area, and Density 2010. http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_10_SF1_GCTPH1.ST05&prodType=table

[8] Alternatively, the Spearman's rank correlation coefficient can be used. It measures monotonic dependence between two variables, while the Pearson correlation coefficient measures linear dependence. We used the latter one, since we visually observed linear dependence.

5. Compute the correlation between vote tally and votes cast for each candidate correlation for the entire state using aggregate county data.

6. Compute the point estimate correlation and the confidence interval for each candidate between vote percentage from each county and the correlation between vote tallies with vote percentage at the precinct level. This statistic will show whether steeper slope causes higher vote percentage.

7. Shuffle precincts to assure randomness, and test randomness with at least 5 non-parametric tests. The shuffling involves two steps: random selection of a county, and the subsequent selection of the precinct with the smallest vote tally. A county is a set of precincts, and this set can be randomly generated from a pool of all precincts in a state.

8. Compute the maximum likelihood estimator for the point estimate of the true number of votes cast for each candidate. This population estimate is inferred from a random sample, which is incrementally drawn without replacement from the array of precincts generated in the previous step.

9. Perform a one-side hypothesis test (using cumulative distribution function of the Hypergeometric distribution) on the number of votes cast for each candidate. This test is performed for each cumulative votes sum in the arrays of precincts from step 6. If the ordering is alphabetical by precinct name, then the slope of cumulative votes flattens relative to the vote tally. This is often not the case with the random ordering from step 7.

10. Analyze whether population density has an impact on the precinct size as does vote tally. This checks whether each candidate gets more votes in urban versus rural areas. This filters out population density factor while computing partial correlation between vote tally and vote count for each candidate.

11. Compute statewide partial correlations[9] between each pair of candidates using precinct level vote percentage. If an anomaly existed and was substantial, negative partial correlations of bigger magnitude would indicate vote flipping in favor of one of two candidates. Other methods show which candidate is a victim, and which one is a beneficiary.

---

[9] http://vassarstats.net/textbook/ch3a.html
http://publib.boulder.ibm.com/infocenter/spssstat/v20r0m0/index.jsp?topic=%2Fcom.ibm.spss.statistics.help%2Falg_partial_corr_partial.htm
http://en.wikipedia.org/wiki/Partial_correlation

V. Intuition behind the Methodology

Testing the hypothesis about a presumptive election fraud that is traceable with statistical data anomalies:

1. Comparing the statewide vote percent with the county or precinct averages or medians provides a quick and simple way to detect inconsistencies in the results. Some candidates may have consistently better (or worse) results in counties (or precincts) with smaller (or bigger) vote tally across multiple states. This anomaly may be conditional upon another factor, such as the stage of the election season or the fact whether the Central Tabulator is used or not.

2. Correlation describes the linear relation between two variables. It ranges from negative 1 to positive 1. Zero Pearson correlation implies non-existence of linear (monotonic) dependency, while positive correlation means that a candidate is likely to gets more vote percentage in the precincts with larger vote tally. The "point estimate" of correlation can be computed from a sample (subset of precincts), or an exact value can be derived entire population (all precincts in a state or a county). Then, a "confidence interval" can be constructed around this point estimate. For example, let the point estimate be 0.3, and let the 95% confidence interval be between 0.2 and 0.4. It means that we statistically confirmed positive correlation.

   If we apply our hypothesis about election fraud being linked with this correlation, we can statistically confirm this fraud on the state-wide or county level. Alternatively, a two tail hypothesis test can be used, with the "null" hypothesis stating that the correlation is zero. In this case, Student-T distribution has to be used for transformed sample correlation.

   Let's summarize the steps to compute the confidence interval for the correlation:

   a) Compute arithmetic mean (average) vote tally count across all precincts, and then compute normalized ratios of vote tally counts (for each precinct) divided by this average.
   b) Compute ratios for each candidate for each precinct by dividing these individual vote counts by the vote tally count within each precinct.
   c) Compute arithmetic mean and standard deviation for all normalized vote tally counts from "a", and normalize them by subtracting this mean, and then dividing by the standard deviation for each of them.
   d) Compute arithmetic mean and standard deviation for all ratios from "b", and normalize them by subtracting this mean, and then dividing by the standard deviation for each of them.

e) Compute Pearson correlation between the vote tally statistic from "c" and all statistics from "d" for each candidate. This is the "point estimate" correlation.

f) Apply Fisher's transformation to these point estimate correlations to get approximately standard normal random variable.

g) Construct confidence interval around this standard normal random variable, and then invert that transformation (using hyperbolic tangent function) back into the interval for the correlation.

If we multiply statistics (c) and (d), and select the top tail (say, the largest 5% of these products, or the ones beyond 3 standard deviations of this population of products), then we would get a list of precincts that would be perfect candidates for an election fraud audit. These outliers would be the precincts with relatively large vote tally and relatively large vote percent deviation from an official result of a candidate. This outlier analysis is out of the scope of this paper though. By the way, these outliers and statistically significant correlations are visually apparent from the plots as well.

Computing correlation between vote tally and vote percentage for each candidate for each county produces similar results as the ones depicted in Figure 1, Figure 3 through Figure 11, but at the county level only. Instead of plotting these functions at the county level, we compute the point estimate averages of these correlations, their confidence intervals, and medians. If these numbers are analyzed across states, candidates, and elections, they can reveal suspicious data anomalies. Normally, these confidence intervals must include zero. If this is not the case, further investigation is recommended.

The following figures provide a graphical view of the problem for 11 states. These states were selected because we had precinct-level data readily available for the entire state, which allows for precise charts and statistics.

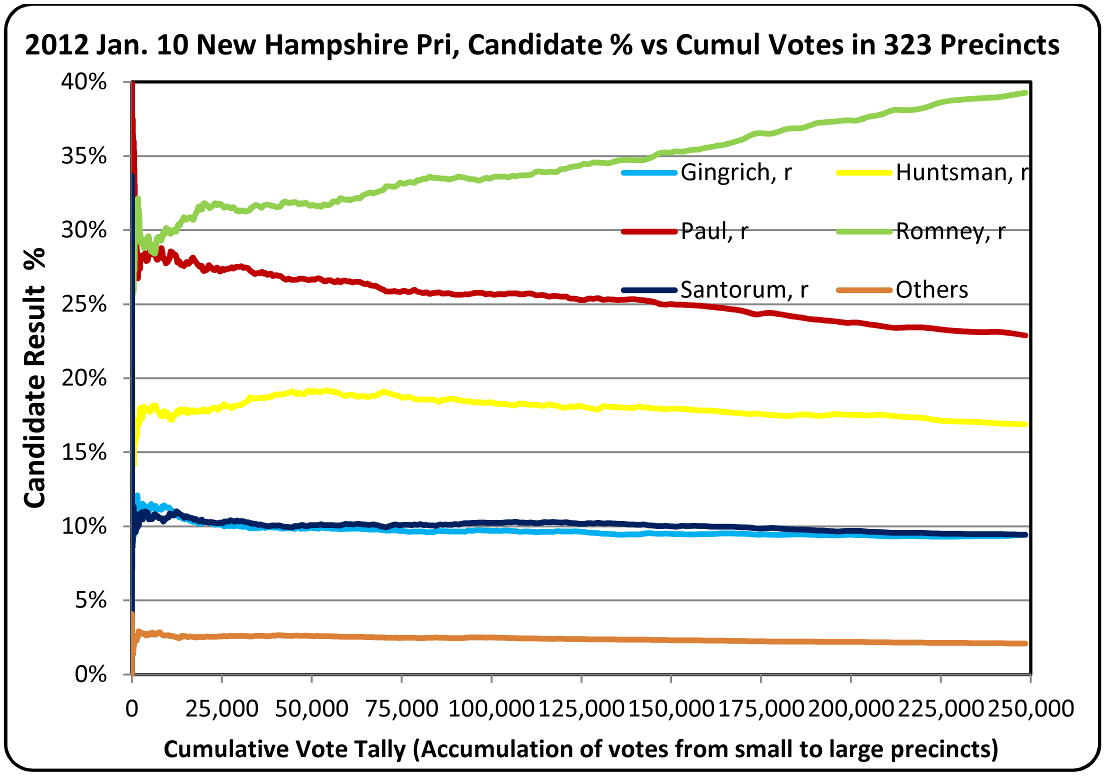

Figure 3 2012 Jan. 10 New Hampshire Pri, Candidate % vs Cumul Votes in 323 Precincts

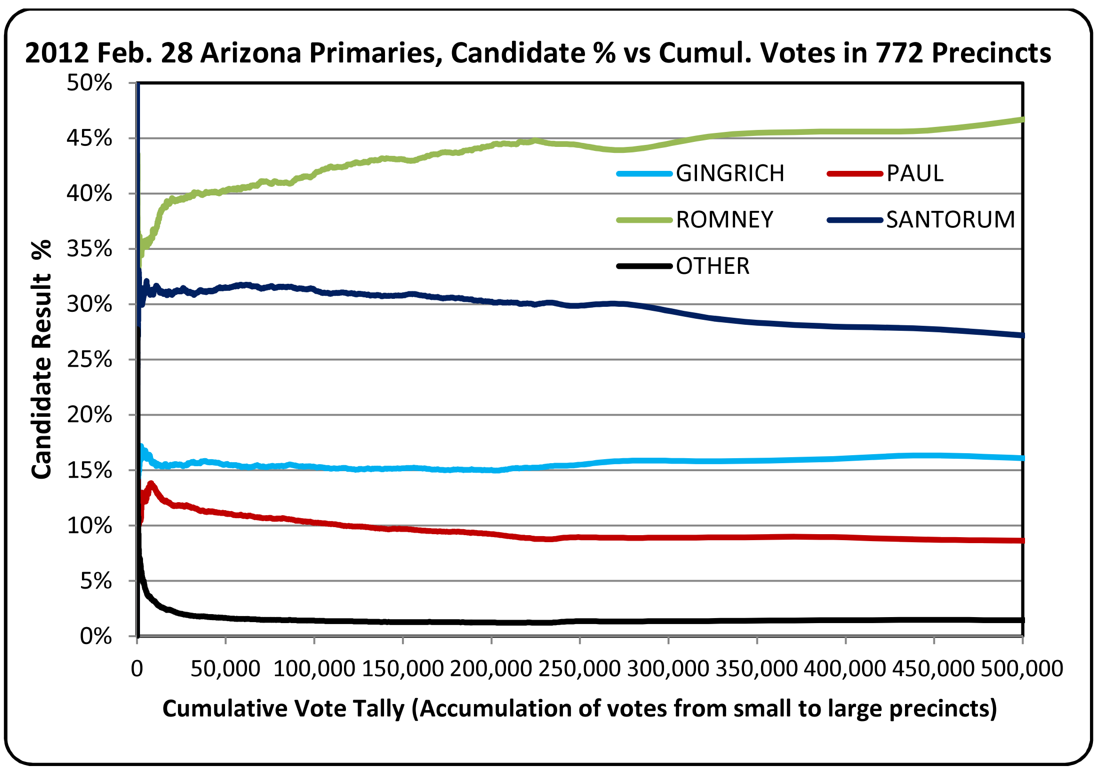

Figure 4 2012 Feb. 28 Arizona Primaries, Candidate % vs Cumul. Votes in 772 Precincts

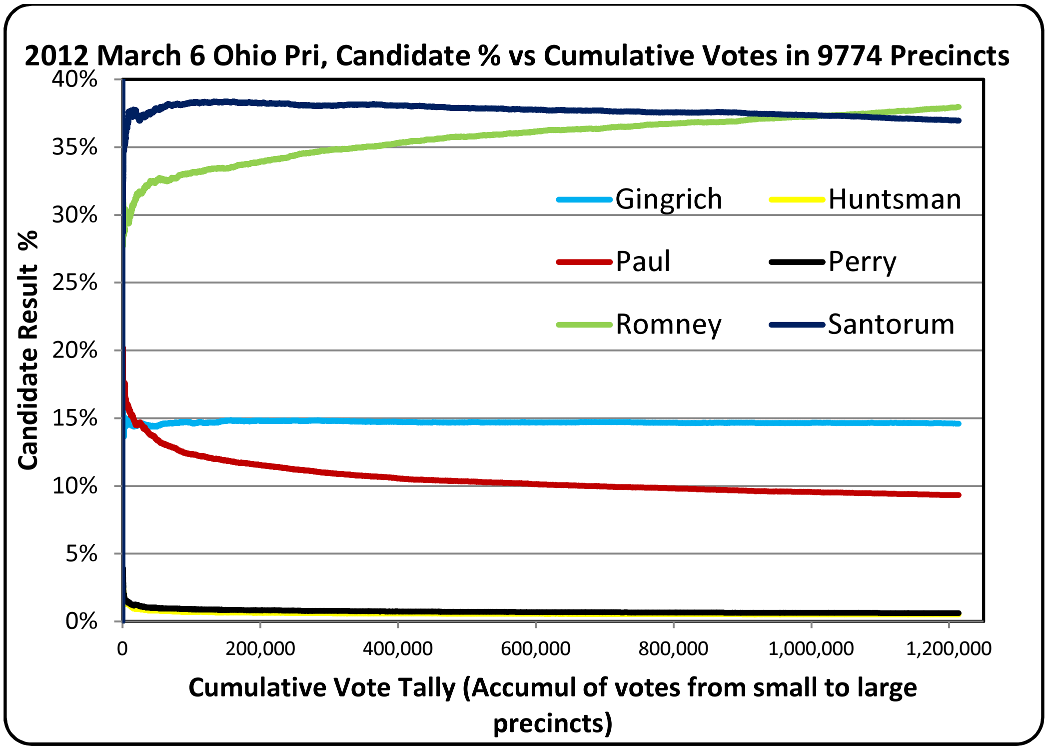

Figure 5 2012 March 6 Ohio Pri, Candidate % vs Cumulative Votes in 9774 Precincts

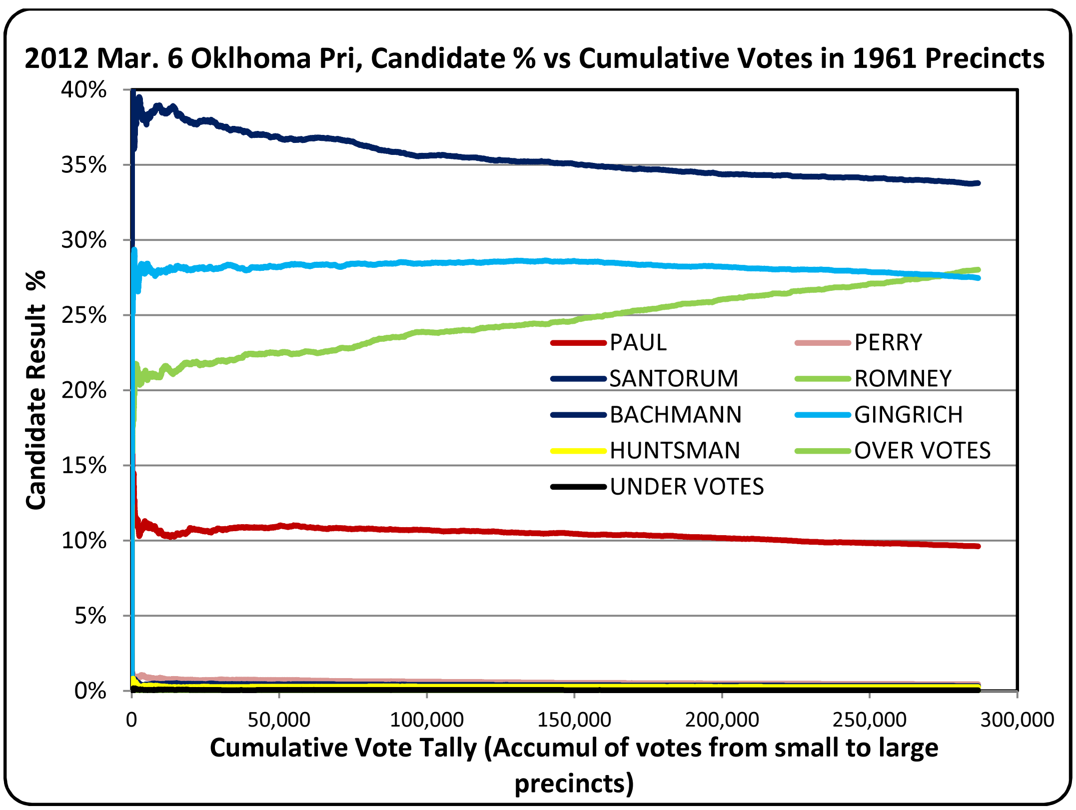

Figure 6 2012 Mar. 6 Oklhoma Pri, Candidate % vs Cumulative Votes in 1961 Precincts

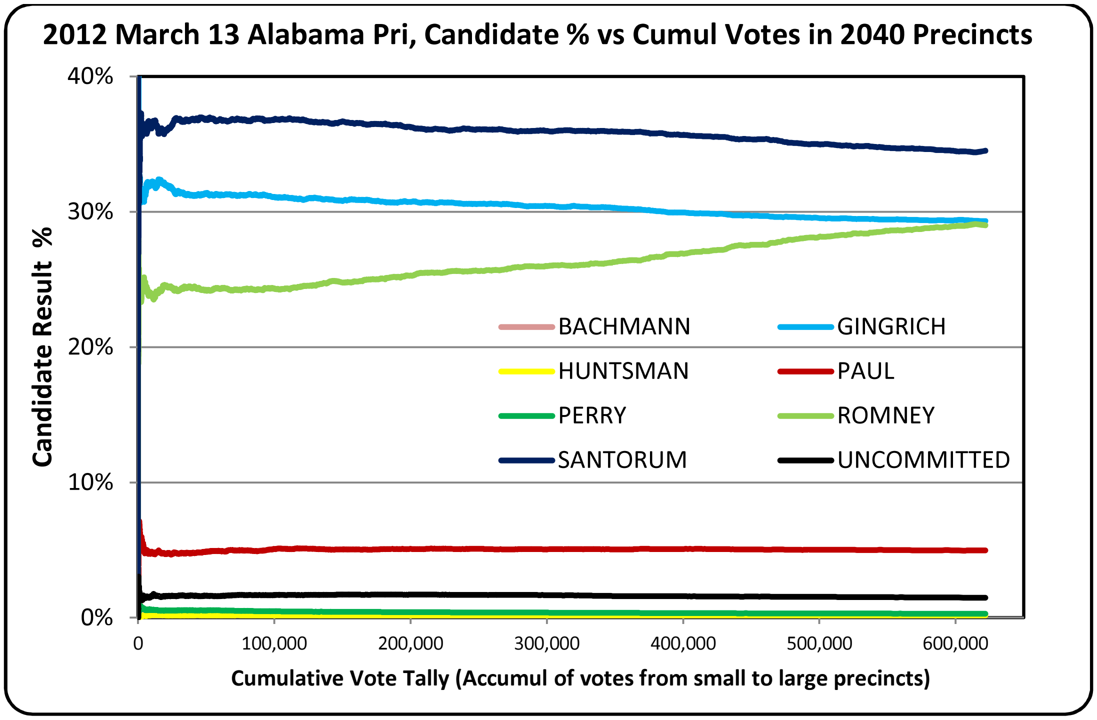

Figure 7 2012 March 13 Alabama Pri, Candidate % vs Cumul Votes in 2040 Precincts

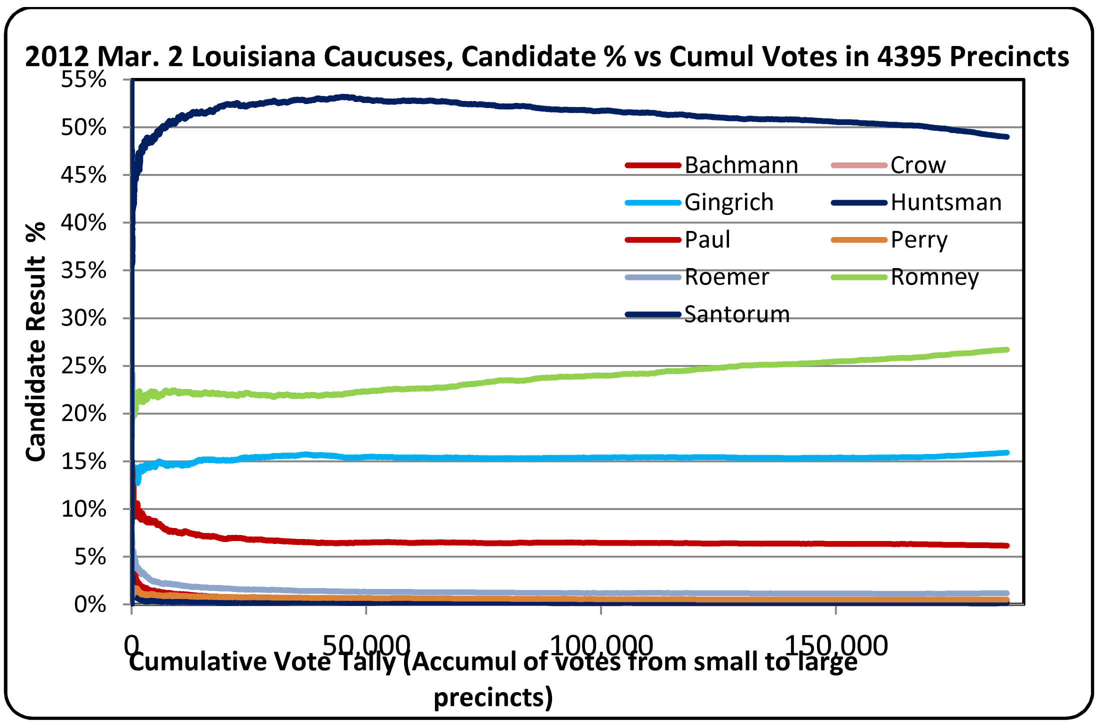

Figure 8 2012 Mar. 2 Louisiana Caucuses, Candidate % vs Cumul Votes in 4395 Precincts

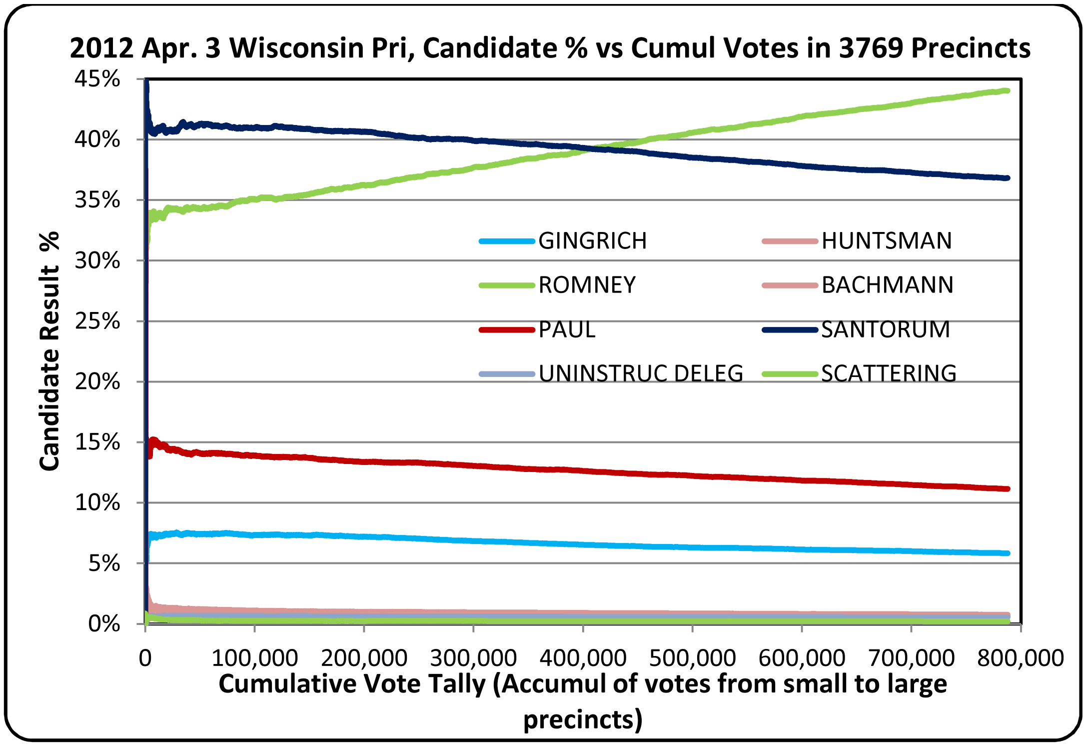

Figure 9 2012 Apr. 3 Wisconsin Pri, Candidate % vs Cumul Votes in 3769 Precincts

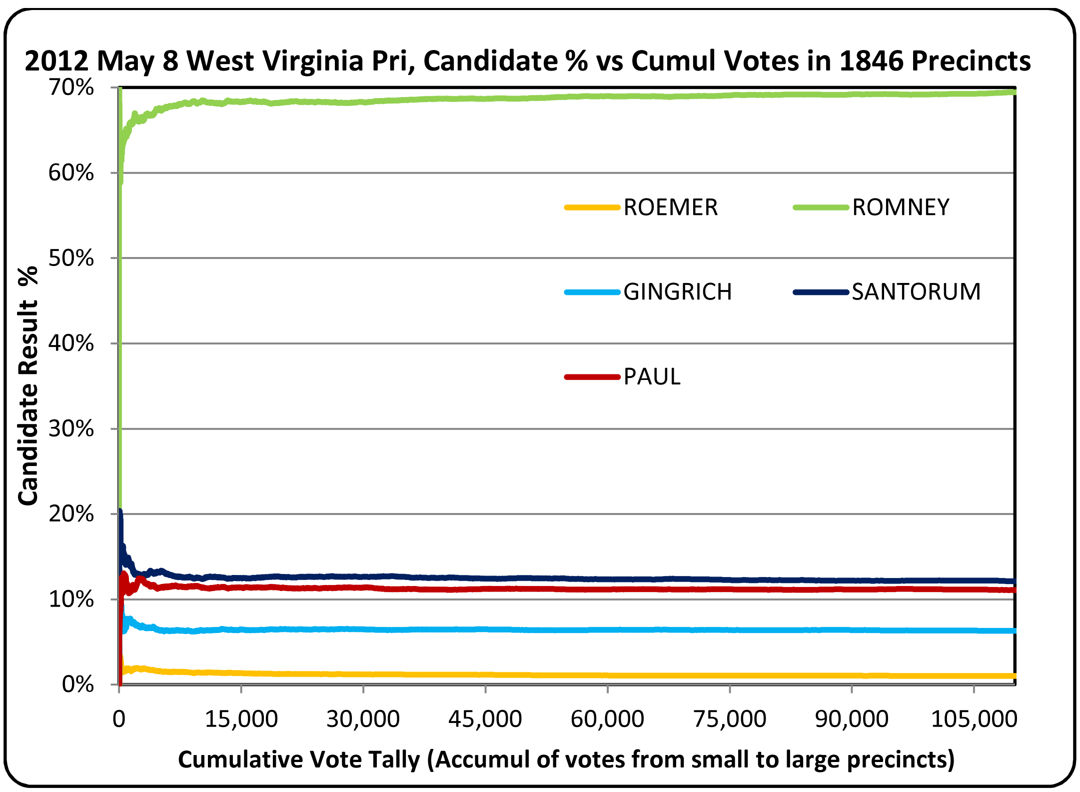

Figure 10 2012 May 8 West Virginia Pri, Candidate % vs Cumul Votes in 1846 Precincts

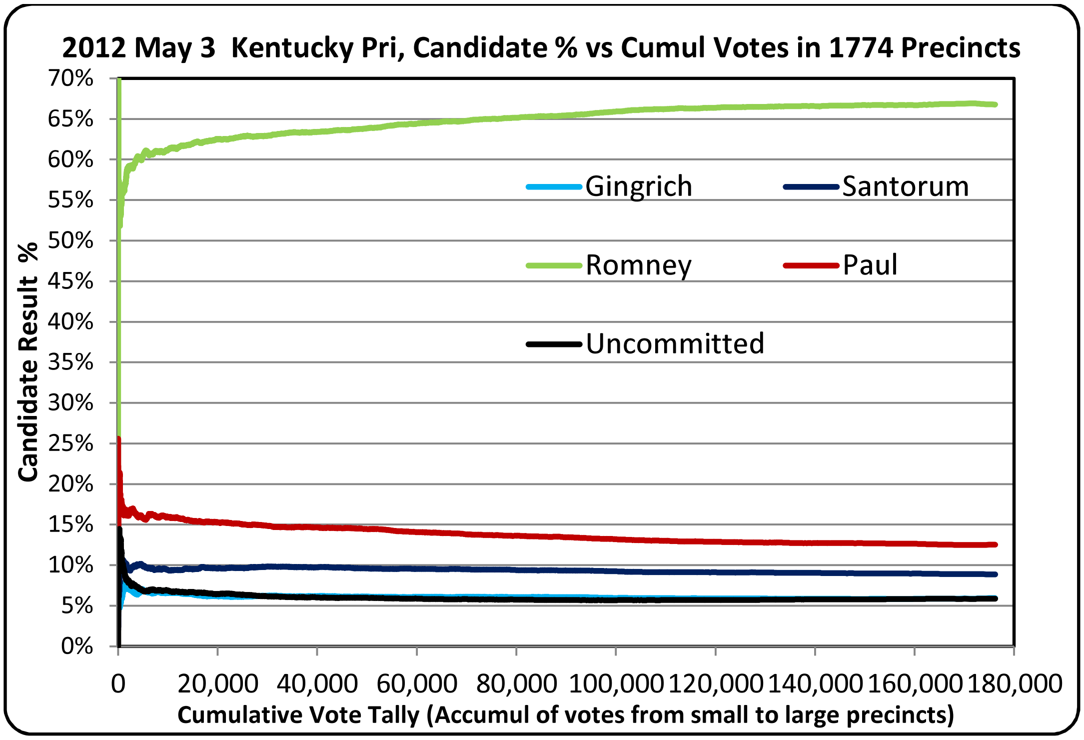

Figure 11 2012 May 3 Kentucky Pri, Candidate % vs Cumul Votes in 1774 Precincts

Table 1 shows the advantage candidate Romney gained through this effect for the above 11 states. Nationwide the total is approximately 1.2 million votes.

Table 1 Vote lost/gained in 11 states[10]

State	Rick Santorum	Newt Gingrich	Ron Paul	Mitt Romney
Alabama	-14034	-11924	956	28522
Arizona	-20682	1993	-15438	35840
Iowa	-761	-957	-2296	7850
Kentucky	-1292	-1179	-6133	10400
Louisiana	-2640	2284	-2879	8429
New Hampshire	-2966	-4122	-12881	23209
Ohio	-10774	1349	-48129	65777
Oklahoma	-12812	-1760	-2578	18382
Utah	1310	-846	-348	1310
Wisconsin	-33982	-12606	-23518	76745
West Virginia	-1032	-94	-529	2239
Totals:	-99665	-27862	-113773	278703

[10] Not all candidates are listed, causing the horizontal sums to not add up to zero.

Election fraud in favor of a candidate produces positive and bigger cross-county average correlation point estimate. Counties with extreme outliers in these averages must be thoroughly audited. If we compute correlation between vote tally and vote percentage for each candidate for each county in a state, split this set of counties into two subsets by some factor (e.g. by counties with and without electronic systems), and then run Student-T test on the population means[11], we might find statistically different average correlations between these two sub-populations. The one with the bigger correlation will indicate fraud (unless someone claims that these machines were assigned to the counties by some factor that inherently causes upslope in these selected counties). These Student-T tests are out of scope in this paper.

3. The next step is computing similar correlations between vote tally and vote percentage for each candidate on the statewide level using both precinct data and county data. Obviously, the precinct data produces more accurate and informative results, but both of these results can be compared with each other. If we detect significant correlation from the precinct data, then we should expect statistically significant correlation from the county data. Election fraud applied to the precinct data is propagated to the county data. We can trace how well the county data shows this fact.

4. Another data anomaly can be detected by computing the point estimate correlation and the confidence interval for each candidate between vote percentage from each county and the result from step 3 (the correlation between vote tally and the vote percentage on the precinct level). The intuition is as follows: Let's assume that there exists an unknown factor that justifies significant non-zero average correlation in step 3 across counties. However, this factor cannot justify the dependency between its magnitude and the vote percentage result in an average county. Meanwhile fraud can justify this dependency: the more election fraud is committed in a county, the higher (lower) the correlation, and the bigger (smaller) the vote percentage the candidate gets.

5. Inferences on correlations can be made on the entire population. However, in order to run hypothesis tests on random samples from this population and show infinitesimal probability of drawing such sample from this population with alleged properties (e.g. "official" vote count for each candidate), we need to generate such random samples, prove that they are random, and still preserve the data anomaly in these samples. The randomness of the samples is proven with the following five non-

[11] http://en.wikipedia.org/wiki/Student's_t-test

parametric tests: Median Run Test, Sign Test, Runs Up and Down Test, Mann Kendall Test, and Bartels' Rank Test. The null hypothesis for all of them is the randomness of the sample. Thus, sufficiently large p-value can statistically prove randomness of samples.

How do we randomly select precincts? Since we are looking at the events that are occurring at the same time, we can define randomness only by space. Specifically, if we have latitude and longitude for each county or precinct, then we can compute distance between each of them and an arbitrary selected origin on the map within the same state. If we draw a sub-sequence from a sequence of precincts (ordered by vote tally), then most of the above randomness tests typically reject randomness. Intuitively we draw precincts from all over the state in this sub-sequence, but they turn out to be slightly clustered in space by their vote tally. In order to randomize the sample, we need to split the selection process into two steps: random selection of a county within a state and non-random selection of a precinct within this county. The first step assigns probabilities to each county depending on its size in the number of precincts before every drawing. In other words, the cumulative mass function must be updated for each county after each random drawing without replacement. During the second step, we draw the smallest unselected precinct (by vote tally) in the selected county.

After ordering the entire population of precincts this way, we can draw a random sub-sequence, and prove its randomness with all 5 tests. Moreover, if we plot vote percentage of each candidate against cumulative vote tally of this randomly ordered sequence, *we can still detect the same slope as in the original deterministic ordering. This approach satisfies the requirement for randomness in the data sample, yet produces the desired slope, which indicates likely election fraud.*

There are two extreme cases, which can help in understanding the method. If there is only one “county” in the state, then the samples are random, but slopes will be flat. If the number of “counties” equals to the number of precincts, then the slopes will be preserved, but the sample will not be random. Thus, for a more general case, we can generate random pseudo counties of predefined size, and then run the tests. The size will define the tradeoff between randomness and preservation of the slope. Since counties are allocated in most states, we can use them in our analysis, instead of generating them first.

6. The Hypergeometric distribution allows making probabilistic expectations about random samples’ properties based on the entire population, or

inferences about the entire population based on a specific instance of a random sample. For example, suppose that the total vote tally (population) in the state is 100, and candidate A actually got 60 votes. Using the above methods, we have drawn a random sample without replacement from this population, which is akin to an exit poll (assuming that it is perfectly random and truthful). If our random sample has size 10, then the probability that 6 respondents voted for the candidate A is 0.2643, the probability that at most 6 respondents voted for the candidate A is 0.6258, and the probability that at least 6 respondents voted for the candidate A is 0.6386.

Alternatively, one can infer the number 60 (the total population vote for candidate A) from the above sample of size 10 with 6 votes for A with the following formula: floor (6 * (100 + 1) / 10). But this is just a point estimate. There are at least two methods to construct the confidence sets (similar to confidence intervals) around point estimates for the hypergeometric distribution: “test-method” and “likelihood-method”. These methods are important for analyzing the exit polls results, but they are not the focus of this report.

However, even the point estimate of the vote counts for each candidate can show potential fraud-based bias in favor of or against one or several candidates. Specifically, if the precincts are randomly ordered (as described in the previous point), and the population point estimate of vote counts keeps on increasing for one candidate while decreases or flat for other candidates, then this serves as another indication of suspicious positive correlation, but viewed from a different angle. The averages can be computed for deviations of these point estimate vote percent results for each precinct from the “official” results. These are the rough corrections of the official results towards the actual results.

7. Let’s look at the third method to detect the same data anomaly. This time we run a series of one-sided hypothesis tests on the vote percentages for each candidate in the samples. If we run these tests on the precincts that are sorted alphabetically by country and/or precinct name, then the anomaly is not detected. However, if precincts are randomly ordered as defined earlier, then the anomaly is extremely pronounced in favor of one specific and the same candidate across states and counties. The following list describes steps to reproduce this analysis:

    a. After ordering precincts randomly as defined in point 5 and running tests for randomness, compute the precinct cumulative vote count sums and vote percentages for each candidate and for the whole vote tally. You may think of these sums as incremental exit polls results. The randomness of these incremental “exit polls” is discussed earlier.

b. Run two hypothesis tests for each candidate at each ordered precinct row with the cumulative counts. We applied the test to a few percentiles: 15%, 30%, 45%, 60%, 75%, 90%, and 100% and used the hypergeometric cumulative density function[12] to run these tests.

The following example illustrates the point. Suppose that the statewide vote tally is 100, and the candidate Bad "officially" got 40 votes, while the candidate Good "officially" got 30 votes, with the rest of the votes distributed among other candidates. We do not know how many votes these two candidates actually got. Let's assume that we added up all "official" votes from 35% percent of precincts with the smallest vote tally. These precincts have only 20 votes cast, and 10 of them were for Mr. Good, while only 5 of them were for Mr. Bad. Evidently, Mr. Bad has to catch up with the remaining 65% of precincts in order to get his 40%, since he has only 25% so far. Meanwhile, we can run a hypothesis test on Mr. Bad: the "null" hypothesis is that he will eventually get at least 40 votes (40%), and an alternative hypothesis is that he will get less than 40% of votes. Since Mr. Bad has a long way to go to catch up, we will reject the "null" hypothesis (say, at 99% confidence level), and we will statistically conclude that Mr. Bad actually got in reality less than 40 votes in total. This is called upper-tail hypothesis test. We can run a lower tail hypothesis test for Mr. Good, who was a victim of this vote flipping. In this case, we will reject a "null" hypothesis (say, at 99% confidence level) that Mr. Good's vote count was less than or equal to 30. Obviously, both tests should and can be applied to both candidates.

c. Finally, we compute a maximum likelihood estimation of the statewide vote count for each candidate for each of these percentile samples: 15%, 30%, 45%, 60%, 75%, 90%, and 100%. These are random samples, but they may often exhibit substantial deviations in vote percent relative to the "official" results. The inference can be reinforced by running the test on different states, counties, and even election cycles, and observing the same anomaly over and over again for the same candidate in the current election cycle.

8. Although we draw geographically random samples for our analysis, and the timeline randomness is not required, someone may argue that we omit the possibility that one and only one of the candidates may be particularly popular in urban areas as compared with his popularity in the rural area. If this is the case, and we find out that larger precincts (as measured by vote

[12] http://www.perlmonks.org/?node_id=466599

tally) tend to flock in the rural areas, where the population density is higher, then the anomaly may be explained by the natural popularity causes, but not by election fraud.

One simple way to refute this argument is to look at the same candidate (Mitt Romney) in the same state in the same election, but 4 years ago. In fact, the same candidate did not have this density-related slope factor in 2008 in Maryland or in 2012 in Utah (see Figure 12 and Figure 13), while he had it in 2012 in the same state in the same type of elections

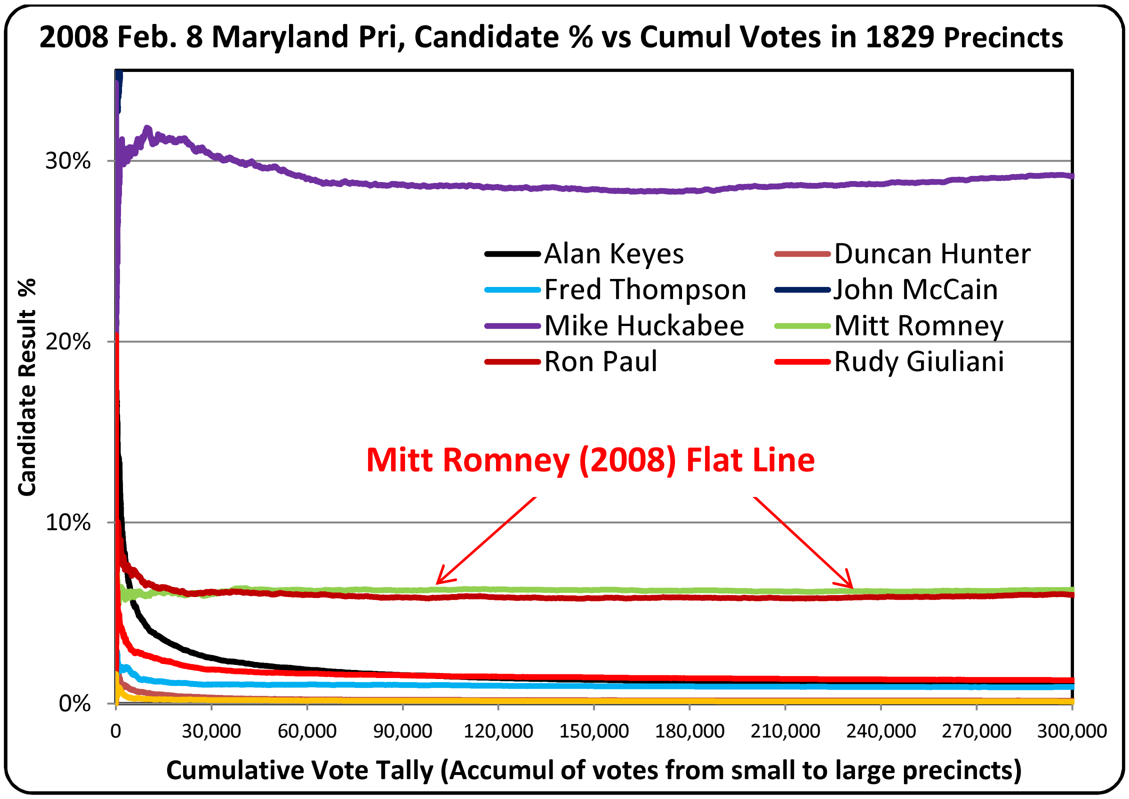

Figure 12 2008 Feb. 8 Maryland Pri, Candidate % vs Cumul Votes in 1829 Precincts

Utah in particular, with a strong demographic of 62% LDS statewide, but with less in cities, should exhibit the sharpest slopes if demographics were a causal factor. Clearly, Utah 2012 flat-lines indicating no vote flipping or demographic effect.

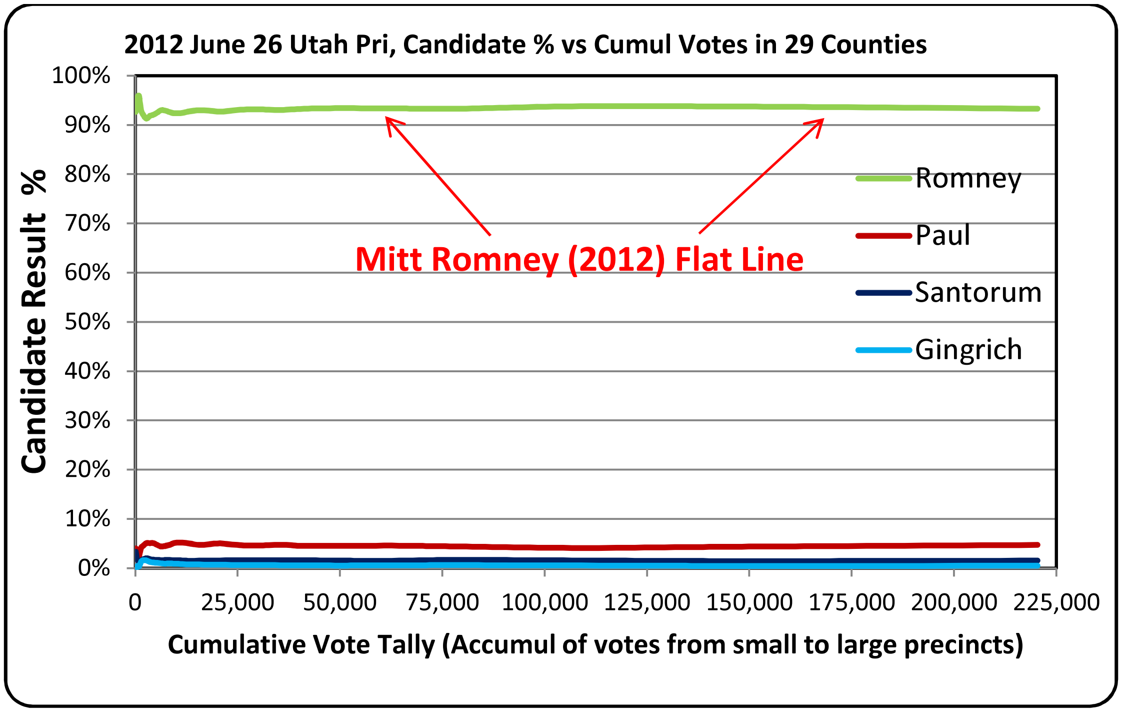

Figure 13 2012 June 26 Utah Pri, Candidate % vs Cumul Votes in 29 Counties

Because we occasionally find very flat results in different states, demographic explanations lose their credibility. Furthermore in nearly all elections not involving Republican candidates we get very flat results, it is therefore fair to assume that something is amiss with elections involving this particular candidate. We have confirmed flat-line results for Democratic, Green, Libertarian and Peace & Freedom parties in the 2008 and 2012 elections.

We employ additional methods to further refute the demographic argument. First, we compute correlations between precinct population densities and their vote tallies. They tend to be positive. Although urban areas have more precincts, but their sizes (measured by the number of registered voters) tend to bigger than the ones in the rural areas. Second, we re-compute correlations between precinct vote tally and candidates' vote count. We observe the data anomaly at this stage. Third, we compute correlations between candidates' vote count (not vote percent!) and precinct population density. Typically, this correlation is positive, since higher density is positively correlated with precinct size. However, if it is bigger for one candidate than others, than this candidate is more popular in urban areas than the others. It does not necessarily mean that he gets more vote percentage in urban areas than other candidates[13]. But we are not seeking this conclusion. We just need to know whether the candidate has more supporters per voter in urban areas than he has in rural areas. If the suspected beneficiary of vote flipping actually has smaller support density in urban versus rural than the victim of vote flipping, then can confirm that we are facing election fraud in this case.

Finally, we can compute partial correlations between precinct vote tally and candidates' vote count by making the population density a control variable. This way we filter out population density factor. If we still observe data anomaly after this correction, then the argument is fully refuted. In fact, we observe very minor impact of the population density factor on the fraud-attributed data anomaly.

9. Vote flipping can be detected by computing partial correlations between each pair of candidates. Partial correlations allow removing other noise factors from the analysis, i.e. assuming that the selected candidates are uncorrelated with the other candidates. All other remaining candidates' factors will be filtered out for each pairwise dependency between each pair. Suppose that we have 5 candidates (A, B, C, D, and E), and we want to compute partial correlation between candidates A and B. If there was flipping between C and D, then this effect will be dropped. If there was flipping between C and A, then this effect will be dropped too. However, if there was flipping between A and B, then it will be emphasized by dropping any natural noise between A (B) and C (D, E). Obviously, the

13 Suppose that we have two candidates: C1 and C2. and we have two precincts: P1 and P1.
Vote tally in P1 is 100, and this precinct is rural with low density; Vote tally in P2 is 200, and this precinct is urban with high density.
Candidate C1 got 40 votes in P1 and 80 votes in P2.; Candidate C2 got 20 votes in P1 and 60 votes in P2.
Candidate C1 is equally popular in urban precincts (80/200 = 40%) as in rural precincts (40/100 = 40%).
Candidate C2 is more popular in urban precincts (60/200 = 30%) than in rural precincts (20/100 = 20%).
So, we conclude that C2 is exposed to the "urban" factor more. It means that C2 may have a positive slope, while C1 is not supposed to have one.
The total percent of C1 versus C2 DOES NOT matter: (40 + 80) / 300 = 40% versus (20 + 60) / 300 = 26.67%. Only SLOPE matters. If there is no fraud, then C1 may have more votes in the urban areas relative to C2, but C1 cannot have steeper slope relative to its own results than C2 has.

natural noise between A and B will remain. However, imagine that all 5 candidates participated in debates, and candidate A was very successful in gaining supporters from the other candidates. There is no reason for only B's supporters to migrate to A in bulk, since A is expected to gain supports from B, C, D, and E relatively equally. If for some reason A gains supports only from B, then this rule must be applied to all states, since presidential debates are broadcast nationally. Instead, we observed what only one and the same candidates is the leader in these negative partial correlations across all states, and this anomalous leader is couple with different candidates from one state to another. This observation is attributable to vote flipping, and it is reinforced by other methods of fraud detection.

VI. Statistical Analysis of Real Election Results

The previous section describes methods, which we applied to 11 states (see Figure 1, 3 through Figure 11). Table 2 contains the statewide results (column 3) for all candidates (column 2) with at least 4% of the votes in each state (column 1). Columns 4 and 5 contain average and median vote percent across all counties in each state for each candidate, while columns 6 and 7 contain average and median vote percent across all precincts. Items highlighted in red deserve serious attention. Items in rose indicate a clear transfer of votes between candidates. When both bounds are either above zero or below zero the anomaly is detected.

Table 2 Candidate votes statistics

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
State	**Candidate**	**State Votes %**	**County Vote Average %**	**County Vote Median %**	**Precinct Vote Average %**	**Precinct Vote Median %**	**County Lower Bound Slope Average**	**County Slope Average**	**County Upper Bound Slope Average**
Iowa	Santorum	24.59%	26.56%	24.98%	24.96%	23.93%	-0.0717	-0.0081	0.0556
Iowa	Romney	24.56%	20.05%	19.18%	20.56%	19.76%	0.1548	0.2147	0.2745
Iowa	Paul	21.45%	20.40%	19.21%	23.12%	21.21%	-0.1768	-0.1120	-0.0472
Iowa	Gingrich	13.32%	14.09%	13.85%	13.49%	12.35%	-0.0034	0.0627	0.1287
Iowa	Perry	10.35%	12.15%	11.57%	11.38%	9.42%	-0.1307	-0.0656	-0.0006
Iowa	Bachmann	4.98%	6.20%	5.59%	5.80%	4.44%	-0.0950	-0.0292	0.0366
New Hampshire	Romney	39.28%	35.55%	34.25%	34.50%	33.63%	0.2264	0.3136	0.4009
New Hampshire	Paul	22.89%	24.37%	24.24%	25.71%	25.52%	-0.4317	-0.3319	-0.2322
New Hampshire	Huntsman	16.89%	17.76%	16.11%	17.76%	16.61%	-0.0740	0.0405	0.1550
New Hampshire	Santorum	9.43%	10.00%	9.96%	9.97%	9.69%	-0.1234	-0.0466	0.0302
New Hampshire	Gingrich	9.43%	10.06%	9.70%	9.71%	9.32%	-0.0804	0.0812	0.2427
Arizona	Romney	46.87%	44.49%	40.31%	41.49%	40.55%	0.1523	0.3183	0.4844
Arizona	Santorum	27.05%	29.20%	30.89%	30.15%	30.77%	-0.3104	-0.1170	0.0765
Arizona	Gingrich	16.02%	16.23%	17.21%	15.35%	15.02%	-0.0936	0.0710	0.2357
Arizona	Paul	8.61%	8.55%	8.49%	10.29%	9.12%	-0.4194	-0.2141	-0.0089
Ohio	Romney	37.96%	32.81%	31.75%	36.00%	35.14%	0.1579	0.2020	0.2461
Ohio	Santorum	36.95%	41.20%	41.41%	37.28%	37.10%	-0.0487	-0.0080	0.0328
Ohio	Gingrich	14.60%	14.76%	14.33%	14.65%	14.13%	-0.1097	-0.0744	-0.0392
Ohio	Paul	9.33%	9.75%	9.35%	10.63%	9.26%	-0.2558	-0.2179	-0.1801
Oklahoma	Santorum	33.78%	36.53%	36.29%	35.73%	35.00%	-0.1553	-0.0946	-0.0339

Oklahoma	Romney	28.03%	24.05%	23.81%	24.43%	23.81%	0.1233	0.1865	0.2497
Oklahoma	Gingrich	27.46%	28.35%	28.63%	27.96%	27.78%	-0.0636	0.0003	0.0642
Oklahoma	Paul	9.63%	9.64%	9.22%	10.50%	9.30%	-0.0952	-0.0320	0.0312
Alabama	Santorum	34.50%	35.31%	34.05%	35.65%	35.29%	-0.1116	-0.0591	-0.0066
Alabama	Gingrich	29.30%	32.00%	31.47%	30.80%	29.80%	-0.0734	-0.0234	0.0265
Alabama	Romney	28.99%	26.02%	25.25%	25.98%	25.00%	0.1318	0.1897	0.2477
Alabama	Paul	4.97%	4.35%	4.48%	5.01%	4.39%	-0.0797	-0.0310	0.0178
Louisiana	Santorum	48.99%	52.49%	52.51%	50.32%	50.00%	-0.0630	-0.0248	0.0134
Louisiana	Romney	26.69%	23.56%	23.44%	23.54%	22.22%	0.0438	0.0883	0.1328
Louisiana	Gingrich	15.91%	16.01%	16.07%	15.07%	13.89%	0.0446	0.0764	0.1082
Louisiana	Paul	6.15%	5.52%	5.11%	7.16%	4.35%	-0.0419	0.0032	0.0482
Wisconsin	Romney	44.03%	38.48%	36.59%	38.91%	37.88%	0.1531	0.2045	0.2559
Wisconsin	Santorum	36.83%	39.69%	39.75%	39.21%	38.64%	-0.1733	-0.1229	-0.0726
Wisconsin	Paul	11.15%	12.05%	11.96%	12.74%	11.51%	-0.1465	-0.0964	-0.0463
Wisconsin	Gingrich	5.84%	7.34%	7.20%	6.57%	6.00%	-0.0170	0.0328	0.0827
West Virginia	Romney	69.56%	68.94%	68.77%	68.74%	69.23%	0.0102	0.0740	0.1377
West Virginia	Santorum	12.09%	12.65%	12.24%	12.48%	11.76%	-0.0761	-0.0034	0.0692
West Virginia	Paul	11.04%	10.60%	10.68%	11.16%	10.53%	-0.1325	-0.0604	0.0117
West Virginia	Gingrich	6.29%	6.69%	6.72%	6.44%	5.88%	0.0014	0.0490	0.0967
Kentucky	Romney	66.77%	64.74%	65.10%	64.35%	65.82%	0.0736	0.1228	0.1720
Kentucky	Paul	12.53%	13.32%	13.03%	14.00%	12.00%	-0.1247	-0.0786	-0.0326
Kentucky	Santorum	8.87%	9.52%	9.63%	9.28%	8.31%	-0.0218	0.0272	0.0761
Kentucky	Gingrich	5.95%	6.53%	6.44%	6.11%	5.13%	-0.0527	-0.0065	0.0396
Kentucky	Uncommitt	5.88%	5.88%	5.71%	6.26%	4.76%	-0.0888	-0.0402	0.0085

In the majority of states, the same candidate has the statewide results above the county or precinct average and median results. It means that “larger” counties or precincts (as defined by the vote tally) tend to prefer this candidate over the others. This surprisingly interesting observation is worth a lot of attention, especially because this unusual feature is not only strong in all states, but very slowly weakens towards the end of the election cycle. The earlier elections are the most important ones. If we assume that this feature is caused by fraud, then the explanation is clear: although fraud was committed at every stage of the cycle (as our analysis shows), it was especially important in the early stages for fixing the public opinion in the desired direction.

Column 9 in Table 2 contains average slopes measurements, which indicate how much, on average across counties, the vote percent depends on the vote tally. The averages are supplemented with the 95% confidence intervals. The same candidate exhibits a striking difference from the others, and in almost all cases this dependency is established by the statistically significant non-zero and positive correlation. On the other hand, other candidates often exhibit an opposite property: correlation is statistically negative. This is a clear property of vote flipping.

Table 3 has a similar structure: each significant candidate (column 2) is analyzed in each state (column 1). Columns 6, 7, and 8 contain the point estimate and the confidence interval for the correlation between the vote tallies in each precinct with the vote percent for each candidate in each state. Columns 9, 10, and 11 contain the same statistics, but computed from the county-by-county aggregates. Both sets typically show statistically significant correlation for the same candidate in multiple states, and often statistically significant negative correlation for other candidates.

Table 3 Candidate votes statistics (cont.)

(1) State	(2) Candidate	(3) County Lower Bound Votes vs. Slope	(4) County Votes vs. Slope Correl.	(5) County Upper Bound Votes vs. Slope	(6) Precinct Lower Bound Slope Correl.	(7) Precinct Slope Correl.	(8) Precinct Upper Bound Slope Correl.	(9) County Lower Bound Slope Correl.	(10) County Slope Correl.	(11) County Upper Bound Slope Correl.
Iowa	Santorum	-0.1936	0.0040	0.2012	-0.0775	-0.0308	0.0159	-0.3288	-0.1405	0.0586
Iowa	Romney	0.0873	0.2799	0.4523	0.3062	0.3479	0.3884	0.2085	0.3899	0.5453
Iowa	Paul	-0.0259	0.1724	0.3576	-0.1792	-0.1336	-0.0874	-0.1089	0.0905	0.2828
Iowa	Gingrich	-0.2261	-0.0300	0.1684	-0.0672	-0.0205	0.0263	-0.2975	-0.1063	0.0931
Iowa	Perry	-0.2982	-0.1071	0.0922	-0.1641	-0.1182	-0.0719	-0.3674	-0.1833	0.0146
Iowa	Bachmann	-0.1248	0.0744	0.2679	-0.1842	-0.1386	-0.0925	-0.3838	-0.2017	-0.0045
New Hampshire	Romney	-0.4628	0.2355	0.7534	0.4051	0.4951	0.5757	0.0780	0.6745	0.9154
New Hampshire	Paul	-0.7065	-0.1384	0.5381	-0.4139	-0.3160	-0.2108	-0.9066	-0.6452	-0.0263
New Hampshire	Huntsman	-0.2802	0.4243	0.8317	-0.2487	-0.1400	-0.0277	-0.7926	-0.3253	0.3827
New Hampshire	Santorum	-0.5149	0.1698	0.7222	-0.2404	-0.1312	-0.0188	-0.8176	-0.3874	0.3204
New Hampshire	Gingrich	-0.3966	0.3105	0.7864	-0.2003	-0.0897	0.0233	-0.8482	-0.4690	0.2279
Arizona	Romney	-0.0934	0.4600	0.7962	-0.0108	0.0628	0.1358	-0.4095	0.1301	0.6022
Arizona	Santorum	-0.2715	0.3027	0.7180	-0.1179	-0.0447	0.0290	-0.6715	-0.2427	0.3078
Arizona	Gingrich	-0.0877	0.4645	0.7983	-0.0613	0.0124	0.0859	-0.5278	-0.0213	0.4964
Arizona	Paul	-0.2585	0.3154	0.7247	-0.1059	-0.0326	0.0411	-0.4969	0.0206	0.5273
Ohio	Romney	0.3083	0.4863	0.6315	0.2081	0.2272	0.2462	0.5709	0.6970	0.7910
Ohio	Santorum	0.1585	0.3561	0.5263	-0.0608	-0.0407	-0.0206	-0.6861	-0.5567	-0.3931
Ohio	Gingrich	-0.1535	0.0578	0.2640	-0.0316	-0.0114	0.0087	-0.2576	-0.0509	0.1603
Ohio	Paul	-0.1341	0.0776	0.2824	-0.2637	-0.2449	-0.2259	-0.3553	-0.1576	0.0536
Oklahoma	Santorum	0.0364	0.2583	0.4559	-0.2012	-0.1581	-0.1143	-0.4420	-0.2420	-0.0190
Oklahoma	Romney	0.0105	0.2339	0.4351	0.2735	0.3142	0.3539	0.1697	0.3792	0.5560
Oklahoma	Gingrich	-0.1564	0.0700	0.2894	-0.0922	-0.0478	-0.0033	-0.3049	-0.0869	0.1398
Oklahoma	Paul	-0.0190	0.2059	0.4109	-0.1498	-0.1060	-0.0617	-0.2265	-0.0026	0.2215
Alabama	Santorum	-0.2199	0.0338	0.2832	-0.1071	-0.0631	-0.0189	-0.3326	-0.1005	0.1432
Alabama	Gingrich	-0.1897	0.0653	0.3120	-0.1286	-0.0848	-0.0407	-0.5772	-0.3912	-0.1667
Alabama	Romney	0.0191	0.2696	0.4883	0.1262	0.1695	0.2122	0.2604	0.4711	0.6390
Alabama	Paul	-0.6256	-0.4437	-0.2160	-0.0486	-0.0044	0.0399	0.0879	0.3214	0.5213
Louisiana	Santorum	-0.2225	0.0246	0.2688	-0.0671	-0.0361	-0.0050	-0.5828	-0.3934	-0.1634
Louisiana	Romney	-0.0397	0.2081	0.4319	0.0730	0.1039	0.1346	0.2507	0.4677	0.6399
Louisiana	Gingrich	-0.3195	-0.0800	0.1692	0.0050	0.0361	0.0671	-0.2784	-0.0350	0.2126
Louisiana	Paul	-0.4164	-0.1901	0.0585	-0.0801	-0.0491	-0.0180	-0.0547	0.1938	0.4196
Wisconsin	Romney	0.0387	0.2679	0.4704	0.3705	0.3981	0.4250	0.3702	0.5543	0.6966
Wisconsin	Santorum	-0.2090	0.0238	0.2541	-0.2535	-0.2230	-0.1920	-0.5676	-0.3868	-0.1704
Wisconsin	Paul	-0.1586	0.0759	0.3022	-0.2447	-0.2140	-0.1829	-0.4474	-0.2407	-0.0096
Wisconsin	Gingrich	-0.2100	0.0228	0.2531	-0.2010	-0.1697	-0.1381	-0.5801	-0.4025	-0.1884
West Virginia	Romney	-0.0312	0.2360	0.4718	0.0574	0.1029	0.1479	-0.0685	0.2004	0.4422
West Virginia	Santorum	-0.1261	0.1441	0.3943	-0.1165	-0.0711	-0.0255	-0.5293	-0.3071	-0.0456
West Virginia	Paul	-0.2189	0.0493	0.3105	-0.0687	-0.0230	0.0227	-0.0472	0.2209	0.4593
West Virginia	Gingrich	-0.1905	0.0787	0.3370	-0.0820	-0.0364	0.0094	-0.5121	-0.2856	-0.0220

Kentucky	Romney	-0.0424	0.1505	0.3326	0.0698	0.1028	0.1355	0.0305	0.2086	0.3739
Kentucky	Paul	-0.0677	0.1256	0.3098	-0.1150	-0.0821	-0.0491	-0.2859	-0.1124	0.0682
Kentucky	Santorum	-0.0803	0.1131	0.2983	-0.0657	-0.0326	0.0005	-0.3068	-0.1349	0.0454
Kentucky	Gingrich	0.0274	0.2179	0.3931	-0.0483	-0.0151	0.0181	-0.3410	-0.1723	0.0072
Kentucky	Uncommit	-0.3623	-0.1833	0.0087	-0.0646	-0.0315	0.0016	-0.1791	0.0002	0.1794

In addition to analyzing this suspicious slope, let's look at how it correlates to the results in each county. Columns 3, 4, and 5 from Table 3 contain point estimate and the confidence interval for this relation. In case of vote flipping, both the beneficiary and the victim(s) of fraud will have statistically significant positive correlation in Column 4. In other words, the more fraud is committed, the more votes the beneficiary gets and the steeper its up slope is; the less votes the victim(s) get and the steeper their down slopes are. Actually, these slopes are not supposed to exist in the first place, but if their steepness determines the vote results on average across counties, then it can be attributed to the election fraud. Figure 12 and Figure 13 show that these curves can be flat for the same candidate, who exhibits anomalous upwards slopes in other cases.

Table 4 contains randomness tests' results for all states. Five different non-parametric randomness tests are applied to the entire randomly shuffled sequence of precincts. The p-values are shown in these tables, and all tests define the null hypothesis as "the sequence is random". The randomness is determined as a distance between an arbitrary origin point (a county location in this state) and each precinct. The location of each precinct is approximated by the location of its county. The confidence interval is 95%. It means that roughly 5% of rejections are acceptable. Even if the rejection percent exceeded 5% by a lot, the issue could be resolved by generating pseudo counties, as described in the earlier section. But in these 10 states the natural counties serve well to assure randomness. The randomness can be tested on sub-sequences as well, and these tests are expected to pass too.

Table 4 Randomness Testing

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State	**Index Start**	**Index End**	**Median Run Test p-value**	**Sign Test p-value**	**Runs Up Down Test p-value**	**Mann Kendall Test p-value**	**BBartels' Rank Test p-value**
Iowa	0	1773	0.9795	0.5317	0.4132	0.4592	0.6181
New Hampshire	0	322	0.5838	0.6338	0.2377	0.2047	0.8213
Arizona	0	721	0.5166	0.6248	0.0010	0.3349	0.6786
Ohio	0	9648	0.0399	0.7794	0.0149	0.0541	0.4685
Oklahoma	0	1960	0.8093	0.9446	0.1111	0.3240	0.5493
Alabama	0	2039	0.7754	0.7536	0.3208	0.1941	0.7772
Louisiana	0	4394	0.6441	0.7246	0.2651	0.9807	0.5794
Wisconsin	0	3768	0.2084	0.8040	0.1292	0.6119	0.6324
West Virginia	0	1845	0.6644	0.6364	0.8664	0.2570	0.8923
Arkansas	0	2320	0.9458	0.4361	0.0599	0.2038	0.4131
Kentucky	0	3537	0.1603	0.3453	0.3354	0.1068	0.3554

Since we have generated random samples (which are biased towards the precinct vote tally, but this is statistically acceptable for randomness), we can draw a random sample and compute the point estimate about the entire population. We draw percentile subsequences at 15%, 30%, 45%, 60%, 75%, 90%, and 100% of the statewide vote tally. The same candidate has consistent deviation from the "official" results in his favor, while others often are supposed to have higher vote percent based on the randomly drawn samples.

Table 5 contains partial correlations (column 7) between vote percentages for all pairs of candidates. The rows are sorted by the partial correlation within each state. One candidate shows consistent leadership in negative partial correlations in all states[14]. Since this method does not explicitly show the beneficiary or victim of vote flipping, other methods pinpoint who is who. But consistent leadership of the same candidate against various other candidates across states clearly shows the data anomaly, which is attributed to vote flipping in these leading pairs.

Table 5 Pairwise Candidate Partial Correlations on Vote Percent

(1) State Name	(2) Candidate Name	(3) Candidate Name	(4) Lower Vote Fraction Correlation	(5) Point Estimate Vote Fraction Correlation	(6) Upper Vote Fraction Correlation	(7) Point Estimate Vote Fraction Partial Correlation
Iowa	Santorum	Paul	-0.392256	-0.352144	-0.310696	-0.545285
Iowa	Romney	Paul	-0.260532	-0.216619	-0.171811	-0.496773
Iowa	Romney	Santorum	-0.346589	-0.304968	-0.262149	-0.490519
Iowa	Gingrich	Paul	-0.271919	-0.228268	-0.183680	-0.383743
Iowa	Romney	Perry	-0.301517	-0.258606	-0.214650	-0.366936
Iowa	Paul	Perry	-0.230582	-0.186039	-0.140717	-0.359224
Iowa	Santorum	Perry	-0.151377	-0.105581	-0.059333	-0.331323
Iowa	Gingrich	Santorum	-0.216435	-0.171624	-0.126091	-0.326155
Iowa	Romney	Bachmann	-0.268565	-0.224835	-0.180181	-0.282903
Iowa	Gingrich	Perry	-0.198771	-0.153653	-0.107884	-0.236129
Iowa	Paul	Bachmann	-0.179201	-0.133777	-0.087784	-0.234236
Iowa	Romney	Gingrich	-0.067573	-0.021100	0.025465	-0.211825
Iowa	Santorum	Bachmann	-0.068923	-0.022456	0.024109	-0.183187
Iowa	Gingrich	Bachmann	-0.143971	-0.098088	-0.051785	-0.141312
Iowa	Bachmann	Perry	-0.000394	0.046146	0.092487	-0.061305
New Hampshire	Romney	Paul	-0.056633	0.052822	0.161023	-0.332354
New Hampshire	Santorum	Huntsman	0.021390	0.130216	0.235991	-0.257487
New Hampshire	Paul	Huntsman	0.023841	0.132625	0.238305	-0.231496
New Hampshire	Romney	Santorum	0.064150	0.172074	0.276020	-0.059138
New Hampshire	Gingrich	Santorum	0.256354	0.355536	0.447310	0.080272
New Hampshire	Paul	Gingrich	0.255763	0.354984	0.446804	0.100720
New Hampshire	Romney	Gingrich	0.246078	0.345917	0.438493	0.129184
New Hampshire	Gingrich	Huntsman	0.299487	0.395684	0.483917	0.154480
New Hampshire	Romney	Huntsman	0.282254	0.379688	0.469369	0.221435
New Hampshire	Santorum	Paul	0.393882	0.482281	0.561840	0.353898
Arizona	Romney	Santorum	-0.396196	-0.332854	-0.266358	-0.463722
Arizona	Romney	Paul	-0.340408	-0.274255	-0.205401	-0.375355
Arizona	Romney	Gingrich	-0.270699	-0.201719	-0.130678	-0.350336
Arizona	Santorum	Paul	-0.207091	-0.136184	-0.063854	-0.253679
Arizona	Gingrich	Paul	-0.193714	-0.122481	-0.049963	-0.189721
Arizona	Gingrich	Santorum	-0.124093	-0.051596	0.021450	-0.175965
Ohio	Romney	Santorum	-0.467631	-0.451894	-0.435871	-0.533605
Ohio	Romney	Paul	-0.256657	-0.237922	-0.219009	-0.361847
Ohio	Santorum	Paul	-0.145074	-0.125484	-0.105796	-0.283094
Ohio	Romney	Gingrich	-0.172913	-0.153489	-0.133946	-0.269026
Ohio	Santorum	Gingrich	-0.102647	-0.082863	-0.063014	-0.195988

[14] The only exception is Iowa, where all candidates lost votes in unison, whereas Romney had a single opposing curve. In Iowa Romney has very small partial correlations as well.

Ohio	Paul	Gingrich	-0.155192	-0.135659	-0.116019	-0.190871
Oklahoma	Romney	Santorum	-0.400072	-0.362221	-0.323138	-0.477757
Oklahoma	Santorum	Gingrich	-0.267051	-0.225451	-0.183013	-0.383747
Oklahoma	Romney	Gingrich	-0.252975	-0.211074	-0.168382	-0.367140
Oklahoma	Santorum	Paul	-0.275645	-0.234238	-0.191964	-0.357492
Oklahoma	Romney	Paul	-0.178963	-0.135774	-0.092063	-0.284757
Oklahoma	Gingrich	Paul	-0.205637	-0.162854	-0.119451	-0.260714
Alabama	Romney	Santorum	-0.263036	-0.222173	-0.180515	-0.260159
Alabama	Romney	Gingrich	-0.224667	-0.183053	-0.140772	-0.231663
Alabama	Gingrich	Santorum	-0.212152	-0.170321	-0.127867	-0.222291
Alabama	Gingrich	Paul	-0.119833	-0.076834	-0.033547	-0.087715
Alabama	Romney	Paul	-0.093005	-0.049807	-0.006422	-0.069187
Alabama	Santorum	Paul	-0.049054	-0.005667	0.037741	-0.029976
Louisiana	Romney	Santorum	-0.225543	-0.197293	-0.168711	-0.232165
Louisiana	Santorum	Paul	-0.180693	-0.151939	-0.122925	-0.181988
Louisiana	Gingrich	Santorum	-0.107905	-0.078590	-0.049138	-0.113199
Louisiana	Romney	Paul	-0.099731	-0.070373	-0.040892	-0.107816
Louisiana	Gingrich	Paul	-0.099353	-0.069993	-0.040511	-0.087636
Louisiana	Romney	Gingrich	-0.062305	-0.032800	-0.003237	-0.061442
Wisconsin	Romney	Santorum	-0.335902	-0.307270	-0.278071	-0.312035
Wisconsin	Romney	Paul	-0.293908	-0.264463	-0.234516	-0.263269
Wisconsin	Romney	Gingrich	-0.209672	-0.178943	-0.147861	-0.135790
Wisconsin	Santorum	Paul	-0.046350	-0.014444	0.017491	-0.111460
Wisconsin	Paul	Gingrich	0.033441	0.065298	0.097023	0.013530
Wisconsin	Gingrich	Santorum	0.080322	0.111962	0.143377	0.060639
West Virginia	Romney	Santorum	-0.539597	-0.506442	-0.471718	-0.678629
West Virginia	Romney	Paul	-0.490194	-0.454741	-0.417786	-0.663455
West Virginia	Romney	Gingrich	-0.360772	-0.320423	-0.278876	-0.515522
West Virginia	Santorum	Paul	-0.136808	-0.091758	-0.046329	-0.424863
West Virginia	Gingrich	Santorum	-0.111277	-0.065989	-0.020428	-0.287415
West Virginia	Paul	Gingrich	-0.122453	-0.077262	-0.031751	-0.274837
Kentucky	Romney	Paul	-0.501853	-0.476785	-0.450917	-0.620298
Kentucky	Romney	Santorum	-0.325628	-0.295850	-0.265485	-0.442056
Kentucky	Romney	Uncommitted	-0.306355	-0.276190	-0.245471	-0.435335
Kentucky	Romney	Gingrich	-0.281098	-0.250465	-0.219322	-0.387027
Kentucky	Santorum	Paul	-0.124816	-0.092242	-0.059470	-0.287669
Kentucky	Paul	Uncommitted	-0.119400	-0.086789	-0.053990	-0.273907
Kentucky	Gingrich	Paul	-0.112642	-0.079986	-0.047157	-0.245106
Kentucky	Santorum	Uncommitted	-0.097121	-0.064374	-0.031488	-0.170506
Kentucky	Gingrich	Uncommitted	-0.105453	-0.072753	-0.039895	-0.162368
Kentucky	Santorum	Gingrich	-0.055662	-0.022751	0.010210	-0.118189

Table 6 shows the impact of population density on vote tally and on our analysis in general. Column 4 contains mostly positive values, which means that urban areas in fact have precincts with bigger vote tally. Column 5 shows our original data anomaly: correlation between vote tally and vote count is unusually high for the same candidate in all states, and often unusually low for the other candidates. Column 6 shows that this “special” candidate is not special in “urban” versus “rural” areas. Other candidates, like Ron Paul, show stronger popularity difference in urban versus rural areas. In other words, *if we assume that both Ron Paul and Mitt Romney got the same vote percent in the entire state, and Ron Paul got more votes in urban areas, we refute the common argument that Mitt Romney is more popular in cities.* Column 7 partial correlation between vote tally and vote count, which are hypothetically assumed to be uncorrelated with the population density. As we can see, the data anomaly is still evident even after this correction, since this density factor did not appreciably favor Mitt Romney.

Table 6 Population Density and Vote Tally Analysis

(1) State Name	(2) Candidate Name	(3) Statewide Vote Fraction	(4) Precinct Population Density vs Total Vote Tally Correlation	(5) Precinct Vote Tally versus Candidates' Vote Count	(6) Precinct Population Density vs Candidate Vote Count Correlation	(7) Precinct Partial Correlation Candidate Vote Count vs Total Vote Tally :exclude Density
Iowa	Santorum	24.59%	0.331914	-0.030837	0.214173	-0.110616
Iowa	Romney	24.56%	0.331914	0.347925	0.314395	0.272004
Iowa	Paul	21.45%	0.331914	-0.133617	0.399496	-0.307847
Iowa	Gingrich	13.32%	0.331914	-0.020502	0.263602	-0.118683
Iowa	Perry	10.35%	0.331914	-0.118246	0.167496	-0.186929
Iowa	Bachmann	4.98%	0.331914	-0.138605	0.139305	-0.197880
New Hampshire	Romney	39.28%	0.522564	0.495135	0.508012	0.312732
New Hampshire	Paul	22.89%	0.522564	-0.315992	0.540699	-0.834530
New Hampshire	Huntsman	16.89%	0.522564	-0.139975	0.481832	-0.524376
New Hampshire	Santorum	9.43%	0.522564	-0.131228	0.519803	-0.553100
New Hampshire	Gingrich	9.43%	0.522564	-0.089662	0.452548	-0.428974
Arizona	Romney	46.87%	0.107802	0.062825	0.111341	0.051440
Arizona	Santorum	27.05%	0.107802	-0.044721	0.102431	-0.056386
Arizona	Gingrich	16.02%	0.107802	0.012351	0.097666	0.001842
Arizona	Paul	8.61%	0.107802	-0.032603	0.116265	-0.045711
Ohio	Romney	37.96%	0.095804	0.227227	0.080318	0.221261
Ohio	Santorum	36.95%	0.095804	-0.040710	0.080044	-0.048759
Ohio	Gingrich	14.60%	0.095804	-0.011446	0.146799	-0.025909
Ohio	Paul	9.33%	0.095804	-0.244914	0.033155	-0.249374
Oklahoma	Santorum	33.78%	0.155112	-0.158061	0.167810	-0.189026
Oklahoma	Romney	28.03%	0.155112	0.314234	0.121099	0.301287
Oklahoma	Gingrich	27.46%	0.155112	-0.047846	0.130251	-0.069475
Oklahoma	Paul	9.63%	0.155112	-0.105983	0.223172	-0.146004
Alabama	Santorum	34.50%	0.272695	-0.063128	0.234769	-0.135957
Alabama	Gingrich	29.30%	0.272695	-0.084801	0.251579	-0.164747
Alabama	Romney	28.99%	0.272695	0.169512	0.309041	0.093156
Alabama	Paul	4.97%	0.272695	-0.004365	0.278015	-0.086757
Louisiana	Santorum	48.99%	0.162993	-0.036080	0.172686	-0.066090
Louisiana	Romney	26.69%	0.162993	0.103859	0.144002	0.082335
Louisiana	Gingrich	15.91%	0.162993	0.036084	0.136624	0.014135

Louisiana	Paul	6.15%	0.162993	-0.049074	0.205863	-0.085582
Louisiana	Roemer	1.18%	0.162993	-0.048415	0.087857	-0.063832
Wisconsin	Romney	44.03%	0.379802	0.398069	0.419784	0.284219
Wisconsin	Santorum	36.83%	0.379802	-0.222984	0.314762	-0.390106
Wisconsin	Paul	11.15%	0.379802	-0.214003	0.275213	-0.358162
Wisconsin	Gingrich	5.84%	0.379802	-0.169705	0.209756	-0.275704
West Virginia	Romney	69.56%	0.188111	0.102887	0.195198	0.068693
West Virginia	Santorum	12.09%	0.188111	-0.071110	0.091791	-0.090365
West Virginia	Paul	11.04%	0.188111	-0.023013	0.261142	-0.076088
West Virginia	Gingrich	6.29%	0.188111	-0.036380	0.004955	-0.037990
West Virginia	Roemer	1.01%	0.188111	-0.090566	0.022875	-0.096619
Kentucky	Romney	66.77%	0.159152	0.102770	0.166073	0.078413
Kentucky	Paul	12.53%	0.159152	-0.082096	0.172968	-0.112739
Kentucky	Santorum	8.87%	0.159152	-0.032643	0.149038	-0.057735
Kentucky	Gingrich	5.95%	0.159152	-0.015122	0.086814	-0.029423
Kentucky	Uncommitted	5.88%	0.159152	-0.031534	0.085472	-0.045888

Finally, let’s run two one-tail hypothesis tests for each candidate in 10 states. Columns 3 and 4 in Table 7 contain percentiles in the vote tally and precinct count from the random sample. Column 5 contain maximum likelihood point estimate of the statewide vote percent result based on the current random sample. We can see the respective deviation from the “official” result in column 6. The negative number indicates a fraud beneficiary, while the positive number: fraud victim. Columns 7 and 8 contain hypothesis test results. Let’s assume that the number of actual votes cast for each candidate equals or exceeds the official vote count. Column 7 contains the p-value for this hypothesis.

Let’s assume that the number of actual votes cast for each candidate equals or is less than the official vote count. Column 8 contains the p-value for this hypothesis. The same candidate exhibits anomalous properties: random samples indicate that his final vote percent in the entire population is supposed to be lower than the “official” result, i.e. the null hypothesis about higher than “official” result is rejected in most random samples. The opposite is true for the other candidates: statistically, their random samples suggest that lower than “official” results are rejected for the majority of random samples at 95% confidence level. In some rare cases some other candidates exhibit fraud-like properties, but in these cases the magnitude of this “fraud” in column 6 is negligible.

Table 7 Random sample testing

(1) State Name	(2) Candidate Name	(3) Vote Tally Percentile	(4) Precinct Count Percentile	(5) Point Estimate Statewide Vote Percent	(6) Point Estimate Statewide Vote Percent Deviation from Official Vote Percent	(7) P-value Reject Bigger Vote Count	**(8)** P-value Reject Smaller Vote Count
Iowa	Santorum	15.03%	36.41%	25.0762%	0.4878%	0.952427	0.049444
Iowa	Santorum	30.05%	55.92%	25.2443%	0.6559%	0.999751	0.000263
Iowa	Santorum	45.10%	69.95%	24.8636%	0.2752%	0.978622	0.022074
Iowa	Santorum	60.06%	81.34%	24.8587%	0.2703%	0.996461	0.003686
Iowa	Santorum	75.03%	89.40%	24.6576%	0.0692%	0.836965	0.166862
Iowa	Santorum	90.12%	96.56%	24.6972%	0.1088%	1.000000	0.000000

Iowa	Santorum	100.00%	100.00%	24.5884%	0.0000%	1.000000	0.000000
Iowa	Romney	15.03%	36.41%	17.9969%	-6.5634%	0.000000	1.000000
Iowa	Romney	30.05%	55.92%	18.9347%	-5.6257%	0.000000	1.000000
Iowa	Romney	45.10%	69.95%	19.8716%	-4.6888%	0.000000	1.000000
Iowa	Romney	60.06%	81.34%	20.9915%	-3.5689%	0.000000	1.000000
Iowa	Romney	75.03%	89.40%	22.5052%	-2.0551%	0.000000	1.000000
Iowa	Romney	90.12%	96.56%	23.5633%	-0.9971%	1.000000	0.000000
Iowa	Romney	100.00%	100.00%	24.5604%	0.0000%	1.000000	0.000000
Iowa	Paul	15.03%	36.41%	25.3070%	3.8524%	1.000000	0.000000
Iowa	Paul	30.05%	55.92%	24.9345%	3.4799%	1.000000	0.000000
Iowa	Paul	45.10%	69.95%	24.4417%	2.9871%	1.000000	0.000000
Iowa	Paul	60.06%	81.34%	23.5880%	2.1334%	1.000000	0.000000
Iowa	Paul	75.03%	89.40%	22.7599%	1.3053%	1.000000	0.000000
Iowa	Paul	90.12%	96.56%	21.9432%	0.4887%	1.000000	0.000000
Iowa	Paul	100.00%	100.00%	21.4546%	0.0000%	1.000000	0.000000
Iowa	Gingrich	15.03%	36.41%	12.9250%	-0.3939%	0.045487	0.956744
Iowa	Gingrich	30.05%	55.92%	12.9044%	-0.4145%	0.002723	0.997428
Iowa	Gingrich	45.10%	69.95%	12.8838%	-0.4351%	0.000027	0.999975
Iowa	Gingrich	60.06%	81.34%	13.1664%	-0.1524%	0.028687	0.972424
Iowa	Gingrich	75.03%	89.40%	13.2785%	-0.0404%	0.238857	0.767142
Iowa	Gingrich	90.12%	96.56%	13.3650%	0.0461%	1.000000	0.000000
Iowa	Gingrich	100.00%	100.00%	13.3189%	0.0000%	1.000000	0.000000
New Hampshire	Romney	15.40%	38.70%	35.8881%	-3.3879%	0.000000	1.000000
New Hampshire	Romney	30.28%	57.59%	36.0922%	-3.1838%	0.000000	1.000000
New Hampshire	Romney	45.29%	71.21%	36.9486%	-2.3274%	0.000000	1.000000
New Hampshire	Romney	60.39%	83.28%	37.2158%	-2.0602%	0.000000	1.000000
New Hampshire	Romney	75.01%	93.19%	37.7120%	-1.5639%	1.000000	0.000000
New Hampshire	Romney	90.22%	97.52%	38.7790%	-0.4970%	1.000000	0.000000
New Hampshire	Romney	100.00%	100.00%	39.2760%	0.0000%	1.000000	0.000000
New Hampshire	Paul	15.40%	38.70%	25.6501%	2.7616%	1.000000	0.000000
New Hampshire	Paul	30.28%	57.59%	25.7595%	2.8711%	1.000000	0.000000
New Hampshire	Paul	45.29%	71.21%	24.9369%	2.0485%	1.000000	0.000000
New Hampshire	Paul	60.39%	83.28%	24.4029%	1.5144%	1.000000	0.000000
New Hampshire	Paul	75.01%	93.19%	23.8551%	0.9667%	1.000000	0.000000
New Hampshire	Paul	90.22%	97.52%	23.3235%	0.4351%	1.000000	0.000000
New Hampshire	Paul	100.00%	100.00%	22.8884%	0.0000%	1.000000	0.000000
New Hampshire	Santorum	15.40%	38.70%	9.2343%	-0.1960%	0.078315	0.924443
New Hampshire	Santorum	30.28%	57.59%	9.6613%	0.2310%	0.995334	0.004872
New Hampshire	Santorum	45.29%	71.21%	9.7636%	0.3332%	1.000000	0.000000
New Hampshire	Santorum	60.39%	83.28%	9.8879%	0.4576%	1.000000	0.000000
New Hampshire	Santorum	75.01%	93.19%	9.6782%	0.2479%	1.000000	0.000000
New Hampshire	Santorum	90.22%	97.52%	9.5185%	0.0881%	0.999998	0.000002
New Hampshire	Santorum	100.00%	100.00%	9.4303%	0.0000%	1.000000	0.000000
New Hampshire	Gingrich	15.40%	38.70%	9.1092%	-0.3167%	0.010625	0.989902
New Hampshire	Gingrich	30.28%	57.59%	9.0858%	-0.3401%	0.000065	0.999939
New Hampshire	Gingrich	45.29%	71.21%	9.1273%	-0.2986%	0.000002	0.999998
New Hampshire	Gingrich	60.39%	83.28%	9.2875%	-0.1384%	0.001847	0.998234
New Hampshire	Gingrich	75.01%	93.19%	9.2738%	-0.1521%	0.000004	0.999996
New Hampshire	Gingrich	90.22%	97.52%	9.3011%	-0.1248%	0.000000	1.000000
New Hampshire	Gingrich	100.00%	100.00%	9.4259%	0.0000%	1.000000	0.000000
Arizona	Romney	15.00%	46.81%	39.3775%	-7.4942%	0.000000	1.000000
Arizona	Romney	30.06%	76.04%	40.6982%	-6.1735%	0.000000	1.000000
Arizona	Romney	45.06%	93.91%	43.2548%	-3.6170%	0.000000	1.000000
Arizona	Romney	62.78%	98.20%	45.2542%	-1.6176%	1.000000	0.000000
Arizona	Romney	85.97%	98.75%	45.6628%	-1.2090%	1.000000	0.000000
Arizona	Romney	99.63%	99.72%	46.9027%	0.0310%	1.000000	0.000000
Arizona	Romney	100.00%	100.00%	46.8718%	0.0000%	1.000000	0.000000
Arizona	Santorum	15.00%	46.81%	31.7839%	4.7327%	1.000000	0.000000
Arizona	Santorum	30.06%	76.04%	31.8002%	4.7490%	1.000000	0.000000
Arizona	Santorum	45.06%	93.91%	30.6378%	3.5866%	1.000000	0.000000
Arizona	Santorum	62.78%	98.20%	28.6016%	1.5504%	1.000000	0.000000
Arizona	Santorum	85.97%	98.75%	27.7677%	0.7165%	1.000000	0.000000
Arizona	Santorum	99.63%	99.72%	27.0177%	-0.0335%	1.000000	0.000000
Arizona	Santorum	100.00%	100.00%	27.0512%	0.0000%	1.000000	0.000000
Arizona	Gingrich	15.00%	46.81%	17.2705%	1.2496%	1.000000	0.000000
Arizona	Gingrich	30.06%	76.04%	16.5522%	0.5313%	1.000000	0.000000

Arizona	Gingrich	45.06%	93.91%	15.8965%	-0.1244%	0.014316	0.985961
Arizona	Gingrich	62.78%	98.20%	15.7820%	-0.2389%	0.000000	1.000000
Arizona	Gingrich	85.97%	98.75%	16.3323%	0.3114%	1.000000	0.000000
Arizona	Gingrich	99.63%	99.72%	16.0219%	0.0010%	1.000000	0.000000
Arizona	Gingrich	100.00%	100.00%	16.0209%	0.0000%	1.000000	0.000000
Arizona	Paul	15.00%	46.81%	10.0412%	1.4275%	1.000000	0.000000
Arizona	Paul	30.06%	76.04%	9.5266%	0.9129%	1.000000	0.000000
Arizona	Paul	45.06%	93.91%	8.9421%	0.3285%	1.000000	0.000000
Arizona	Paul	62.78%	98.20%	8.9823%	0.3686%	1.000000	0.000000
Arizona	Paul	85.97%	98.75%	8.7624%	0.1487%	1.000000	0.000000
Arizona	Paul	99.63%	99.72%	8.6133%	-0.0004%	1.000000	0.000000
Arizona	Paul	100.00%	100.00%	8.6137%	0.0000%	1.000000	0.000000
Ohio	Romney	15.01%	34.47%	32.5303%	-5.4332%	0.000000	1.000000
Ohio	Romney	30.00%	52.31%	33.8041%	-4.1594%	0.000000	1.000000
Ohio	Romney	45.00%	66.29%	34.5875%	-3.3760%	0.000000	1.000000
Ohio	Romney	60.01%	77.95%	35.5759%	-2.3876%	0.000000	1.000000
Ohio	Romney	75.02%	87.63%	36.4666%	-1.4969%	1.000000	0.000000
Ohio	Romney	90.01%	95.76%	37.3486%	-0.6149%	1.000000	0.000000
Ohio	Romney	100.00%	100.00%	37.9635%	0.0000%	1.000000	0.000000
Ohio	Santorum	15.01%	34.47%	39.2283%	2.2740%	1.000000	0.000000
Ohio	Santorum	30.00%	52.31%	38.7033%	1.7490%	1.000000	0.000000
Ohio	Santorum	45.00%	66.29%	38.5615%	1.6072%	1.000000	0.000000
Ohio	Santorum	60.01%	77.95%	38.1242%	1.1700%	1.000000	0.000000
Ohio	Santorum	75.02%	87.63%	37.7183%	0.7641%	1.000000	0.000000
Ohio	Santorum	90.01%	95.76%	37.2867%	0.3324%	1.000000	0.000000
Ohio	Santorum	100.00%	100.00%	36.9543%	0.0000%	1.000000	0.000000
Ohio	Gingrich	15.01%	34.47%	15.6070%	1.0106%	1.000000	0.000000
Ohio	Gingrich	30.00%	52.31%	15.4919%	0.8955%	1.000000	0.000000
Ohio	Gingrich	45.00%	66.29%	15.2412%	0.6448%	1.000000	0.000000
Ohio	Gingrich	60.01%	77.95%	15.0530%	0.4566%	1.000000	0.000000
Ohio	Gingrich	75.02%	87.63%	14.8886%	0.2921%	1.000000	0.000000
Ohio	Gingrich	90.01%	95.76%	14.7008%	0.1044%	1.000000	0.000000
Ohio	Gingrich	100.00%	100.00%	14.5964%	0.0000%	1.000000	0.000000
Ohio	Paul	15.01%	34.47%	11.0983%	1.7682%	1.000000	0.000000
Ohio	Paul	30.00%	52.31%	10.6067%	1.2766%	1.000000	0.000000
Ohio	Paul	45.00%	66.29%	10.2930%	0.9629%	1.000000	0.000000
Ohio	Paul	60.01%	77.95%	9.9917%	0.6616%	1.000000	0.000000
Ohio	Paul	75.02%	87.63%	9.7179%	0.3878%	1.000000	0.000000
Ohio	Paul	90.01%	95.76%	9.4896%	0.1595%	1.000000	0.000000
Ohio	Paul	100.00%	100.00%	9.3301%	0.0000%	1.000000	0.000000
Oklahoma	Santorum	15.02%	39.88%	35.3947%	1.6149%	1.000000	0.000000
Oklahoma	Santorum	30.00%	57.88%	34.7505%	0.9707%	1.000000	0.000000
Oklahoma	Santorum	45.04%	71.44%	34.6029%	0.8231%	1.000000	0.000000
Oklahoma	Santorum	60.11%	81.59%	34.1192%	0.3394%	0.999999	0.000001
Oklahoma	Santorum	75.02%	90.06%	34.0654%	0.2857%	1.000000	0.000000
Oklahoma	Santorum	90.01%	96.69%	33.8565%	0.0767%	1.000000	0.000000
Oklahoma	Santorum	100.00%	100.00%	33.7798%	0.0000%	1.000000	0.000000
Oklahoma	Romney	15.02%	39.88%	22.8543%	-5.1729%	0.000000	1.000000
Oklahoma	Romney	30.00%	57.88%	24.2279%	-3.7993%	0.000000	1.000000
Oklahoma	Romney	45.04%	71.44%	24.8306%	-3.1966%	0.000000	1.000000
Oklahoma	Romney	60.11%	81.59%	26.1424%	-1.8849%	0.000000	1.000000
Oklahoma	Romney	75.02%	90.06%	26.8110%	-1.2162%	1.000000	0.000000
Oklahoma	Romney	90.01%	96.69%	27.6052%	-0.4220%	1.000000	0.000000
Oklahoma	Romney	100.00%	100.00%	28.0272%	0.0000%	1.000000	0.000000
Oklahoma	Gingrich	15.02%	39.88%	28.8577%	1.3976%	1.000000	0.000000
Oklahoma	Gingrich	30.00%	57.88%	28.5957%	1.1357%	1.000000	0.000000
Oklahoma	Gingrich	45.04%	71.44%	28.5040%	1.0439%	1.000000	0.000000
Oklahoma	Gingrich	60.11%	81.59%	28.1538%	0.6937%	1.000000	0.000000
Oklahoma	Gingrich	75.02%	90.06%	27.8626%	0.4025%	1.000000	0.000000
Oklahoma	Gingrich	90.01%	96.69%	27.6415%	0.1814%	1.000000	0.000000
Oklahoma	Gingrich	100.00%	100.00%	27.4601%	0.0000%	1.000000	0.000000
Oklahoma	Paul	15.02%	39.88%	11.4399%	1.8147%	1.000000	0.000000
Oklahoma	Paul	30.00%	57.88%	11.0698%	1.4447%	1.000000	0.000000
Oklahoma	Paul	45.04%	71.44%	10.8177%	1.1925%	1.000000	0.000000
Oklahoma	Paul	60.11%	81.59%	10.3914%	0.7663%	1.000000	0.000000
Oklahoma	Paul	75.02%	90.06%	10.0908%	0.4656%	1.000000	0.000000

Oklahoma	Paul	90.01%	96.69%	9.7737%	0.1486%	1.000000	0.000000
Oklahoma	Paul	100.00%	100.00%	9.6252%	0.0000%	1.000000	0.000000
Alabama	Santorum	15.00%	44.41%	36.3326%	1.8308%	1.000000	0.000000
Alabama	Santorum	30.01%	62.94%	36.0217%	1.5200%	1.000000	0.000000
Alabama	Santorum	45.03%	76.86%	35.6339%	1.1321%	1.000000	0.000000
Alabama	Santorum	60.15%	85.49%	35.5352%	1.0334%	1.000000	0.000000
Alabama	Santorum	75.15%	92.11%	34.9446%	0.4428%	1.000000	0.000000
Alabama	Santorum	90.01%	97.40%	34.4968%	-0.0050%	1.000000	0.000000
Alabama	Santorum	100.00%	100.00%	34.5018%	0.0000%	1.000000	0.000000
Alabama	Gingrich	15.00%	44.41%	30.5337%	1.2309%	1.000000	0.000000
Alabama	Gingrich	30.01%	62.94%	30.3346%	1.0318%	1.000000	0.000000
Alabama	Gingrich	45.03%	76.86%	30.0285%	0.7257%	1.000000	0.000000
Alabama	Gingrich	60.15%	85.49%	29.7367%	0.4338%	1.000000	0.000000
Alabama	Gingrich	75.15%	92.11%	29.4684%	0.1656%	1.000000	0.000000
Alabama	Gingrich	90.01%	97.40%	29.3760%	0.0732%	1.000000	0.000000
Alabama	Gingrich	100.00%	100.00%	29.3028%	0.0000%	1.000000	0.000000
Alabama	Romney	15.00%	44.41%	24.8529%	-4.1395%	0.000000	1.000000
Alabama	Romney	30.01%	62.94%	25.7081%	-3.2842%	0.000000	1.000000
Alabama	Romney	45.03%	76.86%	26.5266%	-2.4658%	0.000000	1.000000
Alabama	Romney	60.15%	85.49%	27.1233%	-1.8691%	0.000000	1.000000
Alabama	Romney	75.15%	92.11%	28.1720%	-0.8204%	1.000000	0.000000
Alabama	Romney	90.01%	97.40%	28.8734%	-0.1190%	1.000000	0.000000
Alabama	Romney	100.00%	100.00%	28.9923%	0.0000%	1.000000	0.000000
Alabama	Paul	15.00%	44.41%	5.5437%	0.5703%	1.000000	0.000000
Alabama	Paul	30.01%	62.94%	5.4248%	0.4515%	1.000000	0.000000
Alabama	Paul	45.03%	76.86%	5.3446%	0.3713%	1.000000	0.000000
Alabama	Paul	60.15%	85.49%	5.2290%	0.2557%	1.000000	0.000000
Alabama	Paul	75.15%	92.11%	5.1111%	0.1378%	1.000000	0.000000
Alabama	Paul	90.01%	97.40%	5.0180%	0.0447%	1.000000	0.000000
Alabama	Paul	100.00%	100.00%	4.9733%	0.0000%	1.000000	0.000000
Louisiana	Santorum	15.03%	51.35%	53.1597%	4.1704%	1.000000	0.000000
Louisiana	Santorum	30.03%	66.85%	52.1909%	3.2015%	1.000000	0.000000
Louisiana	Santorum	45.03%	78.43%	51.9146%	2.9253%	1.000000	0.000000
Louisiana	Santorum	60.01%	87.17%	51.1539%	2.1646%	1.000000	0.000000
Louisiana	Santorum	75.01%	93.88%	50.6335%	1.6442%	1.000000	0.000000
Louisiana	Santorum	90.03%	98.66%	49.8224%	0.8331%	1.000000	0.000000
Louisiana	Santorum	100.00%	100.00%	48.9893%	0.0000%	1.000000	0.000000
Louisiana	Romney	15.03%	51.35%	21.9033%	-4.7894%	0.000000	1.000000
Louisiana	Romney	30.03%	66.85%	23.1710%	-3.5218%	0.000000	1.000000
Louisiana	Romney	45.03%	78.43%	23.6854%	-3.0073%	0.000000	1.000000
Louisiana	Romney	60.01%	87.17%	24.5625%	-2.1303%	0.000000	1.000000
Louisiana	Romney	75.01%	93.88%	25.3205%	-1.3722%	1.000000	0.000000
Louisiana	Romney	90.03%	98.66%	26.1279%	-0.5649%	1.000000	0.000000
Louisiana	Romney	100.00%	100.00%	26.6928%	0.0000%	1.000000	0.000000
Louisiana	Gingrich	15.03%	51.35%	15.1837%	-0.7253%	0.000154	0.999856
Louisiana	Gingrich	30.03%	66.85%	15.2562%	-0.6529%	0.000000	1.000000
Louisiana	Gingrich	45.03%	78.43%	15.3162%	-0.5928%	0.000000	1.000000
Louisiana	Gingrich	60.01%	87.17%	15.4149%	-0.4941%	0.000000	1.000000
Louisiana	Gingrich	75.01%	93.88%	15.3264%	-0.5826%	0.000000	1.000000
Louisiana	Gingrich	90.03%	98.66%	15.4815%	-0.4276%	1.000000	0.000000
Louisiana	Gingrich	100.00%	100.00%	15.9090%	0.0000%	1.000000	0.000000
Louisiana	Paul	15.03%	51.35%	6.8725%	0.7210%	1.000000	0.000000
Louisiana	Paul	30.03%	66.85%	6.7641%	0.6126%	1.000000	0.000000
Louisiana	Paul	45.03%	78.43%	6.6407%	0.4892%	1.000000	0.000000
Louisiana	Paul	60.01%	87.17%	6.5120%	0.3605%	1.000000	0.000000
Louisiana	Paul	75.01%	93.88%	6.4476%	0.2961%	1.000000	0.000000
Louisiana	Paul	90.03%	98.66%	6.3237%	0.1722%	1.000000	0.000000
Louisiana	Paul	100.00%	100.00%	6.1515%	0.0000%	1.000000	0.000000
Wisconsin	Romney	15.02%	40.33%	40.3502%	-3.6781%	0.000000	1.000000
Wisconsin	Romney	30.00%	58.42%	41.5319%	-2.4964%	0.000000	1.000000
Wisconsin	Romney	45.01%	72.46%	41.7841%	-2.2442%	0.000000	1.000000
Wisconsin	Romney	60.00%	82.59%	42.6132%	-1.4151%	1.000000	0.000000
Wisconsin	Romney	75.03%	91.03%	43.0773%	-0.9511%	1.000000	0.000000
Wisconsin	Romney	90.05%	97.11%	43.6941%	-0.3342%	1.000000	0.000000
Wisconsin	Romney	100.00%	100.00%	44.0283%	0.0000%	1.000000	0.000000
Wisconsin	Santorum	15.02%	40.33%	38.8644%	2.0376%	1.000000	0.000000

Wisconsin	Santorum	30.00%	58.42%	38.3993%	1.5725%	1.000000	0.000000
Wisconsin	Santorum	45.01%	72.46%	38.0362%	1.2094%	1.000000	0.000000
Wisconsin	Santorum	60.00%	82.59%	37.5464%	0.7196%	1.000000	0.000000
Wisconsin	Santorum	75.03%	91.03%	37.3071%	0.4803%	1.000000	0.000000
Wisconsin	Santorum	90.05%	97.11%	36.9906%	0.1637%	1.000000	0.000000
Wisconsin	Santorum	100.00%	100.00%	36.8268%	0.0000%	1.000000	0.000000
Wisconsin	Paul	15.02%	40.33%	12.2130%	1.0614%	1.000000	0.000000
Wisconsin	Paul	30.00%	58.42%	11.8641%	0.7124%	1.000000	0.000000
Wisconsin	Paul	45.01%	72.46%	11.9120%	0.7603%	1.000000	0.000000
Wisconsin	Paul	60.00%	82.59%	11.7126%	0.5609%	1.000000	0.000000
Wisconsin	Paul	75.03%	91.03%	11.5090%	0.3573%	1.000000	0.000000
Wisconsin	Paul	90.05%	97.11%	11.2886%	0.1370%	1.000000	0.000000
Wisconsin	Paul	100.00%	100.00%	11.1517%	0.0000%	1.000000	0.000000
Wisconsin	Gingrich	15.02%	40.33%	6.1330%	0.2971%	0.999999	0.000001
Wisconsin	Gingrich	30.00%	58.42%	5.8961%	0.0602%	0.932882	0.068489
Wisconsin	Gingrich	45.01%	72.46%	5.9514%	0.1155%	0.999963	0.000039
Wisconsin	Gingrich	60.00%	82.59%	5.8850%	0.0491%	0.988873	0.011418
Wisconsin	Gingrich	75.03%	91.03%	5.8897%	0.0538%	0.999806	0.000202
Wisconsin	Gingrich	90.05%	97.11%	5.8505%	0.0146%	0.953160	0.048443
Wisconsin	Gingrich	100.00%	100.00%	5.8359%	0.0000%	1.000000	0.000000
West Virginia	Romney	15.02%	32.56%	67.86%	-1.701%	1.1E-07	1.000000
West Virginia	Romney	30.00%	50.92%	68.17%	-1.389%	2.02E-11	1.000000
West Virginia	Romney	45.01%	65.65%	68.64%	-0.922%	1.000000	0.000000
West Virginia	Romney	60.00%	77.36%	68.88%	-0.675%	1.000000	0.000000
West Virginia	Romney	75.03%	87.11%	69.05%	-0.508%	1.000000	0.000000
West Virginia	Romney	90.05%	95.61%	69.29%	-0.27%	1.000000	0.000000
West Virginia	Romney	100.00%	100.00%	69.56%	0.0000%	1.000000	0.000000
West Virginia	Santorum	15.02%	32.56%	12.7686%	0.6796%	0.998331	0.001810
West Virginia	Santorum	30.01%	50.92%	12.5578%	0.4688%	0.999212	0.000843
West Virginia	Santorum	45.01%	65.66%	12.4137%	0.3247%	0.998774	0.001303
West Virginia	Santorum	60.01%	77.36%	12.3639%	0.2749%	0.999753	0.000265
West Virginia	Santorum	75.10%	87.11%	12.2972%	0.2082%	0.999914	0.000094
West Virginia	Santorum	90.07%	95.61%	12.1824%	0.0934%	1.000000	0.000000
West Virginia	Santorum	100.00%	100.00%	12.0890%	0.0000%	1.000000	0.000000
West Virginia	Paul	15.02%	32.56%	11.8382%	0.7970%	0.999824	0.000195
West Virginia	Paul	30.01%	50.92%	11.6683%	0.6271%	0.999994	0.000006
West Virginia	Paul	45.01%	65.66%	11.5339%	0.4928%	0.999999	0.000001
West Virginia	Paul	60.01%	77.36%	11.3178%	0.2767%	0.999865	0.000146
West Virginia	Paul	75.10%	87.11%	11.2564%	0.2153%	0.999973	0.000030
West Virginia	Paul	90.07%	95.61%	11.1434%	0.1023%	1.000000	0.000000
West Virginia	Paul	100.00%	100.00%	11.0411%	0.0000%	1.000000	0.000000
West Virginia	Gingrich	15.02%	32.56%	6.1646%	-0.1299%	0.232293	0.778124
West Virginia	Gingrich	30.01%	50.92%	6.3443%	0.0498%	0.679061	0.330579
West Virginia	Gingrich	45.01%	65.66%	6.2865%	-0.0080%	0.467515	0.542290
West Virginia	Gingrich	60.01%	77.36%	6.3452%	0.0507%	0.811490	0.195364
West Virginia	Gingrich	75.10%	87.11%	6.3452%	0.0507%	0.892255	0.113127
West Virginia	Gingrich	90.07%	95.61%	6.3416%	0.0471%	0.978748	0.023480
West Virginia	Gingrich	100.00%	100.00%	6.2945%	0.0000%	1.000000	0.000000
Kentucky	Romney	15.03%	34.68%	63.0495%	-3.7199%	0.000000	1.000000
Kentucky	Romney	30.01%	53.76%	64.0151%	-2.7543%	0.000000	1.000000
Kentucky	Romney	45.01%	67.58%	64.9535%	-1.8160%	1.000000	0.000000
Kentucky	Romney	60.03%	79.23%	65.4496%	-1.3198%	1.000000	0.000000
Kentucky	Romney	75.09%	88.58%	65.7681%	-1.0014%	1.000000	0.000000
Kentucky	Romney	90.05%	96.41%	66.3817%	-0.3877%	1.000000	0.000000
Kentucky	Romney	100.00%	100.00%	66.7694%	0.0000%	1.000000	0.000000
Kentucky	Paul	15.03%	34.68%	14.4936%	1.9630%	1.000000	0.000000
Kentucky	Paul	30.01%	53.76%	14.0889%	1.5582%	1.000000	0.000000
Kentucky	Paul	45.01%	67.58%	13.4843%	0.9537%	1.000000	0.000000
Kentucky	Paul	60.03%	79.23%	13.2204%	0.6897%	1.000000	0.000000
Kentucky	Paul	75.09%	88.58%	13.0864%	0.5557%	1.000000	0.000000
Kentucky	Paul	90.05%	96.41%	12.7299%	0.1993%	1.000000	0.000000
Kentucky	Paul	100.00%	100.00%	12.5307%	0.0000%	1.000000	0.000000
Kentucky	Santorum	15.03%	34.68%	9.1865%	0.3145%	0.974930	0.026463
Kentucky	Santorum	30.01%	53.76%	9.1269%	0.2549%	0.993248	0.007103
Kentucky	Santorum	45.01%	67.58%	9.1508%	0.2787%	0.999905	0.000101
Kentucky	Santorum	60.03%	79.23%	9.0741%	0.2021%	0.999883	0.000126

Kentucky	Santorum	75.09%	88.58%	9.0503%	0.1782%	0.999998	0.000002
Kentucky	Santorum	90.05%	96.41%	8.9742%	0.1022%	0.999998	0.000002
Kentucky	Santorum	100.00%	100.00%	8.8720%	0.0000%	1.000000	0.000000
Kentucky	Gingrich	15.03%	34.68%	6.5724%	0.6239%	0.999998	0.000002
Kentucky	Gingrich	30.01%	53.76%	6.2744%	0.3258%	0.999922	0.000085
Kentucky	Gingrich	45.01%	67.58%	6.0814%	0.1328%	0.983884	0.016948
Kentucky	Gingrich	60.03%	79.23%	6.0774%	0.1289%	0.997630	0.002527
Kentucky	Gingrich	75.09%	88.58%	6.0473%	0.0988%	0.998966	0.001119
Kentucky	Gingrich	90.05%	96.41%	5.9883%	0.0397%	0.984627	0.016748
Kentucky	Gingrich	100.00%	100.00%	5.9486%	0.0000%	1.000000	0.000000

VII. Conclusions

Slopes on cumulative vote tally charts, which should settle to horizontal lines, are an amazing statistical anomaly. The hypergeometric distribution chart, normally produces after a minor initial oscillation, a smooth horizontal line for the rest of the chart. By applying this distribution to the 2012 Republican primary election data, we exposed a serious election anomaly, which can be seen as obvious slopes favoring one candidate. It is an extraordinary observation and indicates overwhelming evidence of election manipulation. A massive set of detailed data and analysis for all 50 states, beyond the scope of this paper, also confirmed these unlikely results. These highly anomalous election results indicate a widespread, systematic exchange of votes favoring one candidate.

Statistical analysis of the Republican Primaries results from 2012 in Iowa, New Hampshire, Arizona, Ohio, Oklahoma, Alabama, Louisiana, Wisconsin, West Virginia, and Kentucky show strong statistical evidence of election manipulation[15]. The anomaly subsides somewhat towards the end of the election cycle, when completion is weakened by the earlier election results.

Historically, an early vote gain effect snowballs through the various primary states as it benefits the candidate with momentum as well as additional votes. Mitt Romney, based on our analysis, should have (statistically) gotten third rank in Iowa's election (as opposed to second); second rank in New Hampshire (as opposed to the first rank), and so on, resulting most likely to a brokered convention at the Republican National Convention in Tampa, FL.

Some rather large statistical anomalies in states such Ohio have negatively affected opposing candidates by reducing their momentum and fundraising power. Ohio's election (statistically) should have been earned by candidate Rick Santorum. Rank switching in Oklahoma's election also affected candidates.

The statistical analysis clearly shows that other candidates were supposed to get more votes than the official count. Tests were performed on random samples as well as the entire statistical populations represented by the whole state in each case. These facts assure us that the tests have high statistical power, as well as lack of selection bias. Many individual counties (600+) have been analyzed as well, indicating that this type of election fraud is pervasive.

We urge readers of this paper to reproduce our results and publish their findings.

[15] The remaining 39 states' data are available, but these results are beyond the scope of this paper.

VII. Data Sources

1. US Census Bureau: Census 2000 U.S. Gazetteer Files. County locations: http://www.census.gov/geo/www/gazetteer/places2k.html
2. US Census Bureau: Population, Housing Units, Area, and Density 2010. http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_10_SF1_GCTPH1.ST05&prodType=table
3. Iowa Election Results, January 3, 2012: https://www.google.com/fusiontables/DataSource?dsrcid=2475248
4. New Hampshire Election Results, January 10, 2012: http://sos.nh.gov/2012RepPresPrim.aspx?id=12938
5. Arizona Election Results, February 28, 2012: http://results.enr.clarityelections.com/AZ/36496/75798/en/summary.html http://results.enr.clarityelections.com/AZ/36496/75798/en/md_data.html?cid=105&
6. Ohio Election Results, March 6, 2012: http://www.sos.state.oh.us/sos/elections/Research/electResultsMain/2012results/2012precincts.aspx
7. Oklahoma Election Results, March 6, 2012: http://www.ok.gov/elections/The_Archives/Election_Results/2012_Election_Results/index.html
8. Alabama Election Results, March 13, 2012: http://results.enr.clarityelections.com/AL/38312/86349/en/summary.html http://results.enr.clarityelections.com/AL/38312/86349/en/md_data.html?cid=30&
9. Louisiana Election Results, March 24, 2012: http://staticresults.sos.la.gov/03242012/03242012_Statewide.html http://staticresults.sos.la.gov/03242012/03242012_45596.html
10. Wisconsin Election Results, April 3, 2012: http://gab.wi.gov/elections-voting/results/2012/spring-election-presidential-preference
11. West Virginia Election Results, May 8, 2012: http://apps.sos.wv.gov/elections/results/download.aspx?year=2012&eid=8
12. Kentucky Election Results, May 22, 2012: http://results.enr.clarityelections.com/KY/38672/84521/en/select-county.html
13. http://elect.ky.gov/results/2010-2019/Pages/2012primaryandgeneralelectionresults.aspx http://elect.ky.gov/SiteCollectionDocuments/Election%20Results/2010-2019/2012/2012offpriresults.pdf