Free and Fair?

Investigating the Results of the 2010 Sri Lankan Presidential Election

Ole J. Forsberg

2011

Presentation for the Midwest ISA Conference

Free and Fair?

Investigating the Results of the 2010 Sri Lankan Presidential Election

Abstract:

One requirement of democracy is for elections to be "free and fair." One of the requirements for an election to be "free and fair" is that the probability of the vote counting is independent of other factors, especially ethnicity and candidate support. This requirement can easily be tested in many elections, such as in Sri Lanka.

Sri Lanka ended its 26-year civil war between the Sinhalese and the Tamils in 2009 with a win for the ruling Sinhalese. While elections were held in Sri Lanka during the civil war, participation in the Tamil areas was highly restricted. The first post-war election took place in 2010, pitting incumbent president Mahinda Rajapakse against General Sarath Fonseka.

Using the results of the 2010 Sri Lankan Presidential election, I will test the "free and fair" claim. The methods used will focus on the patterns exhibited in the rejected ballots: Higher proportions of rejected ballots occurred in areas supportive of Fonseka, and lower proportions of rejected ballots occurred in areas supportive of Rajapakse. These patterns are consistent both with an unfair electoral system and with persistent electoral fraud.

Author:

Ole J. Forsberg

Department of Political Science
Department of Statistics
Oklahoma State University
Stillwater, OK 74078

Introduction

In 2009, the civil war in Sri Lanka finally ended. The rebel group in the north of the tiny island State finally succumbed to the overwhelming drive of the South to keep the island united under a Sinhalese rule. This war, which saw the deaths of thousands and the introduction of suicide terrorism, ended with a combination of military acumen and natural disasters. The military victory was led by President Mahinda Rajapakse, as the civilian titular head of the military, and General Sarath Fonseka, as leader of the military. The natural disaster was the 2004 Boxing Day Tsunami, which devastated the Tamil navy anchored on the eastern part of Sri Lanka.

The two allies were close friends throughout the end of the war, a fact that may have led to the military ultimately being successful. However, political differences led Fonseka to enter the political fray in 2009, campaigning for the presidential elections scheduled for January 2010.

Polls showed a close race, which was to be expected, as both were held in high esteem by many in Sri Lanka. Expectedly, Rajapakse and Fonseka polled slightly higher in their respective home district. Beyond that, the BBC predicted that the initial results may not show a clear winner and that the ultimate winner of this election may not receive a majority of the votes cast (BBC 2010).

The run-up to voting day had the typical Sri Lankan violence, several died in clashes between the supporters of Rajapakse and others. Election Day, itself, progressed with the usual intimidation and violence at polling stations. However, since no international observers were allowed to observe the election, the only evidence of these events was the bodies and the claims by rival factions (BBC 2010, CMEV 2010).

The official counting began on schedule, and the election winner was known within a few hours. The results of the election, which were supposed to be quite close, were 40% voting for Fonseka, and 58% for Rajapakse (CMEV 2010).

The next day, Fonseka declared the results invalid because of extensive voting irregularities. He gathered his closest advisors around him in the five-star Cinnamon Lakeside Colombo hotel to determine his plan of action. Later that morning, Rajapakse ordered the hotel surrounded by members of the military, effectively separating Fonseka from the media and the outside world. At first, the military stated they were only there to apprehend deserters. Later, the military stated they were there in response to Rajapakse accusing Fonseka of plotting a military coup, then it was for Fonseka's military offences, and then for running for president while in the military.

In prison, Fonseka was released only long enough to attend parliamentary meetings, as his party, the New Democratic Front, won enough votes for him to sit in parliament (CMEV 2010). The government court-martialed him for committing "military offences," eventually finding him guilty of corrupt military supply deals and sentenced to three years in prison (Jayasinghe 2010).

***

Fonseka's allegations of electoral irregularities remain. Evidence, however, is lacking. Those who perpetrated the irregularities, if they occurred, are unlikely to come forward.

If they happened, the current head of State is unlikely to open an inquiry, as he directly benefitted from them. Is there a way to determine if Fonseka's assertions have any credibility?

Yes. This is the primary purpose of this paper.

Testing Democracy

As with most things in Political Science, the definition of democracy is contentious—some would say vigorously so (Snyder and Samuels 2006). And yet, even with a disagreement on the exact terms of democracy, certain themes are apparent: Elections are held. All adults in the country are free to vote as their conscience dictates. The candidates/parties have equal access to the media. Each person's vote counts the same regardless of characteristics of that person or of that vote.

If we were to categorize these aspects, we may note that the first two reflect the level of freedom in the election; the last two, the level of fairness. In both cases, the term 'free and fair' refers both to the candidates/parties contesting the election and to the votes themselves. Robert Dahl (2000, 2007) holds this last aspect, 'one person, one vote,' as the cornerstone of a true democracy. Unfortunately, as with most pithy statements, there is a frustrating vagueness in its meaning.

One interpretation of 'one person, one vote' is based on representation: each individual holds the same 'strength of voice' in the legislature. Those holding this view tend to assert that a proportional representation system is more democratic than a majoritarian system (Dahl 2007, Farrell 2001).

A second interpretation (although not mutually exclusive of the first) sees 'one person, one vote' as being based on the actual vote counts. An election is fair if each vote has the same probability of being declared invalid, independent of candidate, party, and demographics (e.g., age, gender, and ethnicity). If certain groups of people have a higher weight to their votes, then the election is not fair. Such a condition may come about through electoral fraud (e.g., ballot box stuffing or counting fraud) or through an unfair electoral system (e.g., complex ballots in certain districts or ballot paper shortages in certain districts). In the United States, Mebane and others have documented the effects of these latter two examples: Florida in 2000 (Wand, et al., 2001; Brady 2001; Smith 2002), and Ohio in 2004 (Mebane 2005).

***

In this research, I will define electoral fairness—one person, one vote—in terms of the probability that a person's vote will be counted; that is, an election is fair if everyone has the same probability of having their vote counted.

Note that this aspect of fairness is just one requirement for an election to be truly democratic. As such, if an election fails this test, one can conclude that there is sufficient evidence against the election being free and fair. If, however, an election passes tests based on this aspect of free and fair, then we cannot conclude the election actually was truly democratic—the converse is not logically equivalent to the original categorical proposition.

Now that we know what we need to test, we can use its characteristics to formulate our tests. There are currently three general methods for testing claims of fairness in elections: digit tests, distributional tests, and regression tests. The last includes both simple linear regression and more complex methods specifically designed to handle count and proportion data using robust methods. In the following sections, I will explain the various tests as well as their strengths and weaknesses.

Digit Tests

Let us assume that the logarithms of the raw vote counts are distributed Uniformly. This implies the distribution of the leading digit for each count has a Benford distribution. (The probability mass function for the Benford distribution is provided in Table 1.) For instance, under these assumptions, the probability of a leading 3 digit is 0.125, which is less than half the probability of a leading 1 digit.

[Table 1 about here]

As we have a known distribution for the leading digits, we can perform the usual chi-squared test to test the null hypothesis that the raw vote counts are properly distributed (Pearson 1900, Freund and Wilson 2003).

In addition to the usual first-digit test, any specified digit has a known distribution under the null hypothesis. Mebane (2008, 2010) suggests that the second- and third-digit tests are more appropriate. This is due to the sensitivity of the first-digit test to the distributional assumption. The logarithm of vote counts is not distributed Uniformly. They are close, however. The first-digit test will reject more frequently than it should if the range is not large enough. The second- and third-digit tests are not as sensitive to range length. In other words, the first-digit test is more of a global test than the second- and third-digit tests.

To illustrate his point, let us apply the three Benford tests to the official results of the 2008 US Presidential Election. As we are specifying the usual $\alpha=0.05$ Type I error rate, we would naturally expect two US states to fail each of the tests (although not necessarily the same state). Figure 1 shows which states fail the various Benford tests. Note that seven of the states fail the first-digit test, whereas only two states fail either the second- or third-digit tests. Thus, at our usual level of confidence, we can reject the hypothesis that the first-digit Benford test is appropriate ($p=0.0086$). Figure 2 repeats this analysis, but for the 2004 US Presidential election. The first-digit test rejects the assumption of fairness for 11 states. Again, at our usual level of confidence, we reject the hypothesis that the first-digit Benford test is appropriate ($p<0.0001$).

[Figure 1 about here]

[Figure 2 about here]

Distributional Tests

One aspect of election testing that has not been sufficiently used is testing for multiple data-generating processes. Under the 'free and fair' hypothesis, the process giving rise to ballot invalidation should be the same across different categories of electoral districts. If this is not true, then we can conclude that there is a different invalidation process and that there is sufficient evidence against the 'free and fair' hypothesis.

Two characteristics of (many) distributions are the expected value and the variance. The expected value is the first raw moment. The variance is the second central moment. There exist distributional characteristics based on higher moments, e.g. skew and kurtosis, but the mean and variance are the most well-known, and statistical tests are prevalent for both.

As an illustration, let us return to the 2008 US Presidential election in Colorado. Let us divide the counties into Obama-winning counties and McCain-winning countries. Figure 3 is a boxplot of the proportion of the vote invalidated for each of the two types of counties. The usual statistical tests indicate that the means are not significantly different (W=465; p=0.436) and the variances are not significantly different (F=1.986; df=29,34; p=0.056). Similarly, we could have used the Kolmogorov-Smirnov test (Conover 1971) to conclude that the distributions are not significantly different (D=0.2429; p=0.2514).

[Figure 3 about here]

From these results, we can conclude that there is no evidence of a second data-generating process for vote invalidation. Thus, there is no *prima facie* evidence of electoral unfairness in the 2008 US Presidential election in Colorado—as expected.

Regression Tests

The fairness aspect of the 'free and fair' assumption can be numerically tested, provided that the data exists. Recall that one of the requirements for an election to be free and fair is that votes are not differentially counted: each person's vote has the same probability of being counted, independent of candidate/party choice and demographics.

The independent variable is the proportion of the vote in favor of a specific candidate or party. The dependent variable is the proportion of the vote declared invalid. The unit of analysis may ideally be the precinct, but may be any aggregation level as long as the sample size is sufficient and as long as the number of invalidated ballots is greater than five. The former requirement is due to the need for testing power. The latter requirement is due to the sample size requirement for the Normal approximation for the Binomial distribution.

Simple Linear Regression

Simple linear regression is a modeling technique that finds the relationship between two continuous variables. Under ordinary least squares (OLS) fitting, there are several important assumptions. The main assumption is that the dependent variable is Normally distributed, conditional on the value of the independent variable. Two important aspects of this requirement are that the dependent variable is continuous and unbounded.

Proportion data is continuous. It is not unbounded, however. As such, a transformation must be performed to transform the dependent variable into an unbounded variable. There are many appropriate transformations available. The logit transform is the usual, however. It has two important characteristics: it maps a continuous variable onto a continuous variable, and it maps a variable bounded by 0 and 1 (exclusive) onto an unbounded variable.

Thus, the first alteration to the simple linear model is to transform the dependent variable by applying the logit transformation. The second alteration is to use a robust fitting technique. As our research hypothesis is that different categories of ballots are treated differently, there are really two processes at work in the model. Because the data may be generated from two different processes, non-robust fitting methods will be biased (Mebane and Sekhon 2004).

Robust Regression

Robust regression is an umbrella term for a variety of modeling techniques that attempt to fit the majority of the data without being adversely influenced by the minority. Several robust techniques exist for linear regression. Some methods merely adjust for the presence of heteroskedasticity (Huber 1964; Huber 1973; White 1980; White 1982). Others utilize a special class of estimators (Huber 2009; Mebane and Sekhon 2004). Others fit the model using the median (or other quantile) in lieu of the mean (Hao and Naiman 2007; Koenker and Hallock 2001).

Mebane and Sekhon (2004) suggest a multinomial model that utilizes hyperbolic tangent regression. This is based, in part, on Huber (1981), which recommends using M-estimators. As Mebane and Sekhon's suggestion is equivalent to Huber's when the dependent variable has a multinomial distribution, the choice between the two heavily depends on the form of the dependent variable. As the dependent variable in this research is a proportion, I use Huber's recommendation.

***

Thus, for this research, I will diverge from the path of Mebane and Sekhon and transform the dependent variable (the proportion of the vote declared invalid in the electoral district) using the logit transform, and then apply Huber's MM-estimator to model the variable appropriately.

Hypothesis

I will test the hypothesis that the 2010 Presidential election in Sri Lanka was free and fair. As mentioned previously, the fairness requirement implies, among other things, that there is no relationship between "the probability that a vote is declared invalid" and "the candidate for whom the ballot is cast." As such, we can statistically test this hypothesis in several manners: digit tests, distribution tests, and regression tests.

Data and Variables

The Sri Lankan Department of Elections publishes the election results on its official website. The election counts are aggregated at the electoral division level. In Sri Lanka, there are nine provinces (first-level units), 23 districts (second-level units), and 160 electoral divisions. The data is collected at the precinct level and aggregated at the electoral division level. In addition to the electoral divisions, each district counts votes sent through the mail (*postal*) and votes cast by displaced persons (*displaced*) separately. In most cases, the counts for one or both are too small to include in any reliable analysis. To avoid creating bias, all postal and displaced vote totals were removed from consideration.

The electoral division is sufficiently large to allow for the Central Limit Theorem to eliminate issues of non-Normality. The number of votes cast in the smallest electoral division, Maskeliya, Nuwara-Eliya in the Central Province, was 6566 out of the 36,940 registered voters.

***

Since we are examining the process that gave rise to invalidated votes, the dependent variable will be the proportion of the vote declared invalid by the Sri Lankan Department of Elections branch in each electoral division. The Sri Lankan Department of Elections supplied this data.[1] The invalidation rate ranged from a minimum of 0.00477 to 0.04422. The former rate occurred in Galgamuwa, Kurunegala, North Western Province (which supported Rajapakse); the latter, in Chavakachcheri, Jaffna, Northern Province (which supported Fonseka).

The Three Tests

Echoing the previous discussion, I will run three tests of the null hypothesis: digit tests, distribution tests, and (robust) regression tests. The first tests the digit count using the Benford model. The second tests the hypothesis at the provincial level, by comparing the distribution of the invalidation rate at the electoral division level for provinces won by Rajapakse versus those lost by Rajapakse. Finally, the third will test the relationship between the proportion of the vote invalidated and the proportion of the vote in favor of Rajapakse in the electoral division.

Test: The Benford Test

The Benford test compares the reported digits of the vote counts with a hypothesized digit distribution. The Benford test is the least reliable of the three tests as there is no a priori reason to believe that vote counts should be distributed in any specific manner. However, its popularity suggests including it.

[1] Sri Lankan Department of Elections home page: http://www.slelections.gov.lk/. The data is available at http://www.slelections.gov.lk/presidential2010/province.html.

According to the official results, the first-digit Benford test indicates severe departure from the expected digit distribution (B=93.86; p << 0.0001). The second- and third-digit Benford tests, however, indicate no significant departure form expectation (B=8.98, p = 0.344, for second-digit, and B=5.91, p=0.658, for third). As mentioned earlier, the first-digit should be of most concern to us as that digit affects the election outcome much more than does an altered second digit. Because of this, the Benford test indicates electoral fraud.[2]

Interestingly enough, in the provinces won by Fonseka (district numbers 4, 5, 6, 7, and 8), the first-digit Benford test only rejects for one district (district 7). In the provinces won by Rajapakse, the Benford test rejects for eight of 17 districts. The former rate (20%) is not statistically different from what we would expect were the test appropriate (p=0.204). The latter *is* statistically different from what we would expect (p<<0.0001).

[Figure 4 about here]

Figure 4 shows the distribution of the p-values for each of the provinces and for each of the three Benford tests, with a horizontal line at our usual alpha=0.05 level. Note that the second- and third-digit tests both have an appropriate number of failures under the null hypothesis (approximately 5%). The first-digit test, however, indicates nine failures across the State—eight in Rajapakse-leaning provinces.

[2] As mentioned earlier. The Benford test assumes that vote counts have a specific distribution. There is no *a priori* reason for vote counts to be distributed in such a manner. As such, conclusions based solely on the Benford test are naturally suspect.

These results do not support the 'free and fair' hypothesis. The high number of failing Rajapakse-leaning provinces indicates that the reported vote counts are not in accord with the supposed vote-count distribution.

Test: The Distribution of Invalidated Votes

As mentioned above, the invalidation rate in the electoral divisions ranges from 0.00477 to 0.04422. Neither the proportions nor their logits are distributed as Normal random variables, according to the Shaprio-Wilks test (W=0.8602; p<<0.0001). In fact, it appears as though there are two separate populations involved.

Mahinda Rajapakse won 120 of the 160 electoral divisions and seven of the nine provinces. If the election is fair, then there should be no significant difference in invalidation rates between the provinces Rajapakse won and those he lost. In fact, if the processes generating invalid votes are the same, then there should be no difference in distributions between the two. Figure 5 is a boxplot of the invalidation rate for the two categories of provinces. Note that it appears as though there are significant differences between the distributions of invalidated votes in the two. The means (and medians) and the variances appear to be significantly different.

[Insert Figure 5 about here]

The Mann-Whitney non-parametric test indicates a significant difference between the medians of the two categories of provinces (W=3011; p<<0.0001). Furthermore, Fisher's F-test indicates that there is a significant difference in the two variances (F=9.9239; df=23, 135; p<<0.0001). Furthermore, the Kolmogorov-Smirnov test also indicates that the two samples have significantly different distributions ($D=0.7966$; $p \ll 0.0001$).

Thus, we can strongly conclude that the distribution of invalidated votes in the two types of provinces is significantly different. In other words, we can conclude that there are different processes giving rise to invalidated votes in the two. This also constitutes *prima facie* evidence that the election was not free and fair.

Test: Robust Regression

The third and final test will be robust regression on the logit-transformed dependent variable. The dependent variable remains the invalidation rate in the electoral division. The independent variable will be the proportion of the vote in favor of incumbent President Mahinda Rajapakse. If the 'free and fair' hypothesis is correct, then there should be no statistically significant relationship between the two; that is, the effect of proportion in favor of Rajapakse should not be statistically significant.

As mentioned earlier, there are several options when running a regression. As we have one dependent variable and one independent variable, we could use simple linear regression. However, this method assumes that the dependent variable is Normally distributed, conditional on the value of the independent variable. This is not true, here; zero and one bound the dependent variable.

A second option is to use a generalized linear model, using the binomial family and a logit (or similar) link function. This is perfectly acceptable if the sample comes from a single population. If not, then the effect estimates will be biased—perhaps severely so (Mebane and Sekhon 2004).

Mebane and Sekhon (2004) suggest a third option: using a robust regression technique with a high breakpoint. Specifically, they recommend using the hyperbolic tangent estimator. Its advantage is that it has the maximum breakpoint, which means it is very robust to violations of the assumption of a single population.[3] This method is appropriate if the dependent variable is a count.

Alternatively, one can use robust MM-estimators as suggested by Huber (1981). Mebane and Sekhon's process is a formulation of Huber's MM-estimators. The solution I employ is to transform the dependent variable (invalidation rate) using the logit function, and then use robust regression based on Huber's MM-estimator. This has the advantages of being robust to outliers as well as not violating the underlying assumptions, without requiring special software to fit. Furthermore, it should not suggest different conclusions than does the Mebane-Sekhon solution.

[Insert Table 2 about here]

Table 2 provides the results of the robust regression. Under the 'free and fair' hypothesis, none of the independent variables should be statistically significant. Here, we see that the level of support for Rajapakse in the electoral division is statistically significant ($b=-2.733$, $p<<0.0001$). This indicates that there is a higher probability for a Fonseka-supporting ballot to have been invalidated than for a Rajapakse-supporting ballot. This is a clear violation of the fairness hypothesis.

[3] Mebane and Sekhon created the R package `multinomRob` for this very purpose.

In addition to this result, the fact that Rajapakse won the province is statistically significant (b=-0.297; p=0.0004). This indicates that the probability of a vote being declared invalid is significantly lower in provinces won by Rajapakse than in provinces won by Fonseka. In a fair election, this result should also not exist.

[Insert Figure 6 about here]

Figure 6 shows the above results. The scatterplot is of the invalidation rate against the support for Rajapakse in the electoral division, with the winner of the province indicated. The prediction curve, and the symmetric 95% confidence band, indicates the significant—both statistically and practically—relationship between the invalidation rate and the Rajapakse vote. Thus, the table and the graph both strongly indicate severe deviation from the 'free and fair' hypothesis.

Conclusions

This paper dealt with applying and extending current practices in Electoral Forensics to the 2010 Sri Lankan Presidential election. The incumbent won the election, the challenger claimed electoral deficiencies, and the incumbent arrested the challenger. President Mahinda Rajapakse has yet to open an inquiry into the allegations of electoral impropriety. This paper sought to offer a first glance.

All three categories of tests on the official electoral returns indicated deviation from the 'free and fair' hypothesis. The digit tests indicated that the first-digits are not distributed as expected. There are, in fact, far too many 3s and 4s reported for the digits to be appropriately distributed. This *may* indicate electoral fraud in the ballot counting.

The distribution tests indicated that the process of vote invalidation differed between those provinces won by Rajapakse and those won by Fonseka. The Rajapakse-leaning provinces had a significantly lower invalidation and a significantly lower variance on the invalidation rate than did Fonseka-leaning provinces. This result does not necessarily point to electoral fraud. As Fonseka had his greatest levels of support in those provinces with the highest concentration of Tamils, the difference in invalidation processes may be due to an inherent bias against Tamils in the electoral process.

This is difficult to test, however. The Tamil concentration and the support for Fonseka are highly correlated. Thus, it will be difficult to separate the two effects. Furthermore, the Tamil concentration is not well-documented. The Sri Lankan census division did ask ethnicity, but only in four provinces. Thus, any test would have low power.

Finally, the regression test also indicated strong relationship between invalidation rate and support for Rajapakse at the electoral division level. Additionally, there was a statistically significant effect of whether Rajapakse won the province. Both of these statistically significant effects indicate violation of fairness at the candidate level. Those people voting for Fonseka had a significantly—statistically and substantively—lower probability of having their vote counted than those who voted for Rajapakse.

***

These findings, *en masse*, provide sufficient evidence that the election was *not* free and fair. At this point, little can be done to fix the problems of 2010 in time for 2010. The challenge, however, will be to fix the problems of 2010 in time for the next Presidential election—2016. As the incumbent president has no incentive to alter an electoral process that gave him the presidency (again), external pressure will be needed. India, long a power in the region and a player in the Sri Lankan civil war, has the ability to encourage the necessary reform.

However, India has its own troubles. Unfortunately, as long as Sri Lanka remains quiet, there appears to be little chance that external actors will spend enough resources to force a change.

References

Henry E. Brady, Michael C. Herron, Walter R. Mebane Jr., Jasjeet Singh Sekhon, Kenneth W. Shotts, and Jonathan Wand (2001). "Law and Data: The Butterfly Ballot Episode." *PS: Political Science and Politics* 34(1): 59-69.

BBC (2010). "Sri Lanka presidential votes being counted." BBC News. (Posted January 26, 2010) http://news.bbc.co.uk/2/hi/south asia/8478386.stm (Accessed January 26, 2010).

Centre for Monitoring Election Violence (2010). "Final report of election related violence and malpractices: Presidential Election 2010." (26 July 2010) http://cmev.wordpress.com/2010/07/26/

William J. Conover (1971). *Practical Nonparametric Statistics*. New York: John Wiley & Sons.

Robert A. Dahl (2000). *On Democracy*. Yale University Press.

Robert A. Dahl (2007). *On Political Equality*. Yale University Press.

David M. Farrell (2001). *Electoral Systems: A comparative introduction*. Palgrave.

John Fox and Sanford Weisberg (2010). *An R Companion to Applied Regression, Second Edition*. Sage Publications, Inc.

Rudolf J. Freund and William J. Wilson (2003). *Statistical Methods, Second Edition*. Academic Press.

Lingxin Hao and Daniel Q. Naiman (2007). *Quantile Regression*. Sage Publications, Inc.

Peter J. Huber (1964). "Robust Estimation of a Location Parameter." *Annals of Mathematical Statistics* 35(1): 73-101.

Peter J. Huber (1967). "The behavior of maximum likelihood estimation under nonstandard conditions." *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,1*, LeCam and Neyman, editors. University of California Press: 221-233.

Peter J. Huber (1973). "Robust Regression: Asymptotics, Conjectures and Monte Carlo." *Annals of Statistics* 1(5): 799-821.

Peter J. Huber (1981). *Robust Statistics*. John Wiley & Sons Inc.

Amal Jayasinghe (2010). "Sri Lanka jails ex-army chief Fonseka." Agence French-Presse. September 17, 2010.

Roger Koenker and Kevin F. Hallock (2001). "Quantile Regression." *Journal of Economic Perspectives* 15 (4), 143–156.

Walter R. Mebane, Jr. (2004). "The Wrong Man is President! Overvotes in the 2000 Presidential Election in Florida." *Perspectives on Politics* 2(3): 525-535.

Walter R. Mebane, Jr. (2005). "Voting Machine Allocation in Franklin County, Ohio, 2004: Response to U.S. Department of Justice Letter of June 29, 2005." Posted: February 11, 2006. http://www.umich.edu/~wmebane/franklin2.pdf. Accessed: November 6, 2011.

Walter R. Mebane, Jr. (2007). "Election Forensics: Statistical Interventions in Election Controversies." Presentation at the Annual Meeting of the American Political Science Association 2007.

Walter R. Mebane, Jr. (2008). "Election Forensics: Outlier and Digit Tests in America and Russia" Prepared for presentation at The American Electoral Process conference, Center for the Study of Democratic Politics, Princeton University, May 1-3, 2008.

Walter R. Mebane, Jr. (2010). “Fraud in the 2009 Presidential Election in Iran?” *Chance* 23(1):6-15.

Walter R. Mebane, Jr., and Kirill Kalinin (2009). “Comparative Election Fraud Detection.” Presentation at the Annual Meeting of the American Political Science Association 2009.

Walter R. Mebane, Jr., and Jasjeet Sekhon (2004). “Robust Estimation and Outlier Detection for Overdispersed Multinomial Models of Count Data.” *American Journal of Political Science* 48(2): 392-411.

Karl Pearson (1900). “On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling.” *Philosophical Magazine, Series 5* 50(302): 157-175.

R Development Core Team (2010). “R: A language and environment for statistical computing.” R Foundation for Statistical Computing. Vienna, Austria.

Richard Snyder and David Samuels (2006). “Devauling the Vote in Latin America.” In *Electoral systems and democracy*, Larry Jay Diamond and Marc F. Plattner, eds. Johns Hopkins University Press.

Jonathan N. Wand, Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Jr., Michael C. Herron, and Henry E. Brady (2001). “The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County.” *American Political Science Review* 95(4): 793-810.

Richard L. Smith (2002). “A Statistical Assessment of Buchanan's Vote in Palm Beach County.” *Statistical Science* 17(4): 441-457.

William N. Venables and Brian D. Ripley (2002). *Modern Applied Statistics with S, Fourth Edition*. Springer.

Halbert White (1982). “Maximum likelihood estimation of misspecified models.” *Econometrica* 50(1): 1-25.

Halbert White (1980). “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica* 48(4): 817-838.

Digit	1	2	3	4	5	6	7	8	9
Probability	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

Table 1: The probability mass function for the Benford Distribution. Note that the probability of a specific leading digit declines as the digit value increases.

	Estimate	Std. Error	t-value	p-value
Constant term	-2.891	0.080	-35.941	<< 0.0001
Rajapakse Support	-2.733	0.205	-13.350	<< 0.0001
Province won by Rajapakse	-0.297	0.083	-3.589	0.0004

Table 2: Robust regression results for the election outcome. The dependent variable is the logit of the invalidation rate. There is, however, an effect for the level of support in the electoral division for Rajapakse *in addition to* the effect of Rajapakse winning the province. Note that the interaction model indicated no significant interaction between the two independent variables (p=0.1900). As such, I did not report it.

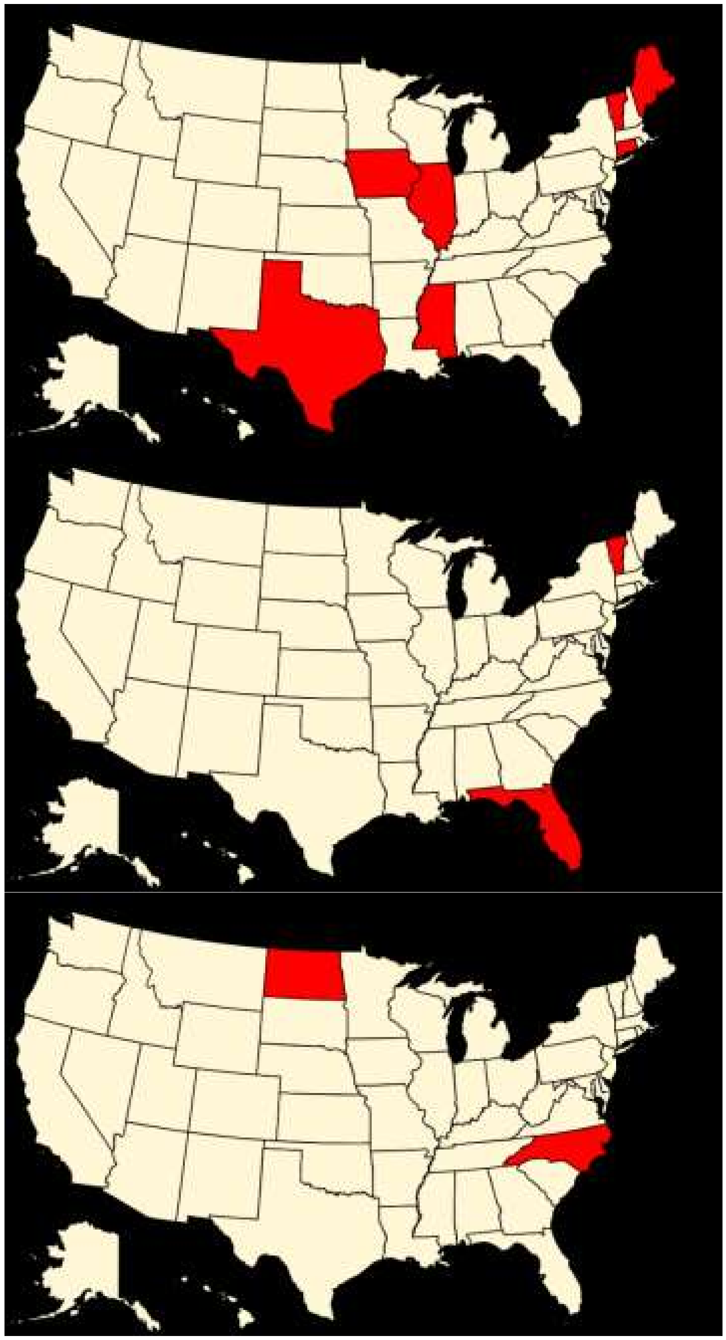

Figure 1: Maps of the 2008 US Presidential election. Red-shaded states fail the respective Benford test (The top map is of the first-digit test; the middle, second-; the bottom, third-digit). Note the excessive number of rejections for the first-digit Benford test.

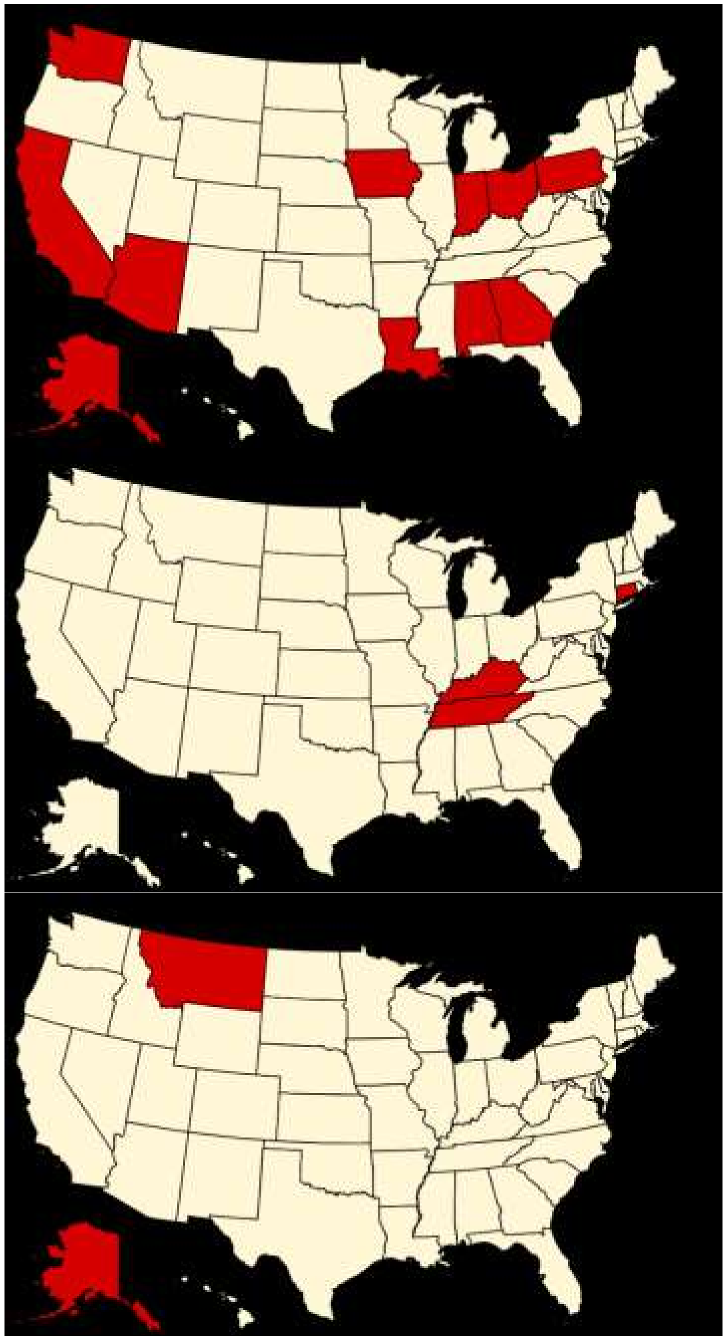

Figure 2: Maps of the 2004 US Presidential election. Red-shaded states fail the respective Benford test (The top map is of the first-digit test; the middle, second-; the bottom, third-digit). Note the excessive number of rejections for the first-digit Benford test.

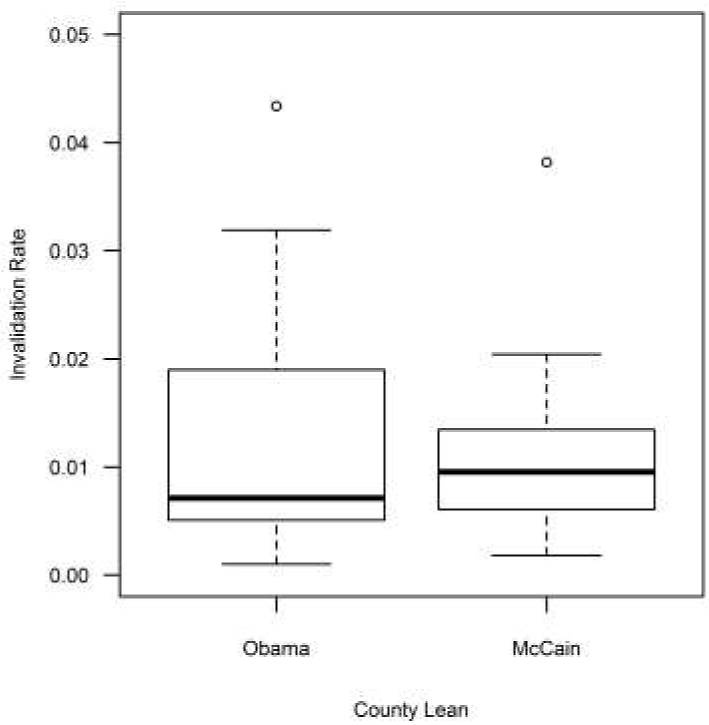

Figure 3: Boxplot of the invalidation rate for two types of counties in the 2008 US Presidential election in Colorado: Those counties supporting Obama and those counties supporting McCain.

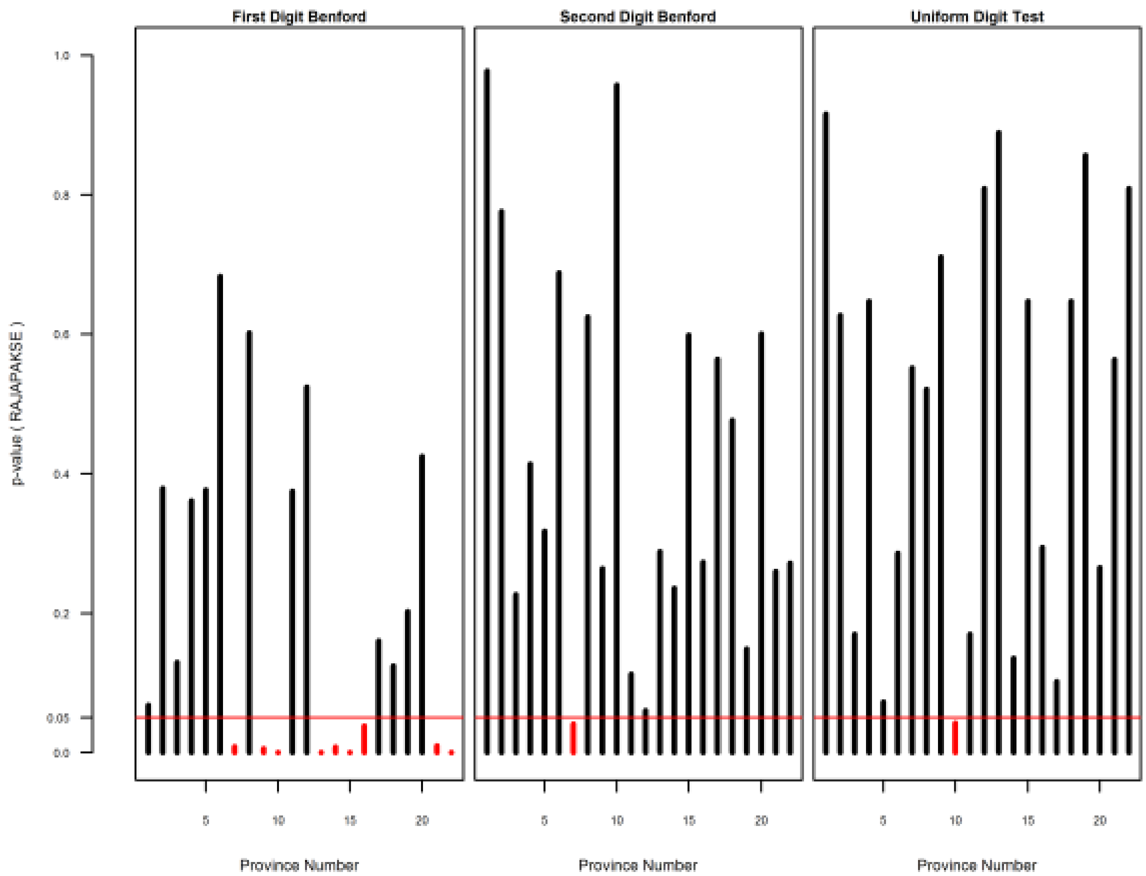

Figure 4: Rejections and p-values for each of the Sri Lankan provinces and each of the three Benford tests. Note the high rejection rate in the first-digit Benford test.

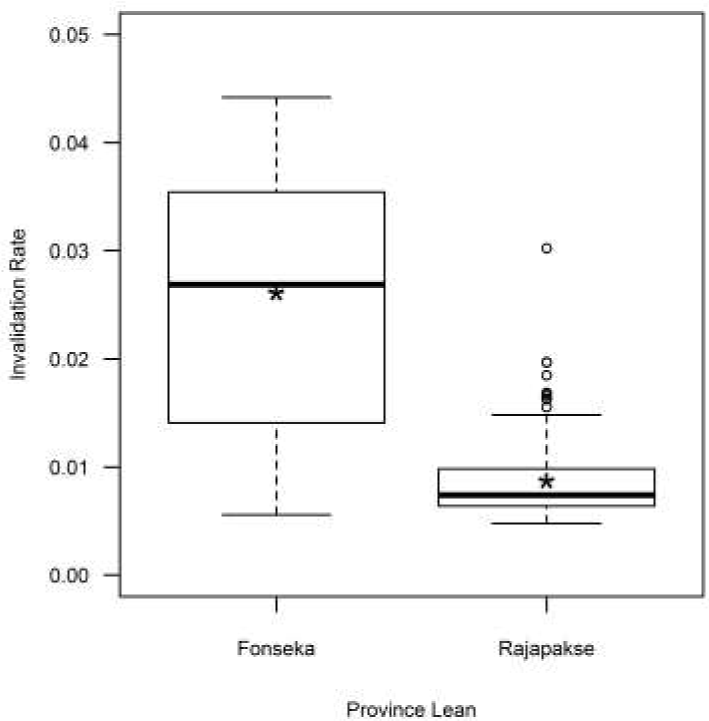

Figure 5: Boxplot of the invalidation rate for two types of provinces in the 2010 Sri Lankan Presidential election: Those counties supporting Rajapakse and those counties supporting Fonseka. The two means are marked with an asterisk.

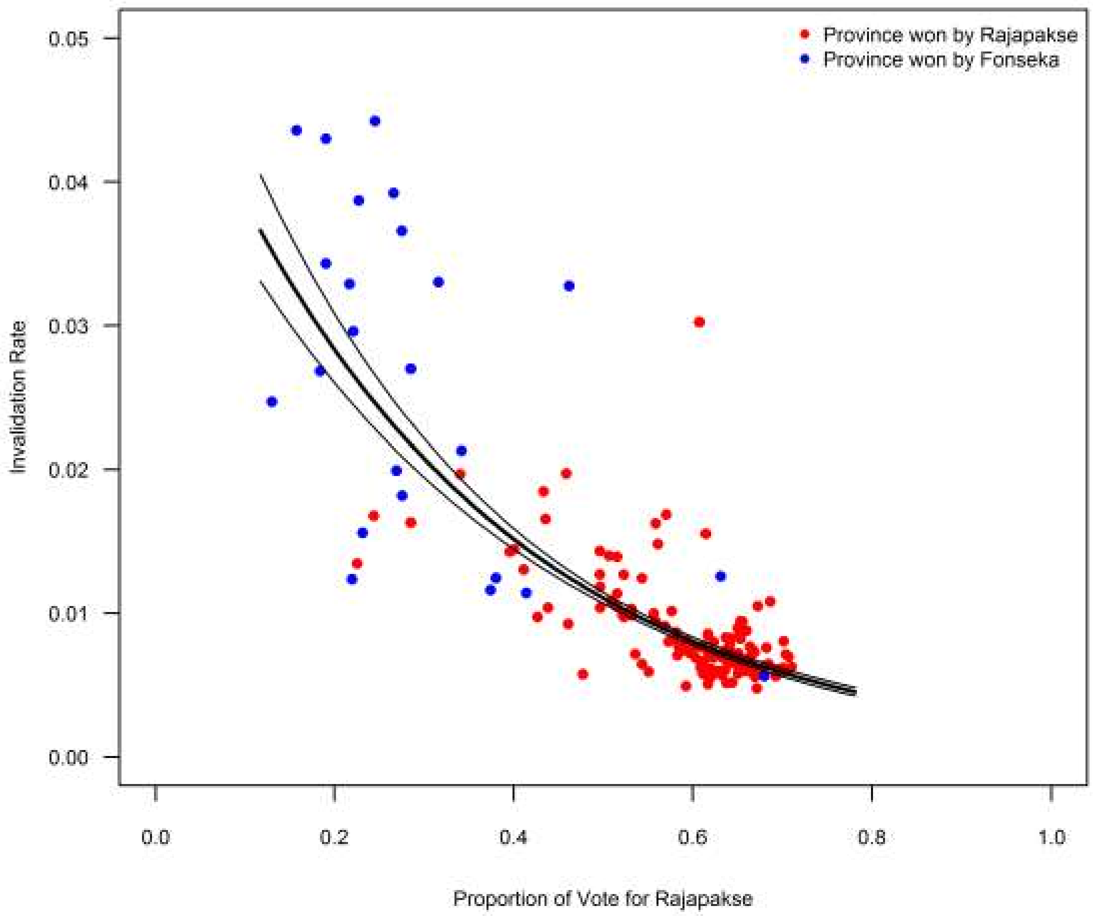

Figure 6: Scatterplot of the invalidation rate against the proportion of the vote in favor of Rajapakse in the electoral division. Also indicated is the winner of the province containing each electoral division. The thick curve is the prediction curve. The thinner lines are the 95% confidence band for those predictions. As a horizontal line cannot fit between the confidence bands, we can conclude, with 95% confidence, there is a statistically significant relationship between the proportion of vote for Rajapakse and the invalidation rate in the electoral division.