Election Forensics: Statistical Interventions in Election Controversies *

Walter R. Mebane, Jr.†

August 28, 2007

Abstract

Controversies about elections are not exactly commonplace, but neither are they rare. In recent years a suite of statistical tools has been developed for diagnosing election anomalies and possibly detecting election fraud. One set of tests focus on patterns in the digits of reported vote counts. I compare the results of such testing to the judgments reached by various sets of election monitors and observers, focusing on four elections: Ohio in the U.S. presidential election in 2004; the presidential election in Mexico in 2006; the 2001 General Election in Bangladesh; and the 2004 presidential election in Indonesia. The statistical tests broadly but not completely agree with observers' conclusions.

*Prepared for presentation at the 2007 Annual Meeting of the American Political Science Association, Chicago, August 30–Sept 2. I thank Owen Lippert for Bangladesh vote count data, Susan Hyde for Indonesia vote count data, and Tom Scherer, Ariel White and Gideon Weissman for assistance.

†Professor, Department of Political Science and Department of Statistics, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (Phone: 734-763-2220; Fax: 734-764-3522; E-mail: wmebane@umich.edu).

Introduction

In recent years the practice of election monitoring has become increasingly institutionalized, with professional personnel in several organizations developing more or less codified standards (e.g. European Union 2002; OSCE ODIHR 2005a). In view of the wide range of circumstances and electoral systems in which elections are conducted, the opportunities for problems and even fraud are many and varied (Lehoucq 2003). It is hardly surprising that the burgeoning field of election monitoring has many gaps, limitations and imperfections (Bjornlund 2004).

One limitation is that in general it is not possible to say whether the officially announced election results are accurate. This flat statement of course calls for qualification. No election is perfect, so the first problem is there is no generally agreed standard to define when the inevitable imperfections are inconsequential. An intuitive criterion for accuracy is that taken all together the imperfections could not possibly have changed the election outcome. Election monitors may have long had such an idea in mind, but the familiar norm that an election should be "free and fair" does not include the idea. Many elections that are declared adequate according to that norm are demonstrably likely to have an inaccurate result. For instance, a European Union Election Observation Mission (EU-EOM) study of the 2001 election in Bangladesh concluded "that the electoral process has guaranteed sufficient conditions of freedom and fairness and represents an important step towards democratic consolidation," notwithstanding "several crucial problems" including "politically motivated violent incidents" and flaws in election administration (European Union 2001). The losing party in that instance challenged the outcome to such an extent that they boycotted the government.

The second problem is that even if everyone were to agree that it is important to verify the accuracy of official results, generally applicable methods for doing so are hard to come by. An intuitive criterion that may guide such verification is the idea that all votes are counted as they were intended to be cast. Immediately this idea runs into the problem of voter suppression: some people may have wished to cast votes but were prevented from doing so. Beyond that is second-order suppression: given existing conditions some do not wish to vote, but they would have done so had conditions been less adverse. This connects with one of the problems that concerned the EU-EOM in Bangladesh in 2001: various practices "contributed to create a heated atmosphere in the country and hindered the development of a peaceful election process. These practices have affected, in some cases, the freedom of choice of the voters" (European Union 2001).

For the votes that were cast, the challenge is to verify that all and only those votes are used in the correct way to allocate the electoral offices. Sometimes this is as simple as deciding which candidate has the most votes, but in many cases it is much more complicated. There may be a proportional representation system that requires extensive calculations to implement, or a mix of personal and party votes that need to be combined, or transferable votes that need to be allocated. Given the detailed election returns, it is important to verify that correct winning candidates or the correct assigned legislative seat shares are announced.

But before that it is important to verify the correctness of the election returns, ideally resolved to the lowest possible level of vote aggregation. Methods such as the quick count (Estok, Nevitte, and Cowan 2002) are intended to help with this. In a quick count, the vote tabulations at a sample of individual polling stations are accumulated using an organization put together outside of the official network. These independently reported counts may then be compared to the officially reported ones. Schemes for precinct-level or voting machine-level post-election audits can have similar motivations (Holt 2007; Norden, Burstein, Hall, and Chen 2007). Feasibility is a primary obstacle to such efforts. For a quick count, for instance, it may be impossible to put together an organization that can report the tabulations from a suitable sample of polling stations. Legal barriers or administrative incapacity may prevent timely and credible post-election audits.

In this paper I begin to consider how well statistical tests that directly examine the officially reported vote counts may help supplement other approaches. In particular I am concerned with tests that focus on the vote counts' second digits. The tests ask whether the frequency distribution of the digits matches the distribution that is implied by Benford's Law (Hill 1995). I refer to that distribution as the 2BL distribution. Such tests have been applied to try to help detect election anomalies at least since they were used in connection with controversy regarding the Venezuelan recall election of 2004 (Pericchi and Torres 2004). Since then, extensive simulation efforts have suggested that vote counts at suitably low levels of aggregation may often be expected to be 2BL-distributed, and that some patterns of systematic manipulation of the counts will cause departures from that distribution (Mebane 2006). Extensive examinations of data from elections in the United States and elsewhere show wide conformity with the 2BL-distribution in vote counts at the precinct level but not at lower levels of aggregation such as voting machines or ballot boxes (Mebane 2006, forthcoming, 2007b). In some cases departures from the 2BL distribution may be explained as natural consequences of the particular configuration of candidate support proportions and precinct sizes in a jurisdiction (Mebane 2007c).

The discussion here is highly preliminary. The approach I take is selective and inductive. I look briefly at four examples where elections were either monitored or closely examined, and I consider whether the indications from testing vote counts' second digits agree with the sense of the evaluations. The examples are voting in Ohio in the U.S. presidential election in 2004, the presidential election in Mexico in 2006, the 2001 General Election in Bangladesh and the 2004 presidential election in Indonesia.

2BL Tests Compared to Evaluations

The test I use is based on a simple chi-squared statistic that compares the observed frequency of the second digits of a set of vote counts to the 2BL expected proportions. The 2BL test statistic is $X^2_{2BL} = \sum_{j=0}^{9}(n_j - Nr_j)^2/(Nr_j)$, where N is the number of precincts having a vote count of 10 or greater (so there is a second digit), n_j is the number having second digit j and r_j denotes the proportion expected to have second digit j according to the 2BL distribution, where $(r_0, \ldots, r_9) = (.120, .114, .109, .104, .100, .097, .093, .090, .088, .085)$. If an observed test statistic is larger than critical values identified by statistical theory, then the conclusion is that the 2BL distribution does not characterize the referent vote counts. Using the chi-squared distribution with nine degrees of freedom for a test of no departure from the expected values gives a critical value of 16.9 for a test at level $\alpha = .05$. In general there are many test statistics to consider, because there are several candidates or there are multiple jurisdictions for each of which we wish to compute the test separately. In such instances we need to make adjustments for having multiple, simultaneous tests. One way to do this is to control the false discovery rate (FDR) (Benjamini and Hochberg 1995) over an entire set of statistics of interest. In this case a test statistic needs to be larger than 16.9 for there to be a significant result at level $\alpha = .05$.

Ohio 2004: During the 2004 U.S. presidential election, Ohio was a closely contested state of pivotal importance for the election outcome. An Election Observation Mission of the Organization for Security and Co-operation in Europe (OSCE) evaluated the election process throughout the country that year and found that "The 2 November 2004 elections in the United States mostly met" international standards for democratic elections, but also it noted a number of problems (OSCE ODIHR 2005b, 1). Among the concerns, the OSCE report particularly noted a number of legal challenges and other allegations that arose in Ohio. After the election, other observers found that election administration in Ohio was replete with irregularities (House Judiciary Committee Democratic Staff 2005). Among other problems, about three percent of would-be voters in the state were prevented from voting due to inadequate provision of voting machines and poor administration, while another one percent of votes were lost due to use of obsolete and error-prone voting machines (Voting Rights Institute 2005).

An examination of 2BL test statistics computed using precinct-level vote count data for each of more than 1,700 U.S. counties in 2004 shows that significantly large statistics occur in relatively few states (Mebane forthcoming), but the 2BL test results from Ohio strongly suggest there are problems there. Table 1 shows the test statistic values for the 19 of Ohio's 88 counties that have $X^2_{2BL} \geq 16.9$ for the precinct vote totals reported for either Kerry or Bush.[1] Even though only one of the 176 statistics for the two candidates is significant when the FDR is controlled—this is the value $X^2_{2BL} = 42.7$ for Kerry's vote counts in Summit County—in all there are 21 statistics greater than 16.9, which is about twelve percent of the statistics. That rate is more than double the nominal test level of $\alpha = .05$. This suggests that deviations from the 2BL distribution are not limited to just Summit County.

*** Table 1 about here ***

Indeed, to some extent the 2BL statistics track the pattern of problems highlighted in other investigations. In many respects, the most extensive problems identified in Ohio in 2004 happened in Cuyahoga County. The county presented the greatest number and range of anomalies in Mebane and Herron (2005). In Cuyahoga, due to confusion related to multiple ballot orders being used at the same polling location, many votes were cast for one candidate but recorded for another (Solov and Suchetka 2004; Associated Press 2004; Jacobs 2007). In Table 1, Cuyahoga is one of only three counties that have $X^2_{2BL} > 16.9$ for both of the candidates. Another county with $X^2_{2BL} > 16.9$ for both candidates is Summit, which is adjacent to Cuyahoga. But other Ohio counties that had demonstrable problems do not have especially large 2BL statistics for the major party presidential candidates. For instance, Franklin County was notorious for having an inadequate supply of voting machines, and the consequent great delays in voting demonstrably reduced voter turnout (Mebane 2005). But the 2BL test statistics for Kerry and Bush in Franklin County are respectively $X^2_{2BL} = 14.0$ and $X^2_{2BL} = 6.0$. The 2BL test does significantly indicate problems with the precinct vote counts in Ohio, but it does not appear to be a particularly useful guide for localizing all of them.

Mexico 2006: The dominant feature of the 2006 election in Mexico is that the vote for president was close and the outcome extremely controversial. The margin of victory was just over one-half of one percent. Many groups observed the 2006 election in Mexico, and most declared the election process basically sound while identifying various problems. The results for president

[1]The precinct data for Ohio in 2004 were collected as part of the DNC study (Mebane and Herron 2005).

were officially challenged. As an EU-EOM report noted, the two party coalitions that received the most votes filed hundreds of challenges that alleged election day irregularities (European Union 2006, 42–43). As part of its response to the challenges, the election court ordered a manual recount of the ballots in 11,839 of the election's 130,788 polling stations or *casillas*, including casillas from 26 of Mexico's 32 states (European Union 2006, 46). The recount resulted in thousands of votes being annulled (European Union 2006, 47).

Five party coalitions ran candidates in the presidential race. The winning candidate was Felipe Calderón of the *Partido Acción Nacional* (PAN), which was the party of the incumbent president, and the candidate receiving the second largest number of votes was Andrés Manuel López Obrador of the *Coalición por el Bien de Todos* (PBT). The third-place finisher was a coalition called *Alianza por México* (APM) that joined the *Partido Revolucionario Institucional* (PRI), which for long time had been the ruling party, and the *Partido Verde Ecologista de México*. Finishing fourth was the *Alternativa Socialdemócrata y Campesina* (ASDC). Fifth was *Nueva Alianza* (NA), a party that was formed as a splinter from the PRI.

Computing 2BL test statistics for each coalition's vote counts at the *seccion* level reveals many significant departures from the 2BL distribution.[2] Table 2 reports the statistics for the secciones in each state. Of the 160 statistics in the table, 48 are significant at the $\alpha = .05$ level when the FDR is controlled; that's every X^2_{2BL} value equal to 20.9 or larger. There is a significant statistic value, with FDR control, for every coalition and for 25 of the 32 states. Clearly the most pervasive and sizable departures from the 2BL distribution occur for the vote counts recorded for the NA party. Table 3 reports for each coalition the number of secciones in each state that have a vote count greater than 9. As the small counts in the NA party column in that table suggest, in most places the NA party received very few votes for president.

*** Tables 2 and 3 about here ***

The small number of votes received by the NA party reflects an election strategy. The leader of that party, who was also its presidential candidate, took various actions to try to shift votes from herself to PAN. As I discuss in Mebane (2007a), there are allegations that some of those actions were illegitimate. The NA leader clearly exhorted her supporters to vote for the PAN candidate, but there are allegations that the exhortations rose to the level of intimidation.

Indeed, there is wider evidence that throughout Mexico votes for president were affected by either strategic voting considerations or undue political influences. They key evidence here is that throughout Mexico voting patterns vary systematically with a measure of which party is strongest in the municipal government in the area where each seccion is located. I develop this argument in Mebane (2007a), and a key finding is the set of 2BL statistics in Table 4. Here I compute 2BL statistics separately for each coalition in each set of party-affiliated municipalities. Many of the statistics are significant even with FDR control across all of the statistics reported in Table 4. Every X^2_{2BL} value equal to 18.9 or larger is significant at the $\alpha = .05$ level by this standard.

*** Table 4 about here ***

Several of the statistics in Table 4 suggest there was some kind of local political influence on the votes recorded for president. For the PAN and PBT vote counts, the mayor's party and significant X^2_{2BL} statistics coincide. X^2_{2BL} is significant for presidential PAN votes only among secciones in municipalities with a PAN mayor, and X^2_{2BL} is significant for presidential PBT votes

[2] I use seccion-level counts and not casilla-level counts because, as I demonstrate in Mebane (2006), the "REDWL" phenomenon means that 2BL-distributed digits should not be expected at the casilla level.

only where there is either a PBT-affiliated mayor or a mayor not affiliated with one of the presidential coalitions. The APM vote counts have significant X^2_{2BL} statistics for every type of mayor except the APM-affiliated mayors. All of the X^2_{2BL} statistics for the NA party vote counts are significant. In Mebane (2007a) I use these and some other results to suggest that PAN and PBT seem to be gaining votes while APM, NA and ASDC are losing votes in places where there are significant departures from 2BL. Votes seem to be going from APM, NA and ASDC to either PAN or PBT. In municipalities where PAN and PBT-affiliated mayors were present, federal candidates affiliated with PAN and PBT, respectively, seem to have tended to gain more votes. The APM coalition seems to have lost votes everywhere except in municipalities with APM-affiliated mayors. Perhaps many voters whose first choice for president was the APM candidate may have noted that candidate's unlikely prospects, observed the closeness of the race between PAN and PBT and decided to abandon the APM candidate in favor of their second choice. Where there were APM-affiliated mayors, perhaps, the voters were somehow deterred from doing this. Or perhaps it was only in those municipalities that APM-favoring voters were able to avoid the massive intimidation they are alleged to have been subjected to.

In any case, the 2BL test statistics here match the highly controversial character of the election, and they suggest interesting directions for further investigation of how the election was conducted. The connections to local political organizations—which are made possible by an extensive data collection effort intended to supplement the presidential vote counts with information about municipality political affiliations (see Mebane 2007a)—are highly suggestive.

Bangladesh 2001: The 2001 General Election in Bangladesh produced an outcome that was close in terms of the vote shares received by the top two parties but not close in the final allocation of legislative seats. A four-party coalition led by the Bangladesh Nationalist Party (BNP) won 40.97 percent of votes and gained 193 seats, while the incumbent leading party, the Bangladesh Awami League (AL), won 40.13 percent of votes and gained 62 seats (Bangladesh Election Commission 2001, 10). Many groups observed the election, and the consensus judgment was in line with the EU-EOM statement quoted above, that the election was sufficiently free and fair. Former U.S. President Jimmy Carter issued a statement that read, in part,

> "It now appears from all international and domestic observer reports that the elections were relatively peaceful and that they were conducted generally in accordance with international standards. These reports run counter to public statements of massive rigging. Any complaints should follow normal legal channels of appeal rather than a complete refusal to recognize the election's legitimacy." (Carter 2001)

Carter's appeal and the observers' judgments did not prevent the AL from rejecting the election result and commencing a two-year boycott of the government. These actions were supported by a study sponsored by the AL that alleged extensive fraud in the election (Centre for Research and Information 2002).

Applications of the 2BL test to polling station vote counts from the Bangladesh elections of 1991 and 1996 do not show extensive departures from the expected 2BL digit pattern, but there are numerous deviations in the vote count data for 2001 (Mebane 2007b). Table 5 summarizes some of the key findings. I compute the 2BL test statistic for the polling station counts separately for each of the 64 *zila* (districts). The official statistical report shows 54 parties participating in the election (Bangladesh Election Commission 2001, 15). I compute the 2BL test statistic separately for the counts for three parties: the AL; the BNP four-party coalition; and the Islami Jatiya Oikya Front (IJOF). I also compute the statistic for the sum of the vote counts of the remaining parties (Others). Because the aggregated Others counts may behave in a systematically different way from the vote counts for a single party coalition, I control the FDR (for $\alpha = .05$) only over the set of statistics for AL, BNP and IJOF. There are 15 significant statistics by this standard: in Table 5 an AL, BNP or IJOF statistic is significant with FDR control if $X^2_{2BL} \geq 24.1$. Table 5 reports the 2BL test statistics for the 12 districts that have a significant statistic for one of those three parties.

*** Table 5 about here ***

All 2BL statistics in Table 5 that are significant with FDR control are for the polling station vote counts recorded for AL or BNP, and there are more significant statistic for the BNP counts than for the AL counts. For each of these parties the most significant departures from the 2BL expected digit pattern occur in Dhaka district. Notwithstanding the sanguine judgments of the groups that observed the election, such results suggest there were extensive irregularities in the election.

Indonesia 2004: Observers described the second round of the 2004 Indonesian presidential election as effective and decisive. Describing the election overall, one observer group wrote, " The Carter Center observed a number of irregularities—many typical of transitional democracies—but, overall, voters were able to exercise their democratic rights in a peaceful atmosphere and without significant hindrance" (Carter Center 2005, 13). An EU-EOM report observed that "the second round Presidential Elections were carried out in an impartial and transparent manner, despite some residual procedural shortcomings" (European Union 2005, 8). The second-round results were not close. Susilo Bambang Yudhoyono defeated the incumbent Megawati Soekarnoputri with 60.7 percent of the votes (Carter Center 2005, 63). No official complaints were filed about these results (European Union 2005, 60).

Applications of the 2BL test to village-level vote counts from the second round show only a few departures from the expected 2BL digit pattern.[3] I compute the 2BL test statistic separately for each of Indonesia's 32 provinces. Table 6 presents the results. Most of the statistics are small. The three Java provinces all have relatively large X^2_{2BL} values. Controlling the FDR for test level $\alpha = .05$ finds four of the statistics from Java provinces to be significant: there is a significant result if $X^2_{2BL} \geq 25.6$.

*** Table 6 about here ***

Discussion

Having only four cases of comparison between the results of the 2BL test and the judgments of diverse sets of election observers and analysts, it is certainly premature to draw confident generalizations about whether the statistical assessment tends to agree with the expert, detailed

[3]Polling station vote counts are available for the second round election, but analysis using polling station counts seems to be affected by "REDWL," similar to the issue with casilla-level counts in data from the Mexico 2006 election Mebane (2006). It may also be that the polling stations have too few voters for 2BL tests against a reference 2BL distribution to be meaningful. Among 537,652 polling stations with nonmissing data, the median number of votes cast is 207. Among the 64,401 villages with nonmissing data, the median number of votes cast in the second round is 1078 (standard deviation 2265).

and nuanced observer reports. Nonetheless the four cases considered here suggest that the 2BL test tends to give results that broadly agree with the observers. In Ohio and Mexico, observers diagnosed many problems with the elections, and the 2BL tests also indicated problems. The 2BL test results are not particularly diagnostic in the Ohio case, in the sense that the tests do not especially flag places known to have experienced serious problems. In the Mexico case, where the 2BL tests were combined with extensive data to measure local partisan configurations, the tests support provocative suggestions about strategic voting and possible voter intimidation. In Bangladesh, the 2BL test results disagree with the sense of observer groups that attempted to confer legitimacy on a contested result, but the test results are very much line with the judgment reached by the losing party, who not only presented a detailed, published case to argue that there was election fraud but also sent a very expensive signal regarding the strength of their beliefs, in the form of a boycott of the postelection government. In Indonesia, observers judged the second round presidential election to be largely although not completely in order, and the 2BL test results concur with that reading.

The results of the preliminary comparative effort seem sufficiently compelling to motivate a more extensive comparison between election monitors reports and statistical tests based on the second digits of low-level vote counts.

References

Associated Press. 2004. "Cleveland Paper Cites Voter Problems." *Associated Press*. December 11.

Bangladesh Election Commission. 2001. "Statistical Report, 8th Jatiya Shangshad Election, October 1, 2001." Election Commission Secretariat, Dhaka, April 2002 edition.

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B* 57 (1): 289–300.

Bjornlund, Eric C. 2004. *Beyond Free and Fair: Monitoring Elections and Building Democracy*. Washington, DC: Woodrow Wilson Center Press.

Carter, Jimmy. 2001. "Postelection Statement by Former U.S. President Jimmy Carter on Bangladesh Elections, Oct. 5, 2001." The Carter Center.

Carter Center. 2005. "The Carter Center 2004 Indonesia Election Report." The Carter Center, June 2005.

Centre for Research and Information. 2002. *A Rigged Election: An Illegitimate Government: Bangladesh Election 2001*. Dhanmondi: Centre for Research and Information.

Estok, Melissa, Neil Nevitte, and Glenn Cowan. 2002. *The Quick Count and Election Observation: An NDI Guide for Civic Organizations and Political Parties*. Washington, DC: National Democratic Institute for International Affairs.

European Union. 2001. "Preliminary Statement—2 October 2001." European Union–Election Observation Mission in Bangladesh 2001.

European Union. 2002. "Handbook for European Union Election Observation Missions." Anders Eriksson, ed. Swedish International Development Cooperation Agency, ISBN 91-586-8777-7.

European Union. 2005. "European Union Election Observation Mission to Indonesia 2004, Final Report." European Union Election Observation Mission, Indonesia 2004.

European Union. 2006. "Final Report: Mexico Presidential and Parliamentary Elections, 2 July 2006." European Union Election Observation Mission, Mexico 2006, 23 November 2006, Mexico City/Brussels.

Hill, Theodore P. 1995. "A Statistical Derivation of the Significant-digit Law." *Statistical Science* 10: 354–363.

Holt, Rush. 2007. "H.R. 811: Voter Confidence and Increased Accessibility Act of 2007 (Introduced in House).".

House Judiciary Committee Democratic Staff. 2005. "Preserving Democracy: What Went Wrong in Ohio." URL http://www.house.gov/judiciary_democrats/ohiostatusrept1505.pdf.

Jacobs, James Q. 2007. “The 2004 Ohio Presidential Election: Cuyahoga County Analysis, How Kerry Votes Were Switched to Bush Votes.” URL http://jqjacobs.net/politics/ohio.html#vocabulary accessed August 21, 2007.

Lehoucq, Fabrice. 2003. “Electoral Fraud: Causes, Types, and Consequences.” *Annual Review of Political Science* 6 (June): 233–256.

Mebane, Walter R., Jr. 2005. “Voting Machine Allocation in Franklin County, Ohio, 2004: Response to U.S. Department of Justice Letter of June 29, 2005.” Unpublished MS.

Mebane, Walter R., Jr. 2006. “Election Forensics: Vote Counts and Benford’s Law.” Paper prepared for the 2006 Summer Meeting of the Political Methodology Society, UC-Davis, July 20–22.

Mebane, Walter R., Jr. 2007a. “Election Forensics: Statistics, Recounts and Fraud.” Paper presented at the 2007 Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 12–16.

Mebane, Walter R., Jr. 2007b. “Evaluating Voting Systems To Improve and Verify Accuracy.” Paper presented at the 2007 Annual Meeting of the American Association for the Advancement of Science, San Francisco, CA, February 16, 2007, and at the Bay Area Methods Meeting, Berkeley, March 2, 2007.

Mebane, Walter R., Jr. 2007c. “Statistics for Digits.” Paper prepared for the 2007 Summer Meeting of the Political Methodology Society, Pennsylvania State University, July 18–21.

Mebane, Walter R., Jr. forthcoming. “Election Forensics: The Second-digit Benford’s Law Test and Recent American Presidential Elections.” In R. Michael Alvarez, Thad E. Hall, and Susan D. Hyde, editors, *The Art and Science of Studying Election Fraud: Detection, Prevention, and Consequences*, .Washington, DC: Brookings Institution.

Mebane, Walter R., Jr. and Michael C. Herron. 2005. “Ohio 2004 Election: Turnout, Residual Votes and Votes in Precincts and Wards.” In Democratic National Committee Voting Rights Institute, editor, *Democracy at Risk: The 2004 Election in Ohio*, .Washington, D.C.: Democratic National Committee.

Norden, Lawrence, Aaron Burstein, Joseph Lorenzo Hall, and Margaret Chen. 2007. “Post-Election Audits: Restoring Trust in Elections.” Brennan Center for Justice, August 1, 2007.

OSCE ODIHR. 2005a. “Election Observation.” OSCE Office for Democratic Institutions and Human Rights (ODIHR), ISBN 83-60190-11-9.

OSCE ODIHR. 2005b. “United States of America 2 November 2004 Elections: OSCE/ODIHR Election Observation Mission Final Report.” OSCE Office for Democratic Institutions and Human Rights (ODIHR), 31 March 2005.

Pericchi, Luis Raúl and David Torres. 2004. “La Ley de Newcomb-Benford y sus aplicaciones al Referendum Revocatorio en Venezuela.” Reporte Técnico no-definitivo 2a. versión: Octubre 01,2004.

Solov, Diane and Diane Suchetka. 2004. “Odd Vote Results Point to Mix-Ups at Some Precincts.” *Cleveland Plain Dealer*. December 10.

Voting Rights Institute, Democratic National Committee. 2005. *Democracy at Risk: The 2004 Election in Ohio*. Washington, D.C.: Democratic National Committee.

Table 1: Ohio 2004 Counties with a Large 2BL Statistic for Either Major Party Presidential Candidate

		Kerry		Bush	
County	n	$n > 9$	X^2_{2BL}	$n > 9$	X^2_{2BL}
Adams	35	35	13.6	35	18.4
Allen	143	143	9.7	143	17.9
Belmont	83	83	17.3	83	5.6
Clermont	191	191	11.4	191	17.5
Cuyahoga	1457	1452	23.7	1396	17.0
Defiance	42	42	18.4	42	14.4
Erie	63	63	7.4	63	17.7
Gallia	35	35	6.4	35	19.4
Greene	142	142	7.3	142	17.5
Hancock	62	62	19.2	62	5.2
Mahoning	313	313	20.3	312	4.7
Morrow	36	36	17.6	36	16.2
Ottawa	78	77	23.5	77	8.4
Paulding	30	30	17.6	30	18.9
Putnam	35	35	18.4	35	5.1
Scioto	106	106	25.2	106	6.0
Shelby	35	35	21.2	35	11.8
Summit	475	475	42.7	474	21.0
Warren	157	157	10.3	157	18.3

Note: n denotes the number of precincts in the referent county, and $n > 9$ denotes the number of precincts with a vote count greater than 9 for the referent candidate. All Ohio counties that have $X^2_{2BL} \geq 16.9$ for the 2004 precinct vote totals reported for either Kerry or Bush are shown.

Table 2: Mexican 2006 Presidential Election: 2BL Test Statistics by State

County	PAN	APM	PBT	NA	ASDC
Aguascalientes	6.6	6.0	14.2	18.9	8.3
Baja California	12.0	13.4	34.9	49.5	20.4
Baja California Sur	6.3	15.6	17.7	13.7	4.0
Campeche	8.3	7.7	19.2	11.8	8.6
Chiapas	8.3	4.7	22.5	23.2	15.8
Chihuahua	9.9	25.1	11.7	131.6	43.6
Coahuila	16.1	5.6	15.3	44.6	12.8
Colima	15.7	10.1	10.6	11.4	6.9
Distrito Federal	10.9	42.6	26.9	178.9	61.6
Durango	5.1	9.8	11.1	51.1	16.3
Guanajuato	13.7	25.4	7.2	117.9	6.6
Guerrero	14.4	28.0	8.8	65.8	11.9
Hidalgo	5.8	19.3	5.4	50.8	4.9
Jalisco	18.1	16.3	24.0	81.6	12.3
Mexico	51.4	15.0	49.8	238.4	18.0
Michoacan	14.0	16.8	5.5	49.7	25.0
Morelos	14.1	8.8	17.3	17.4	11.5
Nayarit	5.6	8.8	7.2	23.7	10.6
Nuevo Leon	8.6	9.3	11.8	7.9	7.3
Oaxaca	16.3	14.4	9.4	38.9	19.8
Puebla	18.5	16.7	20.9	104.3	10.7
Queretaro	16.8	6.3	12.6	25.5	10.1
Quintana Roo	11.6	5.9	6.7	8.8	7.5
San Luis Potosi	9.7	24.9	9.6	43.2	8.1
Sinaloa	39.1	17.1	25.9	36.4	87.9
Sonora	23.1	20.3	3.9	45.5	8.5
Tabasco	7.3	13.6	16.3	6.4	29.0
Tamaulipas	10.7	13.8	9.3	56.8	12.7
Tlaxcala	8.4	11.5	13.5	23.5	12.6
Veracruz	14.0	19.1	9.5	75.2	28.7
Yucatan	20.9	6.6	31.0	46.5	7.5
Zacatecas	10.2	14.5	12.9	35.3	38.5

Note: Entries show X^2_{2BL} statistics for the seccion vote counts for the referent state and candidate.

Table 3: Mexican 2006 Presidential Election: Number of Secciones with Vote Counts Greater than 9, by State

County	PAN	APM	PBT	NA	ASDC
Aguascalientes	593	589	589	187	517
Baja California	1525	1518	1521	460	1253
Baja California Sur	345	339	349	25	199
Campeche	482	488	486	200	186
Chiapas	1787	1922	1919	157	451
Chihuahua	2835	2895	2716	648	1315
Coahuila	1485	1525	1456	226	983
Colima	337	337	338	32	197
Distrito Federal	5557	5524	5559	487	5413
Durango	1330	1354	1155	151	438
Guanajuato	3041	3012	2984	539	1835
Guerrero	1988	2614	2722	289	603
Hidalgo	1600	1695	1689	593	712
Jalisco	3360	3341	3238	1459	2572
Mexico	6171	6149	6187	2567	5308
Michoacan	2585	2636	2678	128	1117
Morelos	911	906	912	423	812
Nayarit	829	875	863	106	252
Nuevo Leon	2134	2134	2068	1048	1545
Oaxaca	2223	2417	2434	64	619
Puebla	2549	2550	2555	687	1511
Queretaro	759	756	759	216	558
Quintana Roo	438	441	442	77	279
San Luis Potosi	1790	1772	1663	213	678
Sinaloa	3720	3743	3528	78	611
Sonora	1360	1345	1315	192	873
Tabasco	902	1134	1136	8	83
Tamaulipas	1734	1734	1714	392	984
Tlaxcala	610	594	610	104	385
Veracruz	4713	4705	4703	332	2077
Yucatan	1078	1077	1038	83	502
Zacatecas	1719	1794	1817	132	496

Note: Entries show the number of secciones in the referent state with a vote count greater than 9 for the referent candidate.

Table 4: Mexican 2006 Presidential Election: 2BL Test Statistics by Municipality Party

	Party	Municipality Party Coalition Membership					
X^2_{2BL}	Voted	PAN	APM	PBT	PAN-PBT	APM-PBT	Other
President	PAN	60.3	7.2	10.2	8.3	10.2	17.4
	APM	44.9	10.5	59.8	22.5	24.8	18.9
	PBT	10.4	3.4	50.5	10.4	12.4	34.7
	NA	387.9	339.8	269.2	76.7	84.7	167.7
	ASDC	4.6	42.9	14.6	33.3	14.7	16.1

	Party	Municipality Party Coalition Membership					
N	Voted	PAN	APM	PBT	PAN-PBT	APM-PBT	Other
President	PAN	17,667	18,341	9,584	1,627	2,459	12,812
	APM	17,620	19,084	10,304	1,663	2,539	12,705
	PBT	17,243	18,570	10,436	1,595	2,412	12,887
	NA	3,740	3,258	2,183	312	582	2,228
	ASDC	10,957	8,173	5,162	680	1,287	9,105

Notes: Top of table: X^2_{2BL} statistics. Bottom of table: N of vote counts ≥ 10. Tests are based on seccion vote counts greater than 9 for the referent party. Each *casilla extraordinaria* used for presidential voting is treated as a separate seccion.

Table 5: Bangladesh 2001 Election: Zila with Significant 2BL Tests for Polling Station Vote Totals, with FDR Control

	AL		BNP		IJOF		Others	
Zila	$n > 9$	X^2_{2BL}	$n > 9$	X^2_{2BL}	$n > 9$	X^2_{2BL}	$n > 9$	X^2_{2BL}
Dinajpur	649	24.7	649	32.3	640	4.1	452	9.9
Naogaon	495	28.4	495	13.2	441	17.0	85	30.7
Natore	409	19.0	409	24.4	306	12.8	200	8.8
Jessore	640	14.8	639	48.9	484	15.2	338	19.2
Mymensingh	1016	21.1	1015	31.3	935	4.0	585	12.5
Dhaka	1666	42.4	1664	74.9	1379	11.3	804	43.8
Narsingdi	446	6.1	445	24.5	356	20.9	145	7.8
Rajbari	242	12.6	241	26.7	218	6.3	79	22.4
Shariatpur	293	24.1	213	5.4	4	5.2	225	6.5
Sunamganj	592	7.0	589	29.4	417	21.2	211	16.7
Noakhali	431	6.0	432	38.6	358	11.3	373	4.9
Chittagong	1400	10.3	1395	40.4	644	30.7	1052	10.1

Notes: 2BL tests, polling station vote counts. $n > 9$ denotes the number of polling stations in each zila with a vote count greater than 9 for the referent candidate. Only zila that have a significant 2BL statistic value (with FDR control) for either AL, BNP or IJOF are shown. FDR control (over all 64 zila) considers only the statistics for the AL, BNP and IJOF vote counts.

Table 6: Indonesia 2004 2BL Statistic for Village-level Vote Counts

		Megawati		Yudhoyono	
Province	n	$n > 9$	X^2_{2BL}	$n > 9$	X^2_{2BL}
Bali	688	688	5.5	687	15.5
Banten	1481	1481	18.0	1480	10.7
Bengkulu	1163	1162	7.4	1163	8.3
DI Yogyakarta	438	438	10.3	438	11.1
DKI Jakarta	157	157	8.1	157	14.7
Gorontalo	447	447	7.1	447	3.8
Irian Jaya Barat	281	218	8.3	239	8.0
Jambi	1197	1188	7.9	1197	5.3
Jawa Barat	5767	5767	34.9	5767	31.1
Jawa Tengah	8565	8562	25.6	8565	21.6
Jawa Timur	8477	8476	19.3	8476	30.5
Kalimantan Barat	1446	1444	9.2	1442	9.2
Kalimantan Selatan	1915	1910	8.2	1914	7.4
Kalimantan Tengah	1115	1112	9.1	1087	10.8
Kalimantan Timur	1053	1047	8.4	987	10.0
Kepulauan Bangka Belitung	315	315	9.9	315	7.0
Kepulauan Riau	245	245	4.8	245	11.0
Lampung	2157	2156	8.2	2157	5.7
Maluku	351	929	17.8	918	9.9
Maluku Utara	589	583	20.3	580	7.9
Nanggroe Aceh Darussalam	5629	5424	5.9	5617	9.5
Nusa Tenggara Barat	782	782	19.1	782	4.7
Nusa Tenggara Timur	1950	1949	5.0	1884	2.1
Papua	557	502	7.0	535	13.6
Riau	1337	1335	2.3	1337	15.3
Sulawesi Selatan	3017	2846	9.7	3015	23.8
Sulawesi Tengah	1340	1286	7.1	1337	8.7
Sulawesi Tenggara	1555	1468	9.5	1555	10.5
Sulawesi Utara	1207	1206	3.4	1207	12.9
Sumatera Barat	2085	2067	7.8	2085	11.4
Sumatera Selatan	2535	2535	2.1	2534	10.3
Sumatera Utara	4560	4491	9.8	4499	4.6

Note: n denotes the number of vialges in the referent province, and $n > 9$ denotes the number of villages with a vote count greater than 9 for the referent candidate.