Identifying Irregularities in Mexican Local Elections

Francisco Cantú University of Houston

A generalized distrust in Mexican local elections raises the question of whether electoral corruption has vanished or remains a prevalent practice in the country. To answer this question, I analyze the 2010 gubernatorial elections, exploiting a feature of the country's electoral system: within each electoral precinct, voters are assigned to polling stations according to their childhood surname. Consequently, the only difference between voters in contiguous polling stations should be their last names. Given that political preferences are seldom correlated with voters' names, I use suspicious differences in turnout levels across contiguous polling stations to identify fraudulent practices. The findings of this article indicate that nondemocratic enclaves that actively obstruct the completion of Mexico's democratic transition still remain today.

The defeat of the Institutional Revolutionary Party (PRI) in the 2000 presidential election marked a watershed moment in Mexican politics. To some extent, the outcome was the result of a series of reforms adopted in the 1990s, which prevented political parties from carrying out electoral manipulation in federal elections. The scope of these reforms, however, was uneven across Mexican subnational governments. As such, popular distrust in the integrity of elections is still quite common at the local level (Hiskey and Bowler 2005; Moreno 2012). The question, then, is to investigate whether post-electoral protests and skepticism about local elections reflect authentic fraud allegations, or whether they are simply a political tool for electoral losers.

Evaluating the integrity of elections is a particularly complicated task, as neither "winners" nor "losers" have incentives to reveal the truth: perpetrators of fraud usually want to keep their activities as hidden as possible (Lehoucq 2003). The alleged victims of fraud, in turn, may be willing to claim the existence of electoral manipulation even when no irregularities are recorded (Eisenstadt and Poiré 2006; Magaloni 2010). To overcome this challenge, I propose a novel empirical strategy to detect electoral fraud. I focus on the case of Mexico and exploit a feature of the country's electoral code: within each electoral precinct, voters are assigned to polling stations according to their childhood surnames; consequently, the only difference between voters at contiguous polling stations should be their last names. Because political preferences are seldom correlated with voters' last names, I identify fraudulent practices by finding unexpected differences in turnout levels and partisan votes across contiguous polling stations. I illustrate this methodology by evaluating the 2010 gubernatorial elections held in 12 states and provide evidence that candidates in three of the states benefited from electoral irregularities. In fact, using this methodology, some of the polling stations that are flagged as suspicious were notorious for violent disruptions or duplicated ballots.

This article contributes to the research on electoral fraud in two ways. First, the findings indicate that nondemocratic enclaves that actively obstruct the completion of Mexico's democratic transition still remain. Second, from a methodological standpoint, the proposed approach contributes to the growing literature that uses statistical tools to evaluate the quality of elections (Beber and Scacco 2012; Fukumoro and Horiuchi 2011; Levin et al. 2009; Mebane 2006; Myagkov, Ordeshook, and Shakin 2009). Specifically, my method proposes an alternative way to identify outlier observations as evidence for fraud (Alvarez and Katz 2008; Hausmann and Rigobón 2011; Jiménez 2011; Wand et al. 2001).

This article is organized as follows. The next section reviews the challenges of measuring the democratization process in Mexico at the subnational level. The third section describes the foundations of the empirical model and explains the methodology I propose for measuring electoral fraud. The subsequent section presents the results

Francisco Cantú is Assistant Professor, Department of Political Science, University of Houston, 429 Philip Guthrie Hoffman Hall, Houston, TX 77204-3011(fcantu10@uh.edu).

I am grateful to R. Michael Alvarez, Allyson Benton, Scott Desposato, Federico Estévez, Stephan Haggard, Eric Magar, Sebastián Saiegh, Peter Smith, and three anonymous reviewers for their suggestions. Participants at seminars at UCLA, USC, ITAM, and CIDE also provided valuable feedback. I thank Mario Cortina, Pablo Mateos, and Heriberto Romo for sharing data on last names in Mexico. All errors are my own. Replication materials may be found at the AJPS Data Archive on Dataverse (http://hdl.handle.net/1902.1/21752).

American Journal of Political Science, Vol. 58, No. 4, October 2014, Pp. 936–951

 DOI: 10.1111/ajps.12097

and their cross-validation across other tests. Finally, the last section discusses the implications of this research.

Local Elections in Mexico

During most of the twentieth century, elections in Mexico failed to function as a legitimate process for selecting public officials. Electoral manipulation included the modification of final vote counts (Castañeda 2000, 231–39; de la Madrid 2004, 814–24), the alteration of ballot boxes (Langston 2012, 19–22), the certification of the results by a biased legislature (Lehoucq 2002), the inflation of voter registration lists (Gillingham 2012; Molinar 1987; Preston and Dillon 2004; Simpser 2012), and the intimidation of opposition supporters (Craig and Cornelius 1995)[1]. Despite instances of centrally planned electoral manipulation (Carbonell 2003, 83–6; Simpser 2013, 113), many of these practices took place in a decentralized manner: local brokers manipulated the ballot boxes in their own regions in exchange for political favors from winning candidates (Langston 2012).[2]

Following the electoral reforms of the 1990s in Mexico, political parties became unable to carry out electoral manipulation at the federal level (Magaloni 2006, 36–38).[3] In the case of local elections, however, the scope of reforms was uneven across the states. Although the Constitution provides a minimum set of rules that each local election must follow, the states are freely able to interpret their right to organize the elections for governors, local legislators, and municipal councils, leaving the quality of these elections dependent on the extra-institutional dynamics of each province (Peschard 2010). Consequently, anecdotes of corruption in local elections are not uncommon, perpetuating a distrust in the fairness of local elections and the impartiality of state electoral institutions (Moreno 2012).

Among the fraudulent techniques described above, a strategy known as *fraude hormiga* may be prevalent.[4] This form of electoral manipulation refers to "the illegal introduction or subtraction of a very few votes in order not to affect the outcome in the polling station—and avoiding its potential nullification—but enough to affect the final outcome in the aggregate" (Crespo 2006, 128–29). In other words, political machines, when they have the opportunity, change the vote counts of the polling stations by a number unlikely to be noticed but considerable enough to be decisive in the aggregate count. The incentives for this type of manipulation come from Mexico's electoral code, which dictates the nullification of vote tallies from polling stations where the number of irregular votes, either by accident or fraud, is greater than the difference between the two leading candidates. However, these grounds for nullification are not applied to the overall result, even in cases where the aggregate number of irregular votes is greater than the difference between the two main competitors. This limitation in the Mexican electoral legislation creates an opportunity for *fraude hormiga*, which becomes more attractive in close elections (Crespo 2008).

The ability to buy off poll workers facilitates this type of fraud. Political parties typically offer money or threaten violence, inducing a poll worker's absence on Election Day (*New York Times* 2010; Raphael 2007). If poll workers do not show on Election Day, political parties install partisan agents who not only allow irregularities to be tolerated at the polling stations but also facilitate altering the vote count in favor of a particular candidate.[5]

2010 Gubernatorial Elections

When the PRI lost the presidency in 2000, many of its members sought consolation in their remaining bastions of political power: the subnational governments (Dresser 2003). Ten years later, the PRI had positioned itself to control not only those regions still controlled by the old guard *priístas* but also states run by other political parties. On July 4, 2010, voters in 12 out of the 32 states in Mexico elected a new governor (see Figure 1). These elections represented a decisive phase in defining the political strength of each party in anticipation of the 2012 presidential election. In most of these states, the election pitted a PRI candidate against an unusual alliance of opposition parties. The coalition, principally formed by the conservative National Action Party (PAN), and the leftist Democratic Revolutionary Party (PRD), was in response to a fear of an overwhelming outcome in favor of the PRI. This strategy was controversial, even among the members of the various coalition parties, but it represented a

[1] Although vote buying is a common issue in many instances (Cornelius 2002; Gibson 2005), this project does not include this type of electoral manipulation.

[2] Larreguy (2012) provides evidence of how local party brokers mobilize clientelistic networks for local elections in Mexico.

[3] For a detailed description of each electoral reform, see Ochoa-Reza (2004) and Craig and Cornelius (1995).

[4] The closest translation of this term in English would be "ant fraud."

[5] Author's interviews with PRI and PRD party brokers, Oaxaca, July 2010.

FIGURE 1 States with Governor Elections on July 2, 2010

pragmatic response to the common practices of local political machines.[6]

The overall results of the election were mixed. While the coalition won in the states of Oaxaca, Sinaloa, and Puebla, the PRI retook Aguascalientes, Tlaxcala, and Zacatecas, the first two of which were formerly governed by the PAN, and the last had had an outgoing governor of the PRD. Although both the PAN and the PRD celebrated seizing three of the PRI's historical bastions, they contested the elections held in Durango, Hidalgo, and Veracruz, claiming fraud. The Federal Electoral Court did not find substantive evidence to support their claims that the elections had been manipulated, and thus certified the legitimacy of the outcomes.[7] Still, a post-electoral poll showed that the percentage of respondents who lacked confidence in the electoral process in their respective states ranged from 25% to more than 50% (Gabinete de Comunicación Estratégica 2010).

Empirical Analysis

Previous studies in Mexico have focused on detecting electoral irregularities (Crespo 2006; Instituto Federal Electoral 2010; Mebane 2006) and determining whether certain inconsistencies affect electoral outcome (Aparicio 2006, 2009; Poiré and Estrada 2006). This article's proposed methodology has two advantages over previous analyses. First, the method I use distinguishes between random and systematic effects at the polling station level. This feature reduces the possibility of confusing cases of deliberate manipulation with cases that should be classified as accidental errors. Second, unlike Aparicio (2009) and Poiré and Estrada (2006), my methodology does not categorize the observations according to the winning candidate in the district; rather, it takes into account the high variance in electoral behavior at the precinct level (Navia 2000).

Identification Strategy

Precincts are the smallest electoral subunits. They group voters into units of 50 to 1,500 (*Código Federal de Instituciones y Procedimientos Electorales* [COFIPE] 2012, Art. 155). Within each precinct, there must be one polling station for every 750 voters. Due to demographic changes after a precinct has been drawn, however, the number of voters in a particular precinct could be greater than 1,500; consequently, some precincts may have more than two polling stations. The first polling station is called the *casilla básica*,[8] and the subsequent polling stations are called *casilla contigua* (e.g., *casilla contigua 1*, *casilla contigua 2*). There is a *casilla contigua* for each additional group of 750 voters in a precinct.[9]

[6]The PRI's leader in the Senate, Manlio Fabio Beltrones, commented that "alliances between enemies that don't respect each other are contrary to nature," and the Interior Minister of the Federal Administration, PAN member Fernando Gómez Mont, said that that kind of coalition "could end up leading to fraud" (*Los Angeles Times* 2010).

[7]See Prats (2010).

[8]The translation in English is "basic polling station."

[9]The translation in English is "contiguous polling station."

The assignment of precinct voters to a particular polling station is strictly alphabetical; that is, voters are distributed among the polling stations according to their last names (COFIPE 2012, Art. 152). If possible, all of a precinct's polling stations must be located in the same building; otherwise, polling stations must be in adjacent locations to provide similar transit access for all voters (COFIPE 2012, Art. 239).

As an illustration, we can focus on the electoral geography of the state of Aguascalientes, which is shown in Figure 2. The lines define the boundaries of the 584 precincts in that state, which are delineated by geographical and demographic characteristics. There are a total of 1,313 polling stations throughout the state. Consider the allocation of voters within precinct 129, which, by the time of the local election in 2010, had 1,416 registered voters assigned by alphabetical order to two different polling stations, each allowing 708 voters on its voting list. Suppose that there is a household in the precinct with two registered voters, and that the last names of these voters are Abasolo and Zurita. Regardless of the fact that they share the same address, voter Abasolo would be assigned to a different polling station than the one assigned to voter Zurita. In short, the only condition for the allocation of voters to precincts is their home address, and within each precinct, the assignment of voters to a particular polling station is dependent upon their surname.[10]

Therefore, for this research, I identify polling stations with potential irregularities by comparing the turnout rates in each unit with those at other polling stations within the same precinct. After classifying each polling station, given its turnout rate relative to other polling stations in the precinct, I measure the change in vote returns for each candidate.

Last Names and Political Behavior. If last names are not correlated with voting behavior, then each polling station should be an unbiased sample of the precinct's voters. Therefore, it is important to ensure that sorting voters by their last names is orthogonal to their voting behavior. Due to the Mexican legislation that protects public records of voter turnout and the identification of a respondent's last name in political surveys, I use four indirect approaches to ensure that there are no differences in sociodemographic variables when grouping individuals by their surnames. The detailed results of this analysis and the databases for this subsection are available in the supporting information of this article.

The first approach uses a national monthly survey of TV viewership conducted by the Social Research Institute. The poll was conducted in July 2011 with a sample size of 1,499 respondents. Given that this poll contains the names of the interviewed individuals, I could identify the first letter of their last names to determine whether there are significant differences in age, gender, and frequency of watching television news, which is the main source of political information in the country (Moreno 2012). I did not find any significant differences in any of these variables when the respondents were sorted by their last names and grouped into halves, thirds, and quartiles.

The second approach analyzes the last names of children who receive benefits from the cash-conditioned transfer program called *Oportunidades*.[11] This program seeks to improve health and education among poor families in Mexico and currently serves more than five million households across the country. The only condition for a family to be targeted by *Oportunidades* is that its income must fall below the poverty line established by the program (Levy 2006). To check for differences of last names across regions, I compare the distribution of the first letter of the last names in every state. The results show only two outlier observations: first, a high proportion of children with the last name Hernández in Hidalgo, and second, the density of last names with Mayan origin and beginning with the letter *C* in Quintana Roo.

To determine whether the population under the poverty line is clustered in just a few last names, the supporting information shows the results at state and municipality levels using Yule's K and Simpson's D, which are two different diversity measurements that show the probability of randomly selecting two individuals with the same last name.[12] The results, state by state, show that

[10]There are two types of polling stations that have additional requirements other than those labeled as either *básica* or *contigua*. There are "special polling stations," or *casillas especiales*, that receive the ballot papers of voters who are temporarily situated outside of their assigned precinct (COFIPE 2012, Art. 270). Similarly, there are "extraordinary polling stations," or *casillas extraordinarias*, which are designed for those precincts whose sociocultural or geographic conditions make it difficult for all voters to commute to the same place (COFIPE 2012, Art. 239). These polling stations represent less than 1% of the observations and are not taken into account in my analysis.

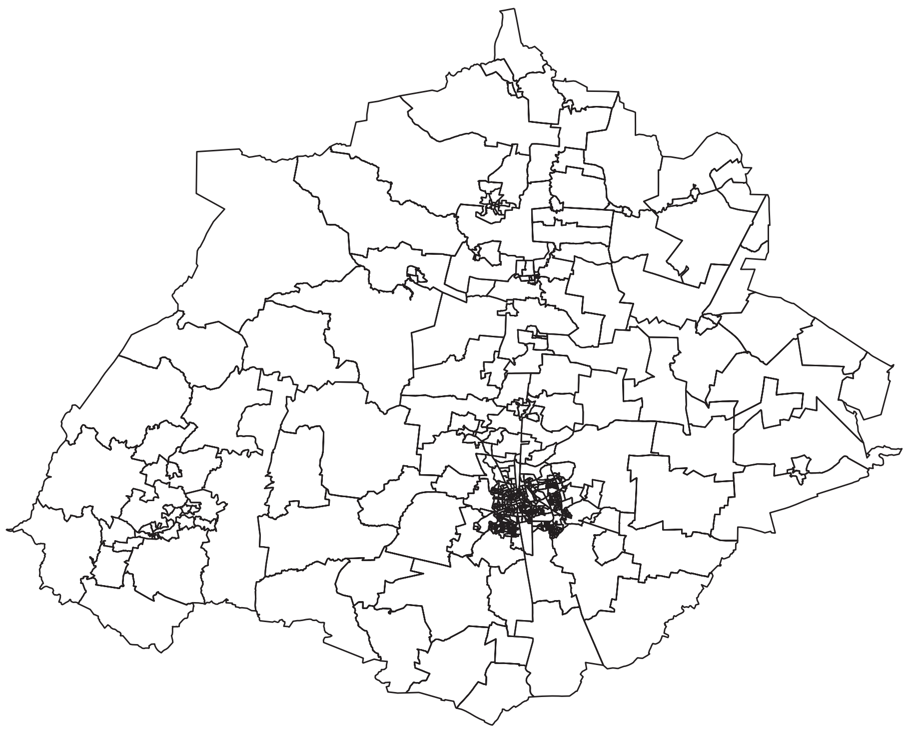
[11]The information is available online at http://www.oportunidades.gob.mx/.

[12]Yule's K is given by

$$K = 10^4 \frac{1}{N^2} \left(\sum_{r=1}^{v} r^2 V_r - N \right), \tag{1}$$

where N is the size of the sample, v is the highest occurring frequency of the last name, and $V_r = r = 1, 2, \ldots, v$ is the number of different surnames with frequency r. The lowest values of K mean that almost every individual in the population has a unique last name, whereas $K = 10,000$ means that all the individuals share the same last name (McElduff et al. 2008, 189). Simpson's D is

FIGURE 2 Electoral Precincts in Aguascalientes



the highest probability of randomly choosing two individuals with the same last name is given by Yule's K in Hidalgo, with a probability of 0.03.

Third, I use the frequency and order of the most common last names on the *Oportunidades*' list to compare them with the most common surnames registered in the electoral roll in each state. As the information of the electoral roll is truncated to show the 100 most common names in every state, I rank the last names found on both lists and compare these rankings using a Spearman correlation analysis. The correlation coefficient for each state is greater than 0.75 in 11 out of the 12 states.

Finally, the online appendix shows one additional test that I utilize. Given the classification of the polling stations that I explain in the following subsection, I explore the probability that observations classified by the methodology as "suspicious" depend on the way voters in a precinct are assigned to polling stations. Given that it is possible to sort polling stations by the alphabetical order in which voters were assigned, I can compare whether, for example, the polling stations with the first letters of the alphabet are more prone to be classified as suspicious than those polling station with other letters in the alphabet.

Methodology

Figure 3 graphically represents the procedure for the analysis of each of the 12 elections held on July 2, 2010. The unit of analysis is the polling station, and the variables are the total number of valid votes (i.e., the number of votes at the polling station that are neither spoiled nor cast for a nonregistered candidate), the number of registered voters, and the number of votes for each candidate.

Classification. If electoral behavior is unrelated to the assignment of voters to particular polling stations within the precincts, I do not expect to find significant differences in the turnout levels among the polling stations in the same precinct. Therefore, I consider the general dispersion of the observations in each precinct as follows.[13]

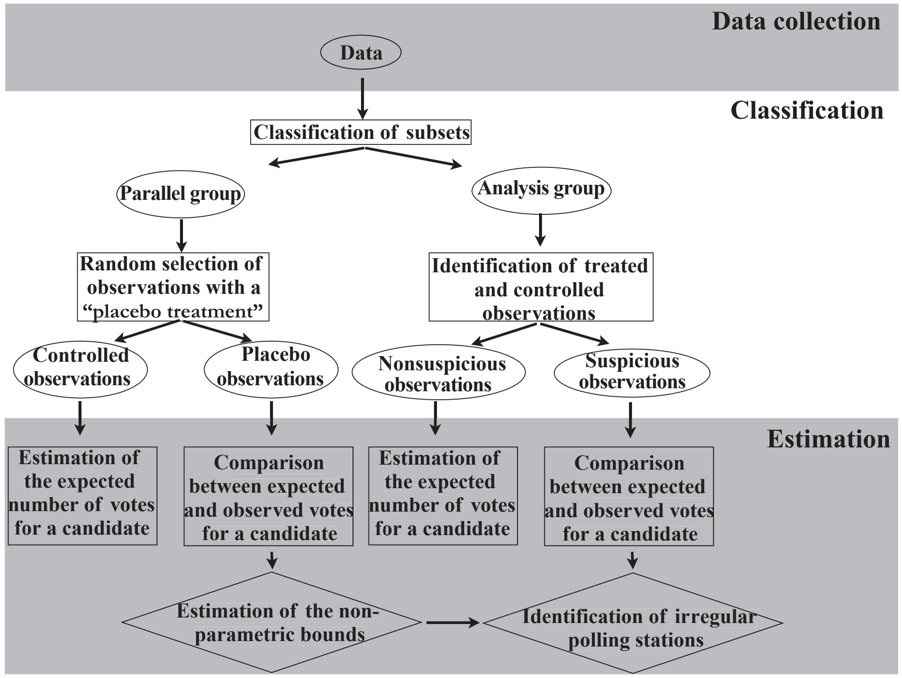
given by

$$D = \frac{1}{\sum_{i=1}^{v} p_i^2} \qquad (2)$$

or the proportion of last names i relative to the total number of last names, p_i (Hunter and Gaston).

[13]This study only considers observations in the database that are free of inconsistencies identified by the electoral authorities. That is, I remove those observations that were marked as ambiguous by the election officials due to the illegibility of the numbers on the official form, the presence of obvious errors made when filling out the form, or the submission of the form without the official

FIGURE 3 Diagram of the Procedure for Detecting Irregular Observations



Consider an election with M different precincts. In every precinct m, I measure the differences in turnout between a polling station i and the rest of the polling stations in the precinct. The turnout estimation, T_i, divides the number of valid votes by the number of registered voters at polling station i. The differences in turnout rates between polling stations are estimated as follows

$$d_m = |T_i - T_j|; \quad i,\ j \in m, i \neq j$$

That is, the intraprecinct turnout difference is calculated as the absolute difference of turnout levels between two polling stations in the same precinct. As an illustration of this methodology, Table 1 shows the vote returns for four electoral precincts in Aguascalientes, where precinct 129 is one of the four. To calculate the differences in voter turnout in precinct 129, I obtain the absolute value of the difference of turnout rates observed between polling stations 129-A and 129-B, which is $d_{129} = |0.64 - 0.55| = 0.09$.

After each precinct is classified according to the turnout differences of its polling stations, the data are divided into two sets of observations. The first subset contains all polling stations without significant differences among turnout levels between the precincts; this subset is labeled the *parallel group*. The second subset

TABLE 1 Vote Returns for Four Different Precincts in Aguascalientes, 2010

Precinct-Polling Station	PRI	PAN	Total Votes (Turnout)
129-A	154	87	278
	(55%)	(31%)	(55%)
129-B	142	138	323
	(42%)	(44%)	(64%)
377-A	159	76	331
	(48%)	(23%)	(45%)
377-B	156	63	331
	(47%)	(19%)	(45%)
434-A	128	138	294
	(43%)	(46%)	(70%)
434-B	105	109	243
	(41%)	(43%)	(60%)
546-A	113	105	227
	(49%)	(46%)	(57%)
546-B	109	98	227
	(43%)	(48%)	(57%)

envelope. These observations represent less than 0.5% of the total observations. I also exclude both extraordinary and special polling stations, which allow voters from outside the precinct to vote or are designed for voters living in communities at a far distance from the regular polling stations but in an insufficient number to constitute their own precinct.

contains those precincts with values of d_m greater than or equal to the quantile 0.95 of the distribution of intraprecinct differences in the state. This subset is labeled the *analysis group.*

In the case of Aguascalientes, the value of the 0.95 quantile is 0.078. Therefore, the analysis group in the example contains the precincts 129 and 434, in which intraprecinct differences in turnout are $d_{129} = 0.09$ and $d_{434} = 0.10$. In contrast, precincts 377 and 546 are assigned to the parallel group.

After dividing the data into two groups, it is necessary to classify the polling stations within each precinct, which I label as either a *suspicious* or a *nonsuspicious* observation. Because the quantity to be estimated is the additional number of votes for a particular party, a suspicious observation must be in the analysis group and has a turnout level above the mean turnout in that precinct. The rest of the polling stations in the analysis group are labeled as nonsuspicious observations.

In the example of Table 1, polling stations 129-B and 434-A are classified as suspicious observations because their turnout rates are higher than the mean turnout in their precinct. In contrast, polling stations 129-A and 434-B are labeled as nonsuspicious observations. To classify the polling stations in the parallel group, I randomly select a group of polling stations in each precinct and assign them a "placebo treatment." In this step, I treat each observation in the subset as a Bernoulli trial with two possible outcomes and a probability p=.5 of being either a suspicious or nonsuspicious observation in the precinct.

Estimation. After classifying precincts and polling stations given their turnout levels, I determine whether the turnout differences also affect the distribution of votes among the candidates. If the turnout differences are caused by unintentional errors during the vote count, the consequences of these errors should affect the political parties in an unbiased way. That is, regardless of the differences in the total number of votes between the polling stations, the proportion of votes for any given party should be the same at different polling stations in the same precinct.

If this condition holds true, consider a precinct m with a suspicious polling station i and a nonsuspicious polling station j. If voters can choose between alternatives x and y, $v(x)$ and $v(y)$ represent the votes for each alternative, making $V = v(x) + v(y)$ the total votes cast in the polling station. I then use this information to estimate the expected number of votes for each alternative in a polling station. For example, for the case of the expected number of votes for alternative x in polling station i, the estimation is[14]

$$e(v_i(x)) = \frac{v_j(x)}{V_j} * V_i; \quad i, j \in A, i \neq j$$

Recall the example described in Table 1. The expected number of votes for the PAN in polling station 129-B is $e(v_{129-B}(PAN)) = \frac{87}{278} * 323 = 101.08$, whereas the expected vote tally in polling station 434-A is $e(v_{434-A}(PAN)) = \frac{109}{243} * 294 = 131.9$. Comparing the number of expected votes for the PAN with the observed number of votes in these polling stations, a difference of more than 36 votes in polling station 129-B and fewer than seven votes in 434-A may be observed.

To evaluate the random error for the expected number of votes, I consider the dispersion of votes for a party within the parallel group, where observations lack significant differences in the total number of votes within the precincts, and I assume no irregularities can be observed. In this group, I also estimate the expected number of votes for the randomly selected suspicious observations. If differences in turnout do not systematically affect an electoral alternative, the differences between expected and observed votes should be similar between the parallel and analysis groups.

Figure 4 shows the results for the PAN in Aguascalientes. The left plot shows the difference between the numbers of expected and observed votes in the parallel group. This plot displays only the observations that were randomly assigned as suspicious, and the expected number of votes is estimated using the unselected polling stations. For example, polling stations 377-A and 546-B are randomly considered "suspicious" in the parallel group, and each station's expected number of votes for each party is estimated by using the vote returns in polling stations 377-B and 546-A.

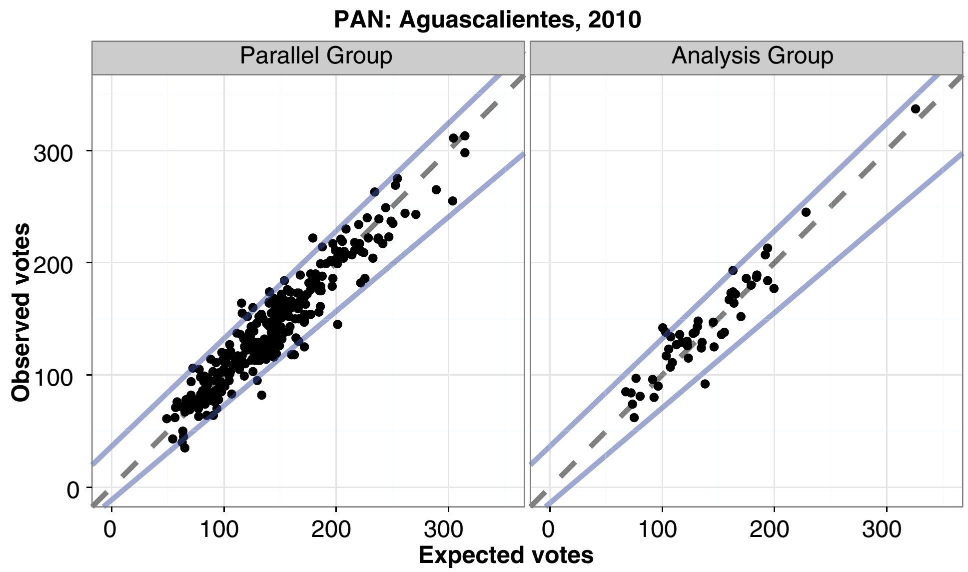
A perfect correlation between the number of expected and observed votes would imply that the proportion of votes for any given candidate is the same across all polling stations in the same precinct, so all observations would fall over the dashed line. Because the measurement is subject to random events that may affect the number of

[14]In the case that the number of nonsuspicious polling stations J in the precinct is greater than one, the expected number of votes for the treated section is

$$e(v_i(x)) = \frac{\sum_{j=1}^{J} v_j(x)}{\sum_{j=1}^{J} V_j} * V_i; \quad i, j \in A, i \neq j.$$

In other words, the estimation of the expected number of votes averages the controlled observations within the precinct.

FIGURE 4 Expected and Observed Votes in Aguascalientes, 2010



votes at any polling station, I estimate the continuous quantiles that include 95% of the observations of the parallel group. In Figure 4, the solid lines in the plot show the Bayesian quantile regression for quantiles 0.025 and 0.975. In other words, these two lines represent the bounds where the comparison between the expected and observed votes should be located 95 out of 100 times.[15]

The final step is to use the bounds from the parallel group to the analysis of the treated polling stations in the analysis group. If turnout differences can be explained by unintentional factors, then the proportion of votes that a party receives should correspond to the polling stations of the precinct (i.e., the observed number of votes will lie within the 95% confidence interval of the expected vote estimation). Otherwise, it is plausible that the irregularities of the turnout levels observed at the polling station disproportionally affect a particular political party.

The right panel of Figure 4 shows the difference between the expected and observed numbers of votes in the analysis group. I use the same bounds from the left plot to consider the variation when no differences in the total number of valid votes were observed. The suspicious observations described in the example are the red crosses on the right panel. The observed number of votes for the PAN in polling station 434-A falls within the solid green lines; that is, despite the difference in the turnout rates between the polling stations of precinct 434, the numbers of votes for the candidates remain similar. In contrast, polling station 125-B displays not only a greater number of valid votes compared with other observations in the precinct but also a higher proportion of votes for the PAN. This difference falls outside the bounds and thus would be considered an electoral irregularity.

Results

The graphical analyses for the incumbent party and the most significant challenger in each of the 12 elections examined in this article are located in the online appendix. In this section, I discuss three of the cases represented in Figures 5 through 7. First, consider Figure 5, which displays the PRI's electoral returns in the state of Hidalgo. In this case, the number of observations outside the upper bound is very similar for both the analysis and parallel groups. Consequently, although the election results were contested by the opposition, the evidence produced by this methodology does not support the claim of fraud.

Second, in the case of Oaxaca, illustrated in Figure 6, the proportion of observations above the upper bound is clearly larger in the analysis group than in the parallel group, which suggests that most of the irregularities benefited the PRI's candidate. Some of the irregularities that the algorithm detects are in the municipality of Tututepec, where citizens filmed a meeting between people

[15]A quantile regression follows a model given by $y_i = x_i^{'}\beta_\tau + \epsilon_i$, where ϵ is the error term restricted to have the τ-th quantile equal to zero. Thus, the objective is to estimate β_τ by minimizing $\sum_{i=1}^{n} \rho_\tau(y_i - x_i^{'}\beta_\tau)$, where ρ is the tilted absolute value function that yields at τ-th sample quantile as its solution (Koenker and Hallock 2001; Kozumi and Kobayashi 2009).

FIGURE 5 Estimation of Expected Votes for the Incumbent Coalition in Hidalgo, 2010

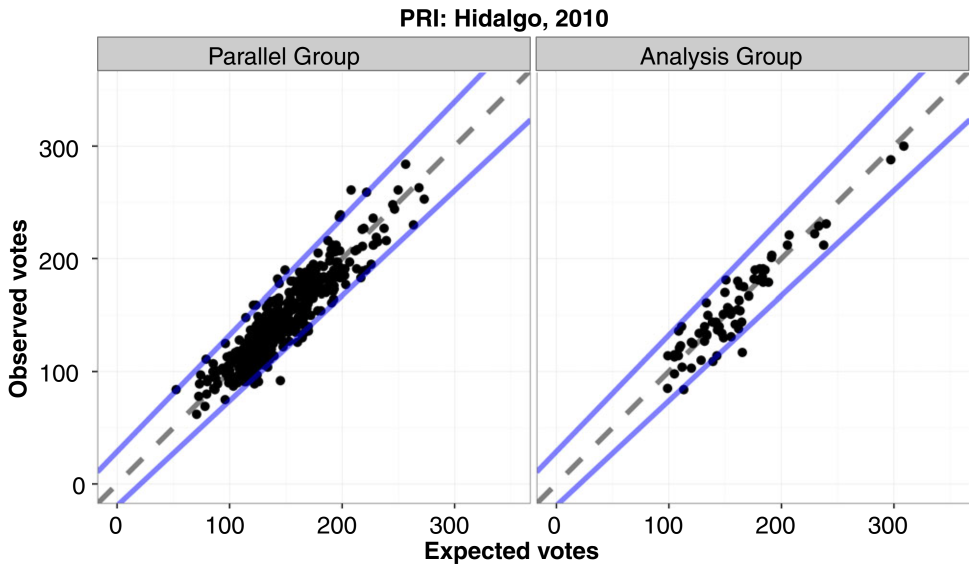
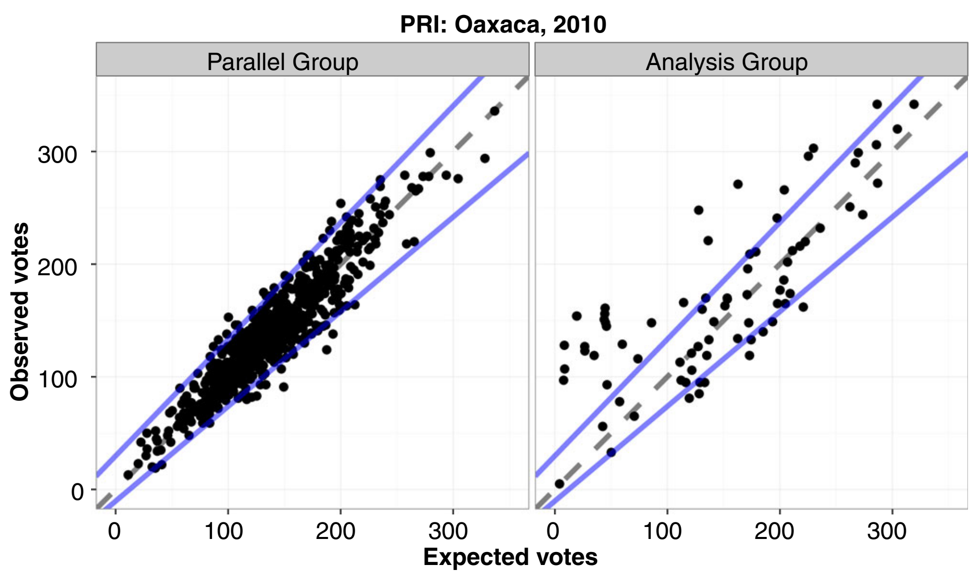


FIGURE 6 Estimation of expected votes for the incumbent coalition in Oaxaca, 2010

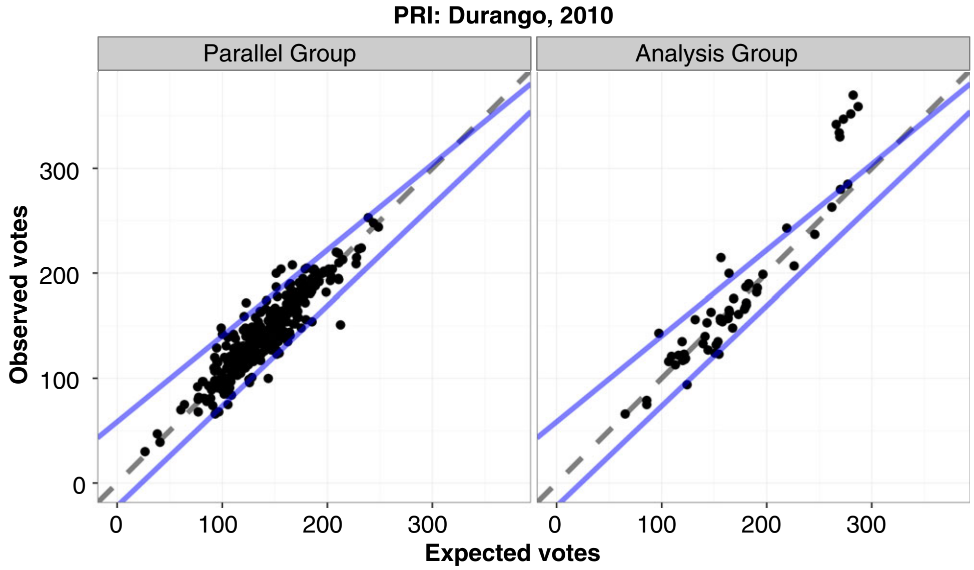


from the local electoral institution and PRI supporters, in which the latter group received paper ballots and other electoral supplies.[16]

Finally, the graphical analysis for Durango, illustrated in Figure 7, shows that most of the electoral irregularities have unusual voting returns for the PRI; for example, there are only two observations in the parallel group that depict vote returns above 250 for the PRI. The results identify specific events that occurred on Election Day. Most of the detected irregular observations in Durango occurred in the city of Gómez Palacio, where an armed group disrupted the electoral process at the polling stations in precinct 447. This one event resulted in two dead police officers and caused voters and poll workers to flee the polling stations. Despite the evidence and the allegations from the challenging coalition, the local electoral court ruled against nullifying the votes in the precinct; electoral officials counted all votes from all of the polling

[16]See http://www.youtube.com/watch?v=Mt3Azm0OtuI.

FIGURE 7 Estimation of Expected Votes for the Incumbent Coalition in Durango, 2010



stations located in this precinct (*El Siglo de Durango*, 2010 6; *La Jornada* 2010, 31).

When the irregularities lack anecdotal evidence to assess their causes, I propose an additional step to evaluate whether the irregularities in an election are the product of innocent accidents or have a systematic bias for a particular candidate. A bias for a particular candidate via irregular turnout in the polling station can be inferred if (1) the evidence suggests that the proportion of observations outside the bounds has a positive bias to the vote returns for a candidate and (2) most of these observations are in the analysis group. To accomplish this task, I use a variation of the add-1-dummy regression outlier model proposed by Polasek (2003). In the original setting, the location of the dummy varies over all observations, creating n different model estimates and using marginal likelihoods as a model choice criterion. In this case, the outlier observations are already identified, and I evaluate the significance of the outliers in modifying the vote returns for a particular candidate. So, rather than creating n different model estimates, I only compare a model with a dummy that identifies the observations in the analysis group and a model that does not distinguish whether each observation is in the parallel or the analysis group.

If the evidence suggests that the proportion of observations outside the bounds has a positive bias to the vote returns of a candidate and that most of these observations are in the analysis group, then a bias for a particular candidate via irregular turnout in the polling station can be inferred. In short, this procedure assesses whether the evidence supports the following two candidate models for the observed data (y_i):

$$M_0 : y_i = \alpha + \beta x_i + \epsilon_i$$

$$M_1 : y_i = (\alpha + \gamma D_i) + \beta x_i + \epsilon_i; \quad D_i = \begin{cases} 1 \text{ if } i \text{ in Analysis Group} \\ 0 \text{ if } i \text{ in Parallel Group} \end{cases}$$

In this case, the observed number of votes (y_i) is either the product of a linear and proportional relationship with the expected number of votes (M_0) or the product of a linear relationship with the expected number of votes plus a positive intercept for those observations in the analysis group (M_1).[17]

[17] For each candidate, a Markov chain Monte Carlo (MCMC) simulation of 100,000 iterations was obtained after removing a burn-in of 10,000. The log Bayes factor estimates were validated with the Chib (1995) method using the MCMC-pack (Martin, Quinn, and Park 2011). The prior distributions for the parameters are

$$\epsilon_i \sim \mathbb{N}(0, \sigma^2)$$

$$\beta \sim \mathbb{N}(1, 1)$$

$$\alpha \sim \mathbb{N}(0, 5)$$

$$\gamma \sim \mathbb{N}(10, \frac{1}{2})$$

$$\sigma^{-2} \sim \mathbb{G}(25, 1)$$

For the sensitivity analysis, the supporting information shows different estimations for the model with different values for σ^{-2}

The summary of the posterior distributions of the parameters for each model and election displayed in the online appendix. I use Bayes, theorem to calculate the posterior probability for each candidate model as follows:

$$Pr(M_k|y) = \frac{p(y|M_k)\,Pr(M_k)}{p(y|M_0)\,Pr(M_0) + p(y|M_1)\,Pr(M_1)};\ k = 0, 1$$

I assign the same prior probability to each of the candidate models to be true (i.e., $Pr(M_k) = 0.5$). With the above quantities, the posterior odds in favor of M_1 are estimated:

$$\frac{Pr(M_1|y)}{Pr(M_0|y)} = \frac{P(y|M_1)}{P(y|M_0)} * \frac{Pr(M_1)}{Pr(M_0)}.$$

The above implies that the posterior odds are a product of the prior odds using the Bayes factor, which is defined as the ratio of the marginal likelihoods:

$$B_{10} = \frac{P(y|M_1)}{P(y|M_0)} = \frac{\int p(y, \theta_1)d\theta_1}{\int p(y, \theta_0)d\theta_0},$$

where θ_0 and θ_1 are the parameters for M_0 and M_1. I use the Bayes factor as a summary of the evidence provided by the data as opposed to the prior supposition. In particular, I assess whether the number of votes for a candidate in those polling stations with suspicious turnout is systematically greater than expected. Following Jeffreys (1961), I use the Bayes factor as evidence against M_0, which assumes that the errors of the linear estimation between the expected and observed number of votes are unbiased disturbances for each candidate.[18]

Figure 8 summarizes the value of the Bayes factor that was derived from the division of marginal likelihoods for each model, $P(y|M_k)$. To illustrate how to calculate the estimation, consider once again the vote returns for the incumbent coalition in Aguascalientes. The log marginal likelihood for M_0 is -1747.992, whereas the log marginal likelihood for M_1 is -1761.776. Consequently, the Bayes factor is estimated as $B_{10} = e^{(-1761.776)-(-1747.992)} \sim 0.00001$. In this case, the value of B_{10} suggests that the evidence against M_0 is negative, which supports a model that does not distinguish for differences in turnout across the polling stations of the precincts.

The Bayes factor finds positive evidence for the alternative hypothesis, M_1, in Durango and strong evidence for M_1 in Oaxaca and Veracruz. That is, the data provide at least positive evidence of a systematic bias in the irregular observations to the benefit of the incumbent candidate, who was a PRI politician in these three cases. There is no reason to assume that random error would benefit the incumbent party, but the distribution of political capital among the candidates makes it plausible that non-accidental irregularities were triggered by the

and shows that the results hold under different specifications.

I also include a similar analysis based on Jiménez (2011) in his analysis of the 2004 referendum in Venezuela. In this case, I consider as the quantity of interest the difference D of comparing the expected $(e(v))$ and observed votes (v) for an alternative x in every polling station i. In the context of this analysis, the alternative x_i is the incumbent party or coalition in every state. If K is the total number of polling stations in both the parallel and analysis groups, then $D = \frac{\sum_{i=1}^{K}(v_{x_i} - e(v_{x_i}))}{K}$.

Out of this population, there is a sample S of size k, and the quantity of interest is the average difference between the expected and observed votes for the alternative x, denoted by $d_k = \frac{\sum_{i\in S}^{k}(v_{X_{\in S}} - e(v_{X_{\in S}}))}{k}$. Jiménez (2011) proposes to vary the size and composition of this sample according to different confidence levels. In contrast, the sample in this examination considers the observations in the analysis group. If the irregularities are innocent or affect both the incumbent and opposition parties in a similar way, d_k should be similar to a random sample ratio. Moreover, if K is large, the variance of the estimator d_k should approximate to $Var(d_k) = (1 - \frac{k}{K})\frac{s^2}{n}$ with $s^2 = \frac{1}{k-1}\sum_{i\in S_k}(d_k)^2$ Lohr (1999).

Therefore, to test the hypothesis that errors are accidental, under the assumption that k and K are large enough:

$$\zeta_k = \frac{d_k - R}{S_k^2} \qquad (3)$$

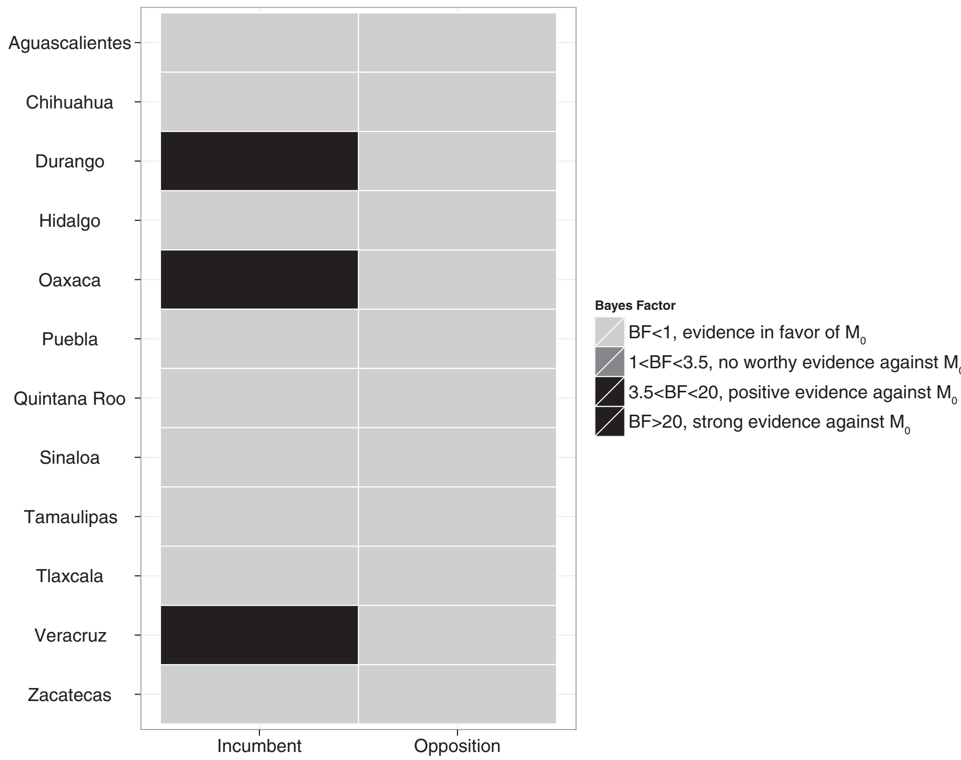
which is expected to be distributed as a standard normal distribution. In this case, a positive and large value of ζ_k means that the observations in the analysis group have a bias in favor of the incumbent party, whereas a negative and large value implies that the opposition party benefited from the irregularities.

The supporting information shows the values of ζ_k across the 12 elections analyzed in this article. The values outside the 95% confidence interval $[-1.96, 1.96]$ are for the elections in Oaxaca and Veracruz. In both cases, it is the candidate of the incumbent party who benefits from the irregularities. There is no reason to assume that random error would benefit the incumbent party, but the distribution of political capital among the candidates makes it plausible that non-accidental irregularities were triggered by the manipulation of votes in favor of candidates supported by the local government and the PRI. Given that the irregularities in Durango only occurred in one precinct, the value for ζ_k in Durango falls within the confidence intervals, and thus shows that this test provides robust results when fraud occurs in several locations.

[18]The interpretation of this factor is as follows (Kass and Raftery 1995):

B_{10}	Evidence against M_0
$B_{10} < 1$	Negative (supports M_0)
$1 < B_{10} < 3$	Not worth more than a bare mention
$3 < B_{10} < 20$	Positive
$20 < B_{10} < 150$	Strong
$B_{10} > 150$	Very strong

FIGURE 8 Bayes Factor for Incumbent and Challenger Parties in Gubernatorial Elections, 2010



manipulation of the vote in favor of candidates supported by the local government and the PRI.[19]

Cross-Validation

The methodology described above finds types of fraud that are often overlooked using alternative forensic tools. To demonstrate this claim, I compare the results from my analysis with some of the standard techniques typically used in electoral fraud analysis and discuss the differences found by this comparison.

I will briefly summarize the three techniques for this analysis. The Second-Digit Benford's Law test employs a chi-square goodness-of-fit test to establish conformity with Benford's Law given the relative frequency of the second significant digit in the vote counts for any given candidate (Mebane 2006). The Last Digit test detects an artificial vote record when the last digit does not resemble a uniform distribution, is serially ordered, and has low frequencies of distant numerals and consecutive digits (Beber and Scacco 2012, 218–20). Finally, the turnout analysis expects that increases in turnout should neither harm nor benefit a candidate in a disproportionate way. In particular, when regressing the number of votes (V) for an alternative as a share of the registered voters in every unit (E) on the turnout level (T), coefficients should fall in the [0, 1] interval (Myagkov, Ordeshook, and Shakin 2009).

The way in which each of the tests provides evidence for manipulation depends upon its specific quantities of interest. For the Second Digit and Last Digit tests, the null hypothesis is rejected when the value of the χ^2 for each test is greater than $\chi^2_{0.05,8}$ for the second digit or $\chi^2_{0.05,9}$ for the last digit. In the case of the turnout analysis, the method raises a red flag when regressing V/E on T, and the coefficient falls outside the interval [0, 1].

Table 2 shows the exact quantities of interest. The Second Digit test finds irregularities in Chihuahua, Quintana Roo, Sinaloa, and Zacatecas, yet it is difficult to assess how many of these results are false positives, as Deckert, Myagkov and Ordeshook (2011) have shown. The Last Digit test finds a significant deviation from the expected uniform distribution in the vote counts of the PRI in Zacatecas. Finally, the turnout analysis does not detect any irregularities. None of these techniques finds irregularities

[19]For a formal explanation on how the asymmetry of resources gives incentives for electoral corruption in Mexico, see Cantú (2013).

Table 2 Analysis of the 2010 Local Elections in Mexico by Other Fraud Detection Algorithms

	Incumbent			Challenger		
	Second Digit (p-value)	Last Digit (p-value)	Turnout (β coefficient)	Second Digit (p-value)	Last Digit (p-value)	Turnout (β coefficient)
Aguascalientes	0.76	0.44	0.26	0.15	0.70	0.63
Chihuahua	0.02	0.38	0.58	0.03	0.76	0.40
Durango	0.48	0.77	0.40	0.37	0.41	0.54
Hidalgo	0.82	0.43	0.50	0.11	0.95	0.46
Oaxaca	0.20	0.43	0.62	0.69	0.68	0.35
Puebla	0.19	0.65	0.45	0.15	0.76	0.46
Quintana Roo	0.00	1.00	0.57	0.99	0.30	0.25
Sinaloa	0.02	0.90	0.29	0.02	0.14	0.70
Tamulipas	0.19	0.10	0.61	0.25	0.34	0.32
Tlaxcala	0.09	0.63	0.47	0.07	0.97	0.40
Veracruz	0.18	0.64	0.39	0.16	0.59	0.45
Zacatecas	0.02	0.45	0.48	0.40	0.02	0.36

in the cases highlighted by the methodology of this article. As such, the evidence suggests that the methodology proposed in this article is a new technique to find electoral irregularities, which complements the extant research tools.

Conclusion

The results of this article have mixed implications for Mexico. For what was a common event in Mexico only a few decades ago, the findings here show a limited number of electoral irregularities in a few states in 2010, thus indicating that fraud is currently the exception rather than the rule. Nonetheless, the findings also suggest that there are still regions where the interests of local powers are in conflict with national-level regulations in Mexico. Specifically, I provide evidence of electoral irregularities that favored the PRI in Durango, Oaxaca, and Veracruz, which represent a quarter of the electoral contests.

This article introduces a novel method to uncover electoral manipulation, a characteristic practice of "brown areas," identified by O'Donnell (1998) as those regions in a country where the interests of local powers might conflict with laws that regulate affairs at the national level. Yet, as recent studies on subnational authoritarianism in Argentina, Brazil, Russia, and the Philippines indicate, the empirical identification of those regions is a challenging task (Borges 2008; Gel'Man and Ross 2010; Gervasoni 2008; Sidel 2004). As long as voters are assigned to polling stations in a way that is not correlated with their electoral behavior, this method can be used fruitfully to analyze electoral contests in other countries.[20]

It should be noted that there are two caveats that require further investigation. First, since the identification of suspicious observations is at the polling station level, the methodology may not identify fraud in those cases where all of the observations in the precinct were tampered with. A plausible extension of this analysis could detect manipulated precincts by comparing turnout differences among those precincts with similar observable characteristics. Second, this article does not evaluate another common type of fraud—namely, the subtraction of votes via nullifying paper ballots. Although this analysis excludes those observations with an atypical number of spoiled votes, it is possible to adapt the methodology to detect differences in voided votes, rather than examining at differences in turnout levels.

Overall, together with an emerging wave of studies on electoral integrity, this article tries to reduce the complexity of fraud denials and allegations in nonconsolidated democracies. On the one hand, this research helps to detect instances in which incumbents might try to "reap the fruits of electoral legitimacy without running the risks of democratic uncertainty" (Schedler 2002, 37). On the other hand, this methodology can evaluate

[20] Examples of elections in which the assignment of voters within electoral wards occurs via family names or voters' identification number are Bolivia, Jamaica, Pakistan, Venezuela, and the U.S. state of Alabama (Alabama League of Municipalities 2011; Electoral Commission of Pakistan 2002; Figueros 1985; Hausmann and Rigobón 2011).

fraud allegations in light of the available evidence to assess the legitimacy of discontent on the part of electoral losers. Moreover, this methodology provides a tool to help observers and authorities evaluate the strengths and weaknesses of an electoral administration.

References

Alabama League of Municipalities. 2011. "Special Report." https://alalm.org/PDF (accessed July 20, 2012).

Alvarez, R. Michael, Jonathan N. Katz. 2008. "The Case of the 2002 General Election. " In *Election Fraud: Detecting and Deterring Electoral Manipulation*, ed. R,. Michael Alvarez, Thad E. Hall, and Susan D. Hyde. Washington, DC: Brookings Institution Press, 149–61.

Aparicio, Javier. 2006. "Fraud or Human Error in Mexico's Presidential Election?"*Voices of Mexico* 77: 7–10.

Aparicio, Javier. 2009. "Análisis estadístico de la elección presidencial de 2006: ?'Fraude o errores aleatorios?"*Política y Gobierno* Volúmen Temático 2009. Elecciones en México. 225–43.

Beber, Bernd, and Alexandra Scacco. 2012. "What the Numbers Say: A Digit-Based Test for Election Fraud Using New Data from Nigeria."*Political Analysis* 20: 211–34.

Borges, André. 2008. "State Government, Political Competition and Education Reform: Comparative Lessons from Brazil."*Bulletin of Latin American Research* 27(2): 235–54.

Cantú, Francisco. 2013. "Cheating and Losing: Electoral Corruption in Competitive Elections." Working Paper. University of California, San Diego.

Carbonell, Miguel. 2003. *El Fin de las Certezas Autoritarias: Hacia la construcción de un nuevo sistema político y constitucional para México*. Mexico City: Universidad Nacional Autónoma de México.

Casta neda, Jorge G. 2000. *Perpetuating Power*. New York: The New Press.

Chib, Siddhartha. 1995. "Marginal Likelihood from the Gibbs Output."*Journal of the American Statistical Association* 90: 1313–21.

Código Federal de Instituciones y Procedimientos Electorales. 2012. Diario Oficial de la Federación. January 14, 2008.

Cornelius, Wayne A. 2002. "La eficacia de la compra y coacción del voto en las elecciones Mexicanas de 2000."*Perfiles Latinoamericanos* 20: 11–32.

Craig, Ann L., and Wayne A. Cornelius. 1995. "Houses Divided: Parties and Political Reform in Mexico. " In *Building Democratic Institutions: Party Systems in Latin America*, ed. Scott Mainwaring and Timothy Scully. Stanford, CA: Stanford University Press.

Crespo, José Antonio. 2006. *2006: Hablan las actas*. Random House Mondadori, Mexico City.

Crespo, José Antonio. 2008. "El modelo del 'fraude hormiga.'*Excelsior*, February 2, 11.

Deckert, Joseph, Mikhail Myagkov, and Peter C. Ordeshook. 2011. "Benford's Law and the Detection of Election Fraud."*Political Analysis* 19(3): 245–68.

de la Madrid, Miguel. 2004. *Cambio de Rumbo*. Mexico, D.F.: Fondo de Cultura Económica.

Dresser, Denise. 2003. "Mexico: From PRI Predominance to Divided Democracy. " In *Constructing Democratic Governance in Latin America*, ed. Jorge I. Domínguez and Michael Shifter. Baltimore: Johns Hopkins University Press, 321–347.

Eisenstadt, Todd A., and Alejandro Poiré. 2006. "Explaining the Credibility Gap in Mexico's 2006 Presidential Election, Despite Strong (Albeit Perfectable) Electoral Institutions." Working Paper. http://www.american.edu/sis/cnas/upload/workingpaper4_elections.pdf (Accessed July 20, 2012).

Electoral Commission of Pakistan. 2002. *Election Laws (Volume 1)*. Pakistan: http://www.ecp.gov.pk/ElectionLaws/Volume-I.pdf (Accesed July 20, 2012).

El Siglo de Durango. 2010. "Una prueba más del fraude electoral fue encontrada en Gómez Palacio." July 15, 6.

Figueros, Mark. 1985. "An Assessment of Overvoting in Jamaica."*Social and Economic Studies* 34(3): 71–106.

Fukumoro, Kentaro, and Yusaku Horiuchi. 2011. "Making Outsiders' Votes Count: Detecting Electoral Fraud through a Natural Experiment."*American Political Science Review* 105(3): 583–603.

Gabinete de Comunicación Estratégica. 2010. "Reporte Postelectoral 2010." Survey. http://www.gabinetece.com.mx/images/uploads/documents/documentos/n˙estudios/reporte˙060810.pdf

Gel'Man, Vladimir, and Cameron Ross. 2010. *The Politics of Sub-National Authoritarianism in Russia*. Burlington, VT: Ashgate.

Gervasoni, Carlos. 2008. "Measuring Variance in Subnational Regimes: Results from an Expert-Based Operationalization of Democracy in the Argentine Provinces."*Journal of Politics in Latin America* 2(2): 13–52.

Gibson, Edward. 2005. "Boundary Control: Subnational Authoritarianism in Democratic Countries."*World Politics* (58): 101–32.

Gillingham, Paul. 2012. "Mexican Elections, 1910–1994: Voters, Violence, and Veto Power. " In *The Oxford Handbook of Mexican Politics*, ed. Roderic A. Camp. New York: Oxford University Press, 53–72.

Hausmann, Ricardo, and Roberto Rigobón. 2011. "In Search of the Black Swan: Analysis of the Statistical Evidence of Electoral Fraud in Venezuela."*Statistical Science* 26(4): 543–63.

Hiskey, Jonathan T., and Shaun Bowler. 2005. "Local Context and Democratization in Mexico."*American Journal of Political Science* 49(1): 57–71.

Hunter, Paul R., and Michael A. Gaston. 1988. "Numerical Index of the Discriminatory Ability of Typing Systems: An Application of Simpson's Index of Diversity."*Journal of Clinical Microbiology* 26(11): 2465–66.

Instituto Federal Electoral. 2010. "Análisis de las Actas de Escrutinio y Cómputo de la Elección de Diputados Federales 2009." Technical report. http://www.ife.org.mx/docs/IFE-v2/ProcesosElectorales/ProcesoElectoral2008-2009/ProcesoElectoral2009-docs/AnalisisActasEscrutinioComputo2009.pdf (Accesed July 20, 2012)

Jeffreys, Harold. 1961. *Theory of Probability*. New York: Oxford University Press.

Jiménez, Raúl. 2011. "Forensic Analysis of the Venezuelan Recall Referendum." *Statistical Science* 26(4): 564–83.

Kass, Robert E., and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90(430): 773–95.

Koenker, Roger, and Kevin F. Hallock. 2001. "Quantile Regression." *Journal of Economic Perspectives* 15(4): 143–56.

Kozumi, Hideo, and Genya Kobayashi. 2009. "Gibbs Sampling Methods for Bayesian Quantile Regression." Technical report, Graduate School of Business Administration, Kobe University.

La Jornada. 2010. "Ratifican triunfo del PRI en Durango." August 16, 31.

Langston, Joy. 2012. *Party Change: Mexico's PRI, from Hegemony to Democracy, 1980 to 2006*. Unpublished manuscript.

Larreguy, Horacio. 2012. "Monitoring Political Brokers: Evidence from Clientelistic Networks in Mexico." Working Paper. Harvard University.

Lehoucq, Fabrice E. 2002. "Can Parties Police Themselves? Electoral Governance and Democratization." *International Political Science Review/Revue Internationale de Science Politique* 23(1): 29–46.

Lehoucq, Fabrice. 2003. "Electoral Fraud: Causes, Types, and Consequences." *Annual Review of Political Science* 6: 233–56.

Levin, Inés, Gabe A. Cohn, Peter C. Ordeshook, and R. Michael Alvarez. 2009. "Detecting Voter Fraud in an Electronic Voting Context: An Analysis of the Unlimited Reelection Vote in Venezuela." In *Proceedings of the 2009 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE '09)*.

Levy, Santiago. 2006. *Progress Against Poverty*. Washington, DC: Brookings Institution Press.

Lohr, Sharon L. 1999. *Sampling: Design and Analysis*. Boston: Brooks/Cole.

Los Angeles Times. 2010. "Mexico's PAN and PRD: Love letters in the making?" February 1. http://articles.latimes.com/2010/feb/01/world/la-fg-mexico-alliance2-2010feb02

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and Its Demise in Mexico*. New York: Cambridge University Press.

Magaloni, Beatriz. 2010. "The Game of Electoral Fraud and the Ousting of Authoritarian Rule." *American Journal of Political Science* 54: 751–65.

Martin, Andrew D., Kevin M. Quinn, and John Hee Park. 2011. "MCMCpack: Markov Chain Monte Carlo in R." http://www.jstatsoft.org/v42/i09/ *Journal of Statistical Software* 42(9): 22.

McElduff, Fiona, Pablo Mateos, Angie Wade, and Mario Cortina. 2008. "What's in a Name? The Frequency of Geographic Distributions of UK Surnames." *Significance* 5(4): 189–92.

Mebane, Walter R. 2006. "Election Forensics: Vote Counts and Benford's Law." Presented at the Summer Meeting of the Political Methodology Society, UC-Davis.

Molinar, Juan. 1987. "Regreso a Chihuahua." *Nexos* 10(111): 21–32.

Moreno, Alejandro. 2012. "Who Is the Mexican Voter?" In *The Oxford Handbook of Mexican Politics*, ed. Roderic A. Camp. New York: Oxford University Press, 571–95.

Myagkov, Mikhail, Peter C. Ordeshook, and Dimitri Shakin. 2009. *The Foresincs of Election Fraud: Russia and Ukraine*. New York: Cambridge University Press.

Navia, Patricio. 2000. "Fair Play at Voting Precincts: A Comparison of Mexican and Chilean Elections." *Democratization* 7(1): 183–89.

New York Times. 2010. "Once-Dominant Party in Mexico Seems Resurgent." July 4, A7.

Ochoa-Reza, Enrique. 2004. "Multiple Arenas of Struggle: Federalism and Mexico's Transition to Democracy. " In *Federalism and Democracy in Latin America*, ed. Edward L. Gibson. Baltimore: Johns Hopkins University Press 255–296.

O'Donnell, Guillermo. 1998. "Polyarchies and the (Un)Rule of Law in Latin America." Presented at the Meeting of the Latin American Studies Association, Chicago.

Peschard, Jacqueline. 2010. "Federal and Local Electoral Institutions: From a National to a Fragmented System." In *Mexico's Democratic Challenges: Politics, Government, and Society*, ed. Andrew Selee and Jacqueline Peschard. Stanford, CA: Woodrow Wilson Press Center. 68–91.

Poiré, Alejandro, and Luis Estrada. 2006. "Allegations of Fraud in Mexico's 2006 Presidential Election." Presented at the 102th meeting of the American Political Science Association.

Polasek, Wolfang. 2003. *Multivariate Regression and ANOVA Models with Outliers: A Comparative Approach*. Economic series. Vienna: Institute for Advanced Studies.

Prats, Juan José Rodríguez. 2010. "El Trife no puede convalidar Durango, Hidalgo y Veracruz: Correligionarios o cómplices?" *Siempre!* 57(2982): 38.

Preston, Julia, and Samuel Dillon. 2004. *Opening Mexico: The Making of a Democracy*. New York: Farrar, Straus and Giroux.

Raphael, Ricardo. 2007. *Los Socios de Elba Esther*. Mexico City: Planeta.

Schedler, Andreas. 2002. "The Menu of Manipulation." *Journal of Democracy* 13(2): 36–50.

Sidel, John. 2004. "Bossism and democracy in the Philippines, Thailand, and Indonesia: Towards an alternative framework for the study of 'local strongmen."' In *Policising Democracy: the New Local Politics of Democratisation.*, ed. John Harriss, Kristin Stokke, and Olle Tornsquist. Basingstone, UK: Palgrave Macmillan, 51–74.

Simpser, Alberto. 2012. "Does Electoral Manipulation Discourage Voter Turnout? Evidence from Mexico." *Journal of Politics* 74(3): 782–95.

Simpser, Alberto. 2013. *Why Governments and Parties Manipulate Elections: Theory, Practice, and Implications*. New York: Cambridge University Press.

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

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Results of the Bayesian regressions and sensitivity analyses (Tables S1-S5).

Summary of the electoral results for Governor elections in Mexico (Table S6).

Tests for the distribution of last names in Mexico (Tables S7-S11 and Figures S1-S15).