PATTERNS OF FRAUD: TOOLS FOR ELECTION FORENSICS

by

JOSEPH DECKERT

A DISSERTATION

Presented to the Department of Political Science
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2013

DISSERTATION APPROVAL PAGE

Student: Joseph Deckert

Title: Patterns of Fraud: Tools for Election Forensics

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Political Science by:

Mikhail Myagkov	Chairperson
Priscilla Southwell	Core Member
William Terry	Core Member
George Evans	Institutional Representative

and

Kimberly Andrews Espy	Vice President for Research and Innovation; Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2013

DISSERTATION ABSTRACT

Joseph Deckert

Doctor of Philosophy

Department of Political Science

June 2013

Title: Patterns of Fraud: Tools for Election Forensics

Election fraud is a serious problem in a number of modern democracies. While election observers can combat this problem to some extent, election forensics aims to provide a low-cost supplement. Forensic tools uncover irregular patterns in aggregate election data which are consistent with fraud. This dissertation improves upon existing tools and establishes methods of controlling for other factors which could cause irregular patterns. These tools are utilized in three cases studies in an effort to better understand the nature of election fraud.

CURRICULUM VITAE

NAME OF AUTHOR: Joseph Deckert

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
California State University, Hayward
Palomar College, Vista, California

DEGREES AWARDED:

Doctor of Philosophy, Political Science, 2013, University of Oregon
Master of Science, Political Science, 2012, University of Oregon
Master of Science, Economics, 2012, University of Oregon
Bachelor of Arts, Political Science, 2005, California State University, Hayward

AREAS OF SPECIAL INTEREST:

Formal Theory and Methodology
Political Economy

PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, Department of Political Science, University of Oregon, 2007-2012

GRANTS, AWARDS, AND HONORS:

Graduate Teaching Fellowship, Political Science, 2007-2012

ACKNOWLEDGMENTS

I would like to express a great deal of gratitude to my advisor, Mikhail Myagkov. His support and assistance was of tremendous value. I also thank Dr. William Terry, who provided a substantial amount of direction and feedback. This research was supported in part by a William C. Mitchell Summer Research Grant from the Department of Political Science, University of Oregon.

For my girls: Molly, Anna, and Allison

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

The history of modern election fraud is as old as the history of modern elections. While democracy expanded in the 20th century and continues to spread worldwide (Huntington 1991), accusations of election fraud are endemic in developing democracies and continue in established regimes (Schedler 2002). While elections are more common, it is not clear that elections have everywhere resulted in democracy.

Election monitoring on the ground has become an important mission of the United Nations along with myriad other international organizations (Hyde 2008; Kelley 2009). Despite the positive impact that monitoring has had on legitimacy, there are several problems with its application. Election monitoring is expensive, and some critics have argued that it is often aimed at a show of legitimacy rather than actual effectiveness (Carothers 1997). Complex elections with multiple monitoring agencies can create problems when the legitimacy of an election is disputed (Kelley 2009).

Election forensics is an alternative approach to gauging the legitimacy of an election based on examining patterns in the data and uncovering "irregularities" which are consistent with election fraud. This statistical analysis has the advantage of being a low-cost supplement to existing monitoring, since it relies on data that is readily available. Election forensics can aid election observers by providing information about the nature and location of these potentially fraudulent outcomes.

Three important caveats apply when using forensic tools. First is that these tools detect large-scale fraud but do not detect small scale fraud (Myagkov and Ordeshook

2009, p. 5). Fraud on the scale of a few hundred or few thousand votes will be lost in the noise in elections with millions of voters. Second, these tools typically require that fraud is present in some locations and not others. Like a polygraph test, we need something of a baseline. In order to distinguish patterns of fraud or absence of fraud, clear differences must actually exist. Unlike a polygraph, we do not need any *a priori* information about where fraud may have occurred. For these two reasons, election forensics cannot replace election monitors but can supplement their work. The third limitation of election forensics is that current methods require some fairly restrictive assumptions. These assumptions pertain to the assumed distribution of vote counts, turnout, and other indicators in the absence of fraud. This dissertation improves on the tools available to address this weakness.

This study makes three main contributions to the field of election forensics. First, I use computer simulations to develop the theoretical foundation of election forensics and improve on existing methods. Forensic tools have focused on finding symptoms that could be caused by fraud. But just as a headache could be caused by a number of different ailments, irregularities in the data could be caused by fraud or a number of other factors. With a more complete theory of how fraud changes patterns in the data we may uncover the actual cause of observed irregularities and relax the assumptions made in previous work. Second, I use the tools of election forensics to examine two instances that occurred during the "third wave" of democratization. While elections have become more legitimate in Russia and Mexico, it appears that the latter has been more successful in actually attaining democracy. Finally, turning to the United States, I use election forensics to assess the claims of political scientists in regards to the establishment of the Solid South. With the end of Reconstruction in 1877, Democrats began disenfranchising the opposition and creating a single-party regime. Experts in Southern politics disagree as to the effectiveness of the myriad tools used to achieve this end, and by using election forensics I can provide estimates of the severity of vote suppression over time. In short, I improve on and expand forensic tools and put them into practice in three significant cases.

Previous Work in Election Forensics

One of the most comprehensive sets of tools available for detecting these patterns is that developed by Myagkov, Ordeshook, and Shakin (2009). The approach that they recommend involves analyzing patterns in the distribution of turnout rates and examining the relationship between the percentage of the electorate which votes for a candidate (absolute support) and turnout rate. They suggest that turnout tends to follow a normal distribution. If some election officials stuff the ballot box with votes for one candidate, those districts will have a turnout rate that is artificially increased. As some data is shifted to the right of the distribution, it will become skewed rather than normal. Additionally, Myagkov et al. assume that the relationship between a candidate's *absolute* support (the percentage of eligible voters who vote for that candidate) and turnout should be positive and roughly equal to that candidate's *relative* support (the percentage of ballots cast for that candidate). This relationship is fairly intuitive. Assume that a candidate receives, on average, 50% of the vote. We would expect that if 10 additional voters came to the polls, roughly 5 would vote for that candidate. In other words, the percentage of votes a candidate receives should not differ whether a district has low or high turnout. If vote inflation occurs in favor of one candidate, however, it will appear that the candidate has stronger support in high turnout districts. This is due to the fact that some of those high turnout districts will have had their vote counts artificially inflated in favor of that candidate.

Although these developing tools have been subject to internal debate (Deckert, Myagkov et al. 2011; Mebane 2011), there remains the promise of an objective research agenda that can evaluate the legitimacy of elections. This dissertation seeks to advance that project.

Formal Models of Election Fraud

The literature on election forensics and the literature on decision-making regarding fraud have largely been disconnected. In the third chapter of the dissertation I attempt to unite these by creating models that will aid in the analysis of election legitimacy.

Magaloni (2006) develops a game-theoretic model aimed at explaining why authoritarian regimes might hold elections, and why they might “tie their hands” when it comes to election fraud. Fearon (2011) also uses game theory to explain in detail the behavior of authoritarian regimes. His coverage of election fraud in the model is essentially a mere assumption that autocratic regimes will engage in as much fraud as is possible without getting caught. Chaves et al. (2009) use choice theory to predict when local election officials will engage in fraud for the authoritarian regime. Their model is aimed at uncovering the relationship between regional characteristics and fraud decisions, but this model uses fairly unsophisticated metrics for fraud.

I put forth two models of regime behavior. The first is developed in a vein similar to that of Magaloni (2006). This model provides a frame of reference for authoritarian regimes, allowing for comparisons across nations. I also discuss the generalizability of Chaves' (2009) model, which has direct empirical applicability. This model analyzes the decisions of local election officials and aids in understanding how choices regarding the type and degree of fraud are made. The Chaves model can be tested with the results obtained from election forensics.

Case Studies

As a key part of developing the toolkit for detecting election fraud, the dissertation covers a number of relevant case studies.

First Case Study: Russia (Chapter 4)

Russian elections were typified by fraud before the dissolution of the Soviet Union in 1991, and this trend continued afterward (Myagkov, Ordeshook et al. 2009). The legislative election of 2007 is one particularly interesting example, given the fallout of election fraud. The republic of Ingushetia reported 98% turnout. The "I did not vote" campaign was started by dissidents, and received signatures and passport numbers from roughly a third of the electorate who asserted that they did not vote in the election (Myagkov, Ordeshook et al. 2009, p. 128). The initiator of this campaign was later arrested, and was shot and killed while in police custody.

Despite attempts at reform, I find that election fraud in Russia remains pervasive. I analyze major national elections from 2007 through 2012 and find no major reductions in observed irregularities.

Second Case Study: Transition to Democracy in Mexico (Chapter 5)

The rise of democracy in Latin America has been tumultuous, with allegations of election fraud in many new democracies (Chaves, Fergusson et al. 2009; Cothran 1994; Lehoucq and Molina 2002). Mexico provides a particularly interesting case. Mexico has long been a nominal democracy, and has been highly stable since the cessation of the Mexican Revolution (Cothran 1994). Mexico began to transition to democracy following economic hardships in the 1980s (Beer, 2003). Although prospects for legitimate democracy seemed bright following the 2000 presidential election, allegations of fraud returned in the 2006 and 2012 elections. Despite fears that the country may be reverting to corruption and illegitimacy, I find that elections in Mexico appear relatively fraud-free. The reforms of the 1990s appear to have a lasting and positive impact on the legitimacy of elections in Mexico.

Third Case Study: Election Fraud in the United States (Chapter 6)

As one of the world's oldest democracies, the United States has a long history of elections and a long history of election fraud (Campbell 2005). In Chapter 6, I analyze the Reconstruction and Redemption eras in the U.S. South. Accounts from the time indicate that vote manipulation was rampant in the form of disenfranchisement (Foner 1988). Scholars disagree in regards to the type and timing of this disenfranchisement. Some argue that extralegal means were employed beginning with the end of Reconstruction (Key 1949), while others argue that legal disenfranchisement had a more serious impact during the height of Redemption (Kousser 1974).

Forms of Fraud

In the most broad terms, there are only two ways of engaging in election fraud: votes can be added, or votes can be subtracted. I will refer to the former as vote inflation, and to the latter as vote suppression. Of course, both could occur at the same time. This is mathematically equivalent to switching votes from one candidate to another, and I will refer to this third possibility as vote stealing. These three types of fraud are explained in Table 1.1, along with examples and the forensic tools that can be used to detect them.

There is ample evidence that all three of these types of fraud have been used to alter election results. While there are many different ways to engage in fraud, any actual mechanism will have one of these three effects. Vote inflation is typically associated with "stuffing the ballot box," where election officials fill out fraudulent ballots or simply count extra votes for one candidate. The same effect could be achieved if a subset of voters votes multiple times. Vote suppression could occur at the official level as well, with election officials destroying votes for a particular candidate. There are myriad other ways to achieve the same effect. Voter intimidation is a common tactic, and violence has often surrounded elections. Vote suppression can be nonviolent as well, and allegations of this type of fraud are common. Limiting the voting equipment available to certain precincts or placing obstacles such as road closures or insufficient parking could be used to target certain voting populations in order to reduce their vote counts. Again, if both inflation and suppression happen at the same time, the result is effectively switching votes from one candidate to another. This one-to-one change can happen more directly, however. One of the oldest traditions in vote fraud is the buying of votes (Campbell

2005). Though crude and of arguable efficiency, this technique could be employed to get voters themselves to switch their vote from the candidate of their preference to another. Of course, election officials could easily engage in this type of behavior by miscounting votes for one candidate as votes for another.

Table 1.1. Types of election fraud

Form of Fraud	Definition	Means	Example Cases	Forensic Tools
Vote Inflation	Adding votes to a candidate's total	Ballot box stuffing	Russia	Myagkov; Benford's Law
Vote Suppression	Subtracting votes from a candidate's total	Voter intimidation, destruction of ballots	United States, Ukraine, Colombia	Myagkov; Benford's Law
Vote Stealing	Switching votes from one candidate to another	Combination of the above, miscounting of ballots	Ukraine, Mexico	Myagkov, with *a priori* knowledge

Conclusion

This dissertation makes three significant contributions to the study of election fraud. First, it expands on the existing tools to analyze election data for evidence of fraud. Second, it tests existing methods and develops new ones for dealing with natural non-normality. Aggregation error due to underlying regional characteristics is one of the largest obstacles facing analysis of election data. By discovering ways of relaxing the homogeneous regions assumption, these tools will be much more applicable. Third, the project uses these advanced tools to perform comprehensive analysis of elections in several regions. While there has been analysis of specific elections using these tools, it has typically been narrow in focus and instrumental to developing the tools. Studies that have undertaken more in-depth analysis of elections have typically relied on weaker measures of fraud. My dissertation bridges the gap between these two sets of literature.

CHAPTER II

METHODS FOR DETECTING AND MEASURING ELECTION FRAUD

Following a brief definition of election forensics, Chapter II will examine in detail methods for detecting election fraud, including indicators, simulation, and alternative ways of identifying election fraud. I will summarize the forensic tools currently employed by election scholars, and expand one tool developed by Myagkov et al. (2009) to expose vote stealing where it exists in addition to vote inflation. After developing simulations, I explain the problems associated with normality assumptions and propose two methods for controlling for variables that might otherwise be problematic for forensic analysis. Finally, I will create a measure of election fraud derived from the forensic tools developed in this chapter.

Election Forensics

Election forensics is the relatively new field of assessing the integrity of elections based on patterns in the election results. The two main schools of election forensics include digit analysis and turnout analysis. The former inspects the frequency of certain digits in vote counts (e.g., Beber and Scacco 2008; Cantu and Saiegh 2011; Mebane 2006), while the latter looks for anomalies in reported turnout (e.g., Herron 2010; Myagkov 2009).

The established literature in both schools has focused on uncovering irregularities that are symptomatic of election fraud. I attempt to take the project further by explaining: (1) how fraud causes irregularities, (2) how other factors can cause irregularities, and (3) how to determine which is the cause when the data are irregular.

Established Indicators of Election Fraud

Digit Analysis

One popular method of analyzing elections for evidence of fraud relies on analyzing the distribution of particular digits in vote counts. These methods have the benefit of requiring little in the way of data as they rely solely on vote counts (Beber and Scacco 2008, p. 2; Mebane 2006). They analyze the distribution of digits in reported vote counts and compare the actual distributions to those expected in fraud-free elections.

In particular, Benford's Law holds that the distribution of digits in data does not follow a uniform pattern. Traditionally, Benford's Law is applied to the first digit in a set of data, and holds that the digit 0 will be the most common, 1 the second most common, and so on. This result is found when data has a logarithmic distribution. If the distribution of the data is skewed such that lower numbers are more common than higher numbers, lower digits will also be more common than higher digits. Naturally occurring data often have such a skewed distribution, such as with income or population size. Data which have a symmetric distribution do not have such a pattern. Instead, digits in symmetric data are likely to be distributed uniformly.

Benford's Law can be observed in a wide range of data, including population sizes, disk space usage, and many other types of naturally occurring data (Fewster 2009). Deviations from Benford's Law have been used as indicators of fraud in accounting and public finance, as they can indicate “non-natural” data generation processes. Since vote counts often appear to have a distribution which conforms to the law, proponents of this method argue that deviations may be indicative of fraud (Cantu and Saiegh 2011; Mebane 2006).

For example, Cantu and Saiegh (2011) use Benford's Law to analyze the first digit of vote counts to examine patterns of fraud in Argentina's presidential elections during the “infamous decade.” Additionally relying on the distribution of the first digit, they develop a Naïve Bayesian classifier which uses information from the first digit of vote counts to categorize elections as fraudulent or clean. They find that the results obtained by using this classifier match well with the historical evidence on the elections they study.

Mebane (2006) argues that vote totals display a Benford compliant distribution not for their first digit, but for their second. He argues that because precinct sizes are artificially carved up to be roughly similar rather than naturally occurring, Benford's Law will not hold for the first digits (1BL). Instead, precinct sizes comply with Benford's Law for second digits (2BL). This pattern remains because vote counts are not the result of a simple random process, but rather a complex process involving mathematical combinations of random variables from different distributions (Mebane 2006, p. 4).

Analysis of the last digit in vote counts is another method of uncovering patterns of fraud along these lines (Beber 2008). While analyzing first or second digits requires assumptions that the underlying data should conform to Benford's Law, proponents argue that analysis of the last digit requires “extremely weak distributional assumptions” (Beber 2008, p. 2). The key assumption is that vote counts should have a last digit that follows a uniform distribution. Since the data are naturally occurring, we should not expect any digit to appear with a greater frequency than any other. If the data has been manipulated, however, particular digits will appear with greater frequency because humans are not good at generating random numbers (Beber 2009, p. 6). Essentially, if vote counts have been generated by corrupt election officials rather than from actual ballots, the last digits will not follow a uniform distribution.

While digit-based methods work well in cases where the available data are limited, they have some weaknesses. Digit-based methods are limited in the forms of fraud which they can detect. Mebane points out that 2BL analysis works best on what he calls “repeaters,” a form of vote-inflation, or when fraud is correlated with a candidate's level of support (Mebane 2006, pp. 15-18). Analysis of the last digit only detects fraud where actual numbers have been generated by humans making up numbers, and as such will not detect common forms of fraud such as ballot box stuffing. Additionally, the relevance of Benford's Law in election data has been called into question (Deckert et al. 2011). While it is true that election data often conform to Benford's Law, it is less clear that election fraud will affect the distribution of digits. Moreover, deviations from Benford's Law are difficult to attribute to fraud. In essence, analysis relying on Benford's Law can easily result in both type I and type II errors and it is not clear how common these errors are or what can be done to correct for them.

Given their potential strength when data are limited, digit-based methods have a place in the election forensics toolbox. Due to their limitations, however, this dissertation focuses on alternative methods that have stronger theoretical foundations and can be expanded to incorporate better data.

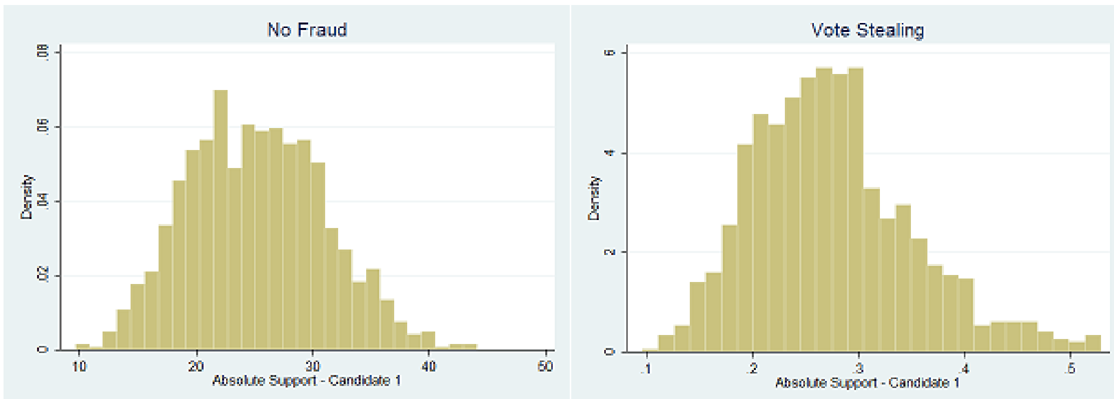
Distributions of Turnout

In a clean election, we expect the distribution of turnout rates to be normal at low levels of aggregation (Myagkov et al. 2009, p. 45). Most elections in stable, competitive democracies have roughly normal distributions of turnout at precinct or district levels. We expect that the distribution will be single-peaked and relatively symmetric (an assumption which will be relaxed in subsequent sections). Suspicious distributions are either skewed, having a "fat tail" where either high or low levels of turnout are more common, or they have a second peak. The intuition for this result is fairly simple. In the case of ballot box stuffing, when votes are added turnout is increased. If turnout is originally normal, and votes are illegitimately added in a subset of polling places, some of the data will be shifted to the right as in Figure 2.1, resulting in an asymmetric distribution. If votes for a candidate are discarded, the opposite pattern will be visible. As turnout is effectively lowered in some areas, parts of the distribution will be shifted to the left, resulting in a wider left-hand tail on the distribution. If both vote inflation and vote destruction occur at the same time, the effect on the distribution of turnout will be lessened. If the inflation and destruction happen at a one-to-one ratio (i.e., vote stealing), turnout may appear normally distributed despite the occurrence of election fraud.

The key assumption made here is that the districts in question are "reasonably homogeneous and that variations in turnout are the consequence of random or uncorrelated factors" (Myagkov et al. 2009, p. 33). If this assumption does not hold, turnout may not follow a normal distribution without election fraud.

Proposition 2.1: Assuming that true turnout has a normal distribution, ballot box stuffing will lead to skewness in the distribution. A skewed distribution of turnout is therefore consistent with vote inflation.

Figure 2.1. Turnout before and after vote inflation. Synthetic data.



Absolute Support and Turnout

Absolute support refers to the percentage of the eligible electorate that votes for a particular candidate. If a candidate gets half of the votes cast in a region where turnout was 50%, that candidate's absolute support will be 25%. Relative support refers to the percentage of votes cast that a candidate receives.

In clean elections, absolute support and turnout are expected to be positively correlated. As more ballots are cast, each candidate should expect to receive more votes. In a regression of absolute support on turnout, we expect the coefficient on turnout to be roughly equal to a candidate's average level of relative support for the election. If a candidate receives 50% of the vote and there is an increase of 2 voters, we would expect that one of those two would vote for that candidate. This pattern is affected by ballot box stuffing, in which every additional vote goes to the candidate who is benefiting from the fraud. In some districts there are additional votes, but all of those votes go to a single candidate. This instance of fraud will be visible in a graph of absolute support and turnout as the regression line of the recipient of fraudulent votes will be pulled upwards by those data points that are shifted.

Simulating Election Fraud

In order to examine the effects of different types of fraud on election returns, I use stylized data from simulated elections. Cantu and Saiegh (2011) note that synthetic data has several advantages for developing and evaluating forensic techniques:

> First, we can generate more data sets than what would be available using only real data. Second, properties of synthetic data can be tailored to meet various conditions which may not be clearly observable in the real data. Third, variations of known frauds (or new frauds) can be artificially created to study how these changes affect performance parameters, such as the detection rate (Lundin, Kvarnstrom, and Jonsson 2002). (Cantu and Saiegh 2011, p. 411)

By using synthetic data it is possible to directly observe the impact of fraud on the patterns in the data. Furthermore, data that violate the normality assumptions of Myagkov et al. (2009) can be generated. The simulations can be programmed to result in distributions of turnout which are not normal due to factors other than fraud. Such data can be used to develop and test new techniques which relax those assumptions. This would not be possible with "real world" data, as the underlying data generating process is unknown. In real world data, we cannot be sure as to the precise impact of election fraud or regional characteristics on the distributions of turnout and the other indicators of fraud.

The simulations used here are an extension of the simulations developed by Deckert, Myagkov, and Ordeshook (Deckert, Myagkov et al. 2011). These simulations begin with a two-dimensional policy space. Two candidates are positioned in that space, with their locations being set exogenously. Voters are placed in the policy space, their position dependent on a combination of random variables. Once placed, voters make their turnout decision. This decision is also based on a random variable. Those who decide to vote cast their vote for whichever of the two candidates is located closer to them in the policy space.

Voter i's ideal point (x_i, y_i) is given by:

$x_i = b_{1x} * v_{1i} + b_{2x} * v_{4i} + e_{ix}$,

$yi = b_{1y} * v_{2i} + e_{iy}$,

where v_{1i} and v_{4i} are random-normal variables and b_{1x} and b_{2x} are constants that determine the effect of those variables on the voters' x position. The noise terms have a normal distribution such that $e_{ix} \sim N(0, \sigma_x^2)$ and $e_{iy} \sim N(0, \sigma_y^2)$ for each voter i.

Voters decide to vote if:

$t_i = b_{1v} * v_{3i} + b_{2v} * v_{4i} + e_{it} > T$,

where ti is the voter's "turnout score," v_{3i} and v_{4i} are random-normal variables and b_{1v} and b_{2v} are constants that determine the effect of those variables on whether or not the voter actually votes; e_{it} is a random-normal error term, and T is the "turnout threshold"—the "turnout score" that a voter must exceed in order to cast a vote.

In each district, the characteristic variables of the voters (v_{1i}, v_{2i}, v_{3i}, and v_{4i}) are distributed random normal, such that:

$v_{1i} \sim N(m_1, \sigma_1^2)$, $v_{2i} \sim N(m_2, \sigma_2^2)$, $v_{3i} \sim N(m_3, \sigma_3^2)$, and $v_{4i} \sim Exp(m_4)$

where m is the district level mean for that variable. Here, m is distributed normally between districts, such that:

$m_1 \sim N(M_1, \sigma_{m1}^2)$, $m_2 \sim N(M_2, \sigma_{m2}^2)$, $m_3 \sim N(M_3, \sigma_{m3}^2)$, $m_4 \sim Exp(M_4)$.

In these simulations, fraud takes the form of either vote inflation or vote stealing.

Fraud is always applied in favor of candidate 1. When vote inflation occurs, candidate 1's vote total is increased by a given percentage and capped by the number of nonvoters. When vote stealing occurs, a percentage of candidate 2's votes are given to candidate 1. In either case, the percentage varies between districts and follows a random-normal distribution. The likelihood of fraud may vary based on the level of support for candidate 1.

Alternative Indicators of Election Fraud

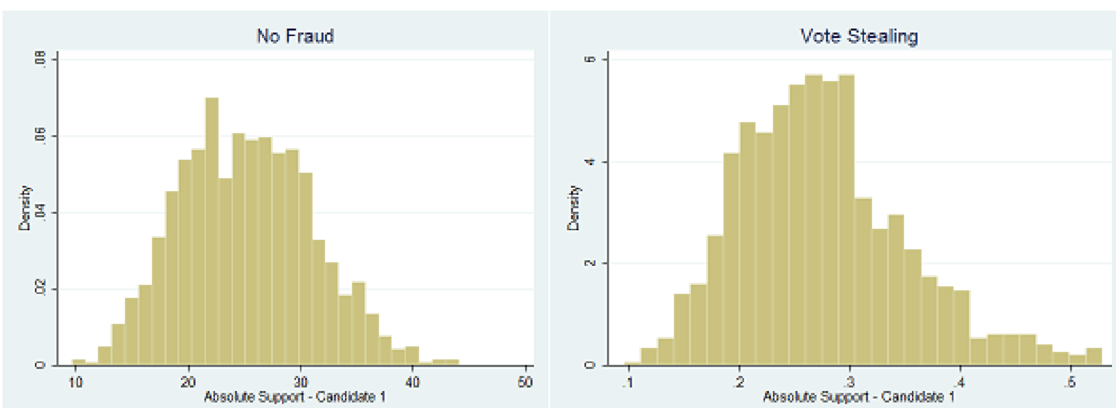
Analysis of Support

If we assume that both relative support and absolute support for candidates are normally distributed across districts, election fraud will cause shifts in these distributions in the same way that fraud causes shifts in the turnout distribution. In elections where turnout is low and the variance of absolute support is high, absolute support may have a truncated normal distribution. Since absolute support cannot be below zero, the left-hand tail may be short. In these cases, analysis of the distribution of absolute support might not be as useful as analysis of relative support.

While analysis of turnout and the absolute support-turnout relationship are able to discover patterns consistent with vote inflation and vote suppression, they are less suited for distinguishing between clean elections and vote stealing. Vote stealing may appear in the relationship between absolute support and turnout, but is difficult to uncover without a priori knowledge (Myagkov, Ordeshook et al. 2009, p. 47), due to the fact that vote stealing will not appear in turnout distributions, and its patterns in absolute support versus turnout could be drowned out by noise and other factors.

Although turnout is not affected by vote stealing, levels of support are. Therefore, looking at the distributions of relative support can be analyzed for consistency with vote stealing. Analysis of absolute support can provide information as to the type of fraud employed. If vote manipulation comes in the form of vote inflation, the absolute support histogram for the perpetrator will shift to the right as in Figure 2.2. The distribution of absolute support for the victim will not change. On the other hand, if vote suppression is driving the shift, the victim will see a leftward shift in distribution of absolute support, while the perpetrator's distribution will be unchanged. In either case, turnout will shift as well. If vote stealing is the underlying mechanism of fraud, turnout will remain normal but *both* candidates will see shifts in their distributions of absolute support. It is likely that any fraudulent election will have instances of all three types of fraud. This diagnostic, however, will help determine the extent of the three forms of fraud.

Figure 2.2. Absolute support for a candidate before and after vote stealing.



Corrections for Natural Non-Normality

Analysis of the distributions of turnout and support rests on the assumption that, in a clean election, turnout and support will be normally distributed. Further, analysis of the turnout and absolute support relationship assumes that the correlation between the two should not be very different than a candidate's relative support—that is, support for a particular candidate is not a function of turnout.

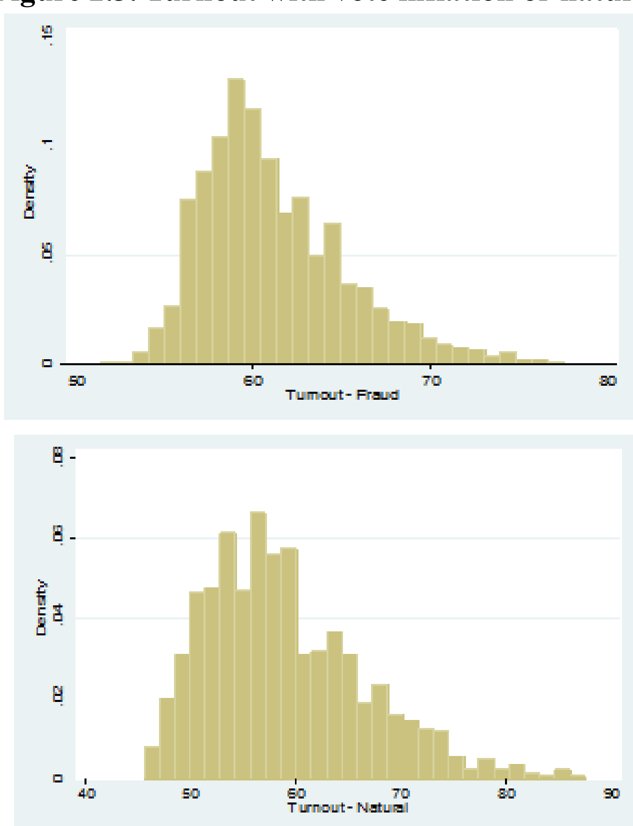
An underlying variable may cause turnout to be skewed if the variable is correlated with turnout and does not have a symmetric distribution. If this same variable is also correlated with support for a candidate, it might appear that the election results are irregular due to vote inflation.

As a simple example, higher levels of wealth might be correlated with increased turnout. In addition, wealth may also be correlated with support for the candidate who is further to the right in the ideological space. At the precinct level, wealth may have an asymmetric distribution. There may be a small number of wealthy precincts and a relatively large number of poorer precincts. This could result in a skewed distribution of turnout. Those few wealthy high-turnout districts may also have higher than average support for the more right-leaning candidate. This could lead to the erroneous conclusion that ballot box stuffing occurred in favor of that candidate.

I refer to this problem as "natural non-normality," a specific form of heterogeneity that leads to distributions of turnout and support which are skewed by variables other than election fraud. Figure 2.3 illustrates this problem. Both distributions appear skewed with a wide right-hand tail. In one case this skewed effect is caused by vote inflation, and in the other, by natural non-normality. Myagkov, et al. (2009) analyze the “flow of votes” from one election to the next. They expect that candidates will receive a similar number of votes from one election to the next and that substantial deviations from this pattern indicate foul play. I propose two alternative methods for addressing this problem.

In this section, I develop two alternatives to address the problem of natural non-normality. The first, the “Jaws” method, does not require any additional data. Instead, it looks at patterns in the residuals obtained by regressing turnout on absolute support. The second, the residual analysis method, requires information on additional variables that affect turnout and support. It analyzes the residuals obtained by regressing turnout or support on these variables.

Figure 2.3. Turnout with vote inflation or natural non-normality.



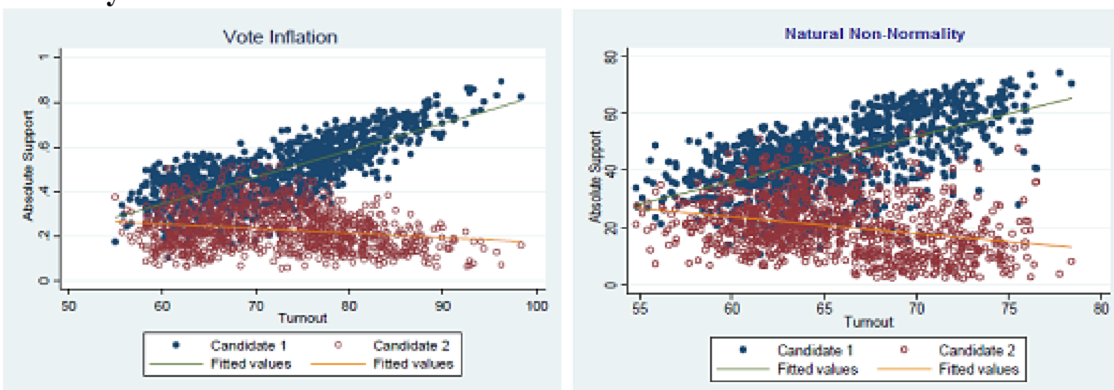
The Jaws Method

Natural non-normality leads to patterns in the data similar to those seen as a result of vote inflation. A crucial difference exists, however, that may reveal the underlying process. At the high end of turnout, most districts fall above the regression line for the candidate who benefited from vote inflation, and most districts fall below the regression line for the candidate whose votes were not altered.

Each district has two entries on the graph, one for each candidate. As vote inflation occurs for one candidate, both of these entries are shifted. For the candidate who receives additional votes, the data point is shifted up and to the right at a 45 degree angle. If there are 100 eligible voters in a district, and 1 more vote is added to the candidate, this candidate's absolute support is increased by 1% and turnout is also increased by 1%. This shift will virtually always place the district above the regression line, because the maximum value for the slope of the regression line is 1 (each vote going to the candidate in question). The data point for the other candidate will be shifted also, but to the left. This outcome is due to the fact that the candidate's vote count, and therefore absolute support, does not change.

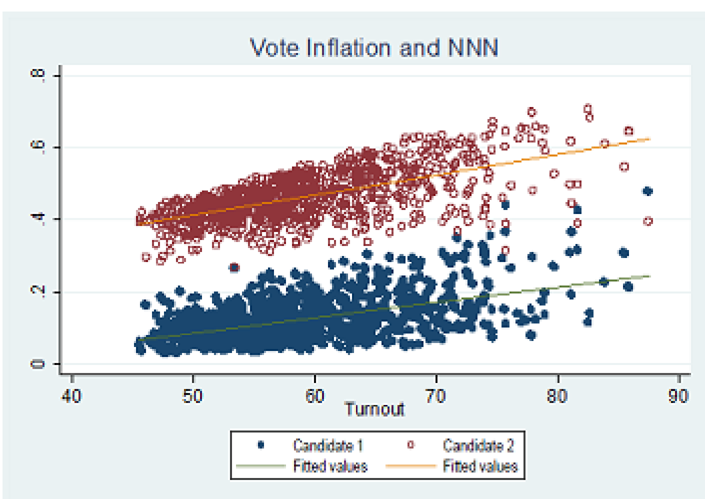
The turnout in the district is changed, however, so this data point will be shifted to the left. When votes are inflated for one candidate after the fact, the data is shifted in a way that forces the regression lines away from each other, like opening jaws. By contrast, the districts with natural non-normality have data which is distributed on both sides of the regression lines, regardless of levels of turnout. Here the data is not shifted systematically, but is instead the result of naturally-occurring variation. This pattern can be observed in Figure 2.4. At the high end of turnout the data points tend to lie above the regression line in the presence of fraud, while they are distributed on both sides of the line in the presence of natural non-normality.

Figure 2.4. Turnout versus absolute support. Vote inflation and natural non-normality.



This method poses two drawbacks. First, it is not always clear that outliers are a result of election fraud. Second, when there is both election fraud and natural non-normality, this method cannot really tell the difference (see Figure 2.5).

Figure 2.5. Both NNN and vote inflation.

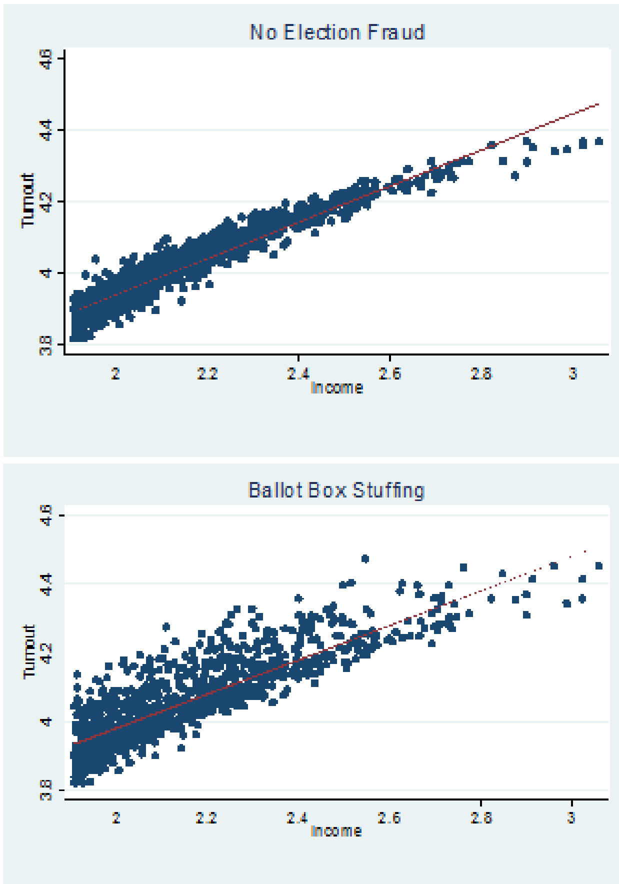


The Residual Analysis Method

We can determine whether or not turnout is normally distributed when controlling for other factors by regressing turnout on those variables and evaluating the normality of the residuals. This method requires more information than vote counts and turnout. Without additional information, however, election forensics may remain susceptible to something akin to omitted variables bias. In order to determine whether irregularities are caused by fraud or by other factors, we need information on those other factors.

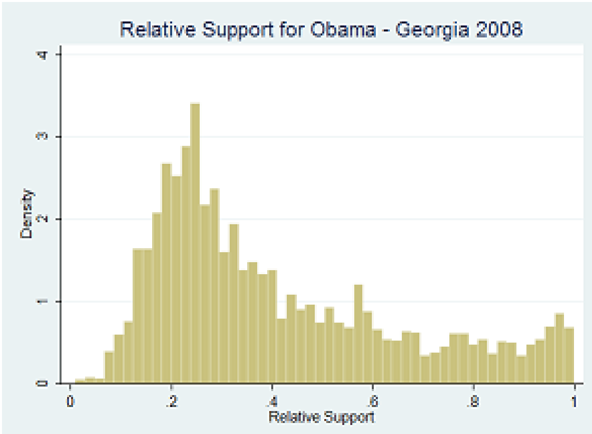
This method assumes that the residuals obtained by regressing turnout or support on explanatory variables are normally distributed in clean elections. If fraud is an important variable in determining turnout or levels of support, then the regression may display irregularities. As turnout and support are artificially increased in some precincts, the data shifts in a predictable way. If turnout is plotted on the y-axis and an explanatory variable (e.g., income) is plotted on the x-axis, as vote inflation occurs some of the data points shift vertically as their reported number of votes is increased. This pattern can be seen in Figure 2.6. When the regression is performed, the regression line also moves vertically. In an extreme case, the data points will appear to form two distinct lines. Since the regression line would fall in the middle, the residuals will have a bimodal rather than a normal distribution. The shift in the distribution of the residuals is similar to the shift in turnout or absolute support as discussed in the preceding sections.

Figure 2.6. Simulation data.



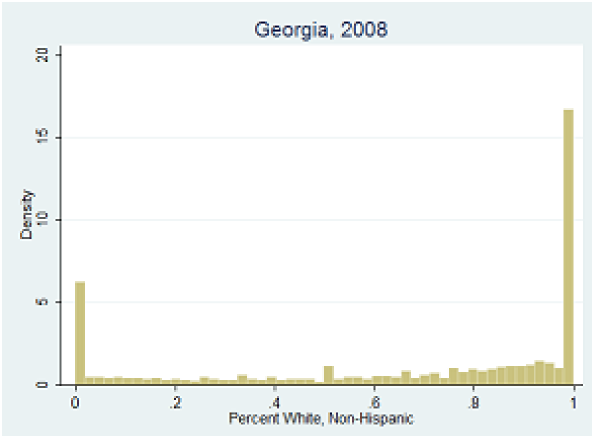
As an example, I analyze the U.S. Presidential election returns in Georgia in 2008. While there is almost certainly some degree of fraud and corruption in U.S. elections, we assume that it is at an insignificant level. For this reason, the U.S. Presidential elections are often used to test forensic techniques for false positives. If we look at the distributions in the 2008 Presidential election in Georgia (Figure 2.7), the data is consistent with a significant amount of ballot box stuffing for candidate Barack Obama.

Figure 2.7. Distribution of relative support for Obama, Georgia 2008.



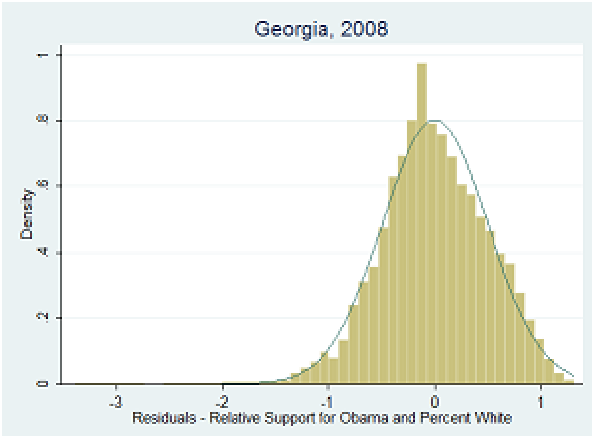
One possible, non-fraudulent explanation for Georgia's non-normality is the idiosyncratic racial politics of the South. In the politics of the U.S. South, race is an important variable, and arguably played a key role in many areas in the 2008 presidential election (see, example.g., Philpot 2009). White voters tend to be more conservative and vote Republican, where African-American voters tend to vote Democrat. Data on racial demographics is available from the U.S. Census Bureau, and is compiled with election results by the Public Mapping Project. In order to measure racial composition, I use the percentage of the census district that is white and nonhispanic. This variable is not distributed normally or symetrically in the data, as evident in Figure 2.8. Instead, there tend to be high degrees of concentration—districts which are almost entirely white or nonwhite.

Figure 2.8. Percent of population that identifies as white.



Given what we know about voter preferences, this may cause the nonsymmetry seen in the distribution of Obama's relative support. If Obama does better in predominantly nonwhite areas than in predominantly white areas, we would not be surprised at the spread in his support or that it has a large concentration at the low end. If we regress his degree of relative support on racial composition, the residuals appear to be normally distributed. The percentage white variable is logged in order to account for nonlinearity in the relationship, which would otherwise influence the distribution of residuals. Figure 2.9 shows the distribution of the residuals from regressing relative support for Obama on race. Controlling for race, Obama's relative support is normal and not consistent with vote inflation.

Figure 2.9. Residuals from regressing turnout on percentage white.

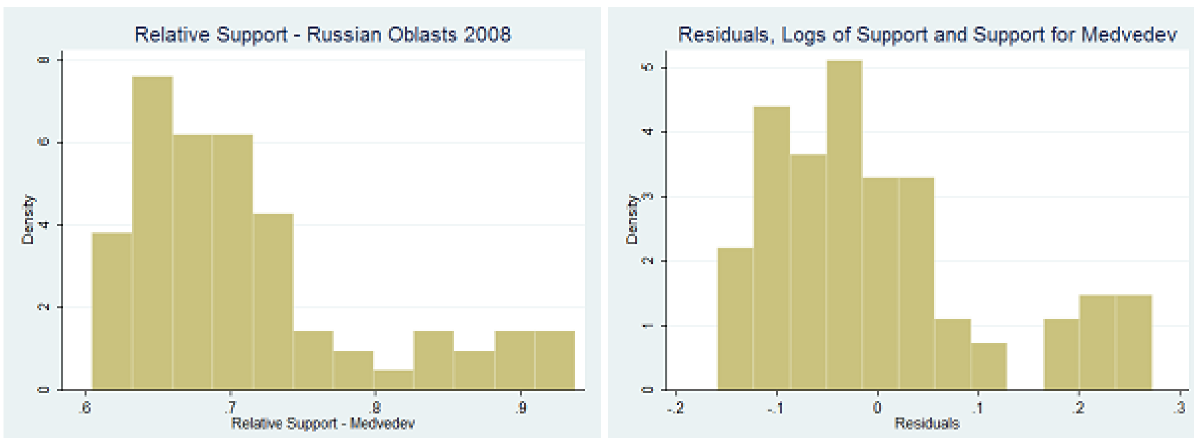


In contrast, the Russian Presidential election of 2008 presents a case where irregularities persist even after controlling for other important variables. This election was likely plagued by fraud. Analysis of turnout and the relative support for Dmitri Medvedev, Vladimir Putin's hand-picked successor, show irregularities. Both turnout and his absolute support have wide right tails, consistent with vote inflation. However, it is possible that this phenomenon is driven by the urban-rural divide. This divide is regarded as one of the most significant cleavages in Russian politics (see, e.g., Berezkin, et al, 1999; Gehlbach, 2000). Traditionally, communist candidates tend to do better in more rural areas. Population density does not have a normal distribution among the 83 federal subjects (akin to states) of Russia. This, rather than vote inflation, could drive the observed skewness in Medvedev's relative support.

The irregularities in the data persist even after controlling for population density. I use population density to measure the degree of urbanization. This data is available at the oblast level from the ArcticStat Socioeconomic Circumpolar Database. When regressing Medvedev's relative support on population density, the relationship is confirmed. Here

Medvedev receives more of the vote in more densely populated regions. However, the distribution of the residuals retains the fat right tail from the distribution of his support. Figure 2.10 shows that this pattern, consistent with vote inflation, persists even when controlling for population density. This outcome suggests that population density is not the driving force in the skewness in the distribution. Instead, the data still appears to be consistent with vote inflation.

Figure 2.10. Relative support for Medvedev and residuals from regressing relative support on population density.



To further evaluate this technique, I run 324 elections with varying parameter values (see Appendix A) and fifty precincts in each. In half of these elections, vote inflation occurs in favor of one candidate. In the other half, there is no fraud but there is natural non-normality. Table 2.1 shows the results of skewness-kurtosis tests for the normality of the residuals obtained by regressing the logarithm of turnout on the logarithm of a control variable in these elections. False negatives occur when the null hypothesis of normality is not rejected at the .1 level despite the presence of vote inflation. False positives occur when the null hypothesis of normality is rejected at the .1

level despite the absence of any fraud. In these simulations, the residual analysis method results in a false negative rate of 35.8% and a false positive rate of 15.4%. Many of these errors can be prevented with visual inspection. These numbers indicate that the residual analysis method can be fairly successful in detecting fraud while controlling for other factors. Furthermore, the test is relatively conservative, which may be desirable from a methodological point of view, placing the burden of proof on those arguing that fraud is present.

Table 2.1. Skewness and kurtosis tests for normality of residuals from simulated elections.

No NNN, Vote Inflation	False Negative Rate: 35.8% (58/162)
NNN, no fraud	False Positive Rate: 15.4% (25/162)

Measuring Fraud

Developments in election forensics have largely focused on detecting anomalies. Efforts to actually measure election fraud have been less common. In this section I use simulation data to evaluate the measures. The simulations vary in the presence of natural non-normality, the type of fraud (vote inflation or vote stealing), the average degree of fraud (the percentage of votes inflated or stolen), the probability that fraud occurs in any given precinct, and the relationship between the likelihood that fraud occurs and support for one candidate or the other. Fraud may be equally likely across all precincts, be more likely for a candidate who has strong support in the precinct, or be more likely for a candidate who has weak support in the precinct. Appendix A presents more detailed information on the parameter values used. Each election in the simulation contains 50 districts, and simulation parameters vary in several ways. Candidate 1's relative support varies between 40%, 50%, and 60%. Additionally there may be natural non-normality caused by a non-normal underlying variable, or all of the underlying variables may be normally distributed. The nature of fraud was allowed to vary a great deal. Two types of fraud were used: vote inflation and vote stealing. The probability that fraud occurred in areas with low support for Candidate 1 varied between 20%, 35%, and 50%. The probability that fraud occurred in areas with high support for Candidate 1 also varied between 20%, 35%, and 50%. The amount of fraud that occurred also varied, with votes being inflated or stolen at rates of 20%, 35%, and 50%. These parameters result in 324 possible combinations. Each combination was simulated 200 times, for a total of 64,800 simulated elections.

I propose a measure based on the skewness value for the distributions of the data. The measure of skewness for a sample captures the magnitude and direction of asymmetry. Because the methods of fraud under study typically result in an asymmetric shift in the data, the degree of asymmetry should capture the degree of election fraud which occurred. For a probability distribution, skewness is defined as: $E[(x-\mu)^3]$.

A statistical test for symmetry or normality is not appropriate for two reasons. First, we do not wish to estimate the likelihood that the data is asymmetric, but the degree to which it is asymmetric. Second, tests for symmetry are highly sensitive to the number of observations. In real world data, with a large number of observations and distribution that is only approximately normal, symmetry is much too easily rejected. Table 2.2 shows the results of skewness-kurtosis tests on turnout in Minnesota in 2008. On visual inspection, the distribution appears fairly normal and symmetric. Normality is rejected based on the skewness-kurtosis test. Only by analyzing a smaller, random sample can the data pass this test.

Table 2.2. Skewness-kurtosis tests for normality of turnout in Minnesota, 2008

Number of observations	Chi Squared Test Statistic	P value
4,119 (all precincts)	.	0.0000
200	40.41	0.0000
100	26.20	0.0000
75	10.52	.0052
50	3.44	.1788

In each case I use the absolute value of skewness values in order to assess the magnitude of the skewness.

Hypothesis 2.1: In clean elections without natural non-normality, skewness = 0 for turnout. In elections with vote stealing, skewness = 0 for turnout. In elections with vote inflation, skewness ≠ 0 for turnout.

Table 2.3 shows the results of the simulations. The distributions for turnout are not skewed when there is neither natural non-normality nor vote inflation. Vote inflation results in skewness in the distributions. The skewness increases as fraud increases. This indicates that ballot box stuffing will create a shift in the distribution, as argued by Mygakov, et al. (2009).

Table 2.3. Skewness measures for turnout in the absence of natural non-normality

Type of Fraud	Mean Skewness, Turnout	Standard Deviation
None	.001	0.340
Vote Stealing	.001	0.342
Vote Inflation, 20%	.251	.380
Vote Inflation, 35%	.458	.378
Vote Inflation, 50%	.624	.360

Hypothesis 2.2: In clean elections with natural-non normality, skewness ≠ 0. Skewness = 0 for the residuals of the regression of turnout on non-normal explanatory variables.

In the elections with natural non-normality, skewness for turnout has a mean of . 308 with a standard deviation of .07. Regressing turnout on the control variable results in residuals which have skewness of -.509 on average. This situation suggests that skewness is less useful when analyzing residuals, since it may be too sensitive. In these cases it is essential to visually inspect the resulting distributions, and future work should seek to derive better measures.

Hypothesis 2.3: When vote stealing occurs, the skewness of absolute support will increase for both candidates. When vote inflation occurs, skewness of absolute support will increase for the beneficiary of fraud and not change for the victim of fraud.

Skewness of absolute support appears to be a good predictor of fraud. A *t* test for the difference of means between the prefraud skewness and postfraud skewness for candidate 1 rejects the null hypothesis with a *t* value of -42.058. Skewness increases in the presence of fraud, confirming Hypothesis 2.3. For candidate 2, a *t* test for difference of means suggests that skewness is a valid measure for vote stealing. The test rejects the null hypothesis with *t* value of -37.904. Table 2.4 provides summary statistics for the skewness of absolute support for each candidate in the absence of natural non-normality. The pattern is less extreme in these cases than it is for turnout when vote inflation is the method of fraud.

Absolute support sometimes has a high value for candidate 1 even in the absence of fraud. Since absolute support tends to be fairly low, this can be caused by the truncation of an otherwise normal distribution. Visual inspection remains a useful tool for evaluating these cases.

Table 2.4. Skewness measures for absolute support

Type of fraud	Candidate	Mean Skewness, Absolute Support	Standard Deviation
None	Candidate 1	.221	.301
Vote inflation or Vote stealing	Candidate 1	.307	.420
None	Candidate 2	.093	.294
Vote Inflation	Candidate 2	.090	.294
Vote Stealing	Candidate 2	.245	.314

Hypothesis 2.4: In an election with natural non-normality, the residuals obtained by regressing absolute support on non-normal explanatory variables will have higher skewness in the presence of fraud than in the absence of fraud.

When analyzing the residuals of this regression, the data from the simulations do not appear to confirm Hypothesis 2.4. In the set of 64,800 simulations with 50 districts each, the mean of skewness for the residuals in the absence of fraud is -.740 and the mean of skewness for the residuals in the presence of fraud is -.631. When increasing the number of districts, the data do support Hypothesis 2.4. This suggests that the skewness measure's sensitivity to outliers may make it more appropriate when there are a large number of districts. In simulations with 1,000 districts in each election, the residuals obtained by regressing absolute support on the log of the control variable have an increase in skewness following election fraud. In the absence of fraud, the mean skewness is .115 with a standard deviation of .067. Following election fraud of either type, the mean skewness is .227 with a standard deviation of .207. The t test for difference of means rejects the null hypothesis with a t value of -6.861.

On review, it appears that the measure of skewness has the potential to serve as a measure for fraud. Future research should continue to test and refine this measure. In later chapters I will use this measure to take a deeper look at election fraud and attempt to uncover factors that influence the use of various forms of fraud.

Conclusion

In this chapter I have tackled two serious obstacles in the field of election forensics. By using regression based techniques, I developed a tool for analyzing electoral returns for irregularities while controlling for important factors such as wealth, racial diversity, or other demographic issues. Then, based on a comprehensive set of forensic tools, I have suggested a measure for election fraud.

In subsequent chapters of this dissertation, I put these tools to use analyzing elections from three different regions. From a methodological perspective, this provides an opportunity to test these tools on cases where we have established literature and first-hand accounts as to the degree and types of fraud employed. The tools developed here can describe the nature of election fraud allowing for a level of analysis for these elections than has been undertaken before. From a comparative perspective, understanding the nature of election fraud in these three cases can inform theories about election fraud and democratic transitions more broadly.

CHAPTER III

FORMAL THEORIES OF ELECTION FRAUD

In this chapter I develop a general model of whether or not a “machine party” (a single party capable of controlling and manipulating the election process) holds elections and how much fraud such a party decides to engage in. This model helps guide election forensics in ways that previous models do not, which will be discussed more fully in the conclusion. In the context of election forensics, this type of modeling exercise is novel. Previous models of fraud (e.g., Chaves et al. 2009; Fearon 2011; Magaloni 2006) were not motivated by forensics.

Next, I discuss the generalizability of the model found in Chaves et al. (2009). This model seeks to predict fraud at a more local level. Rather than focusing on top-down fraud, this model focuses on bottom-up fraud. In other words, local election officials have incentives to report vote counts favorable for the machine party and are subject to local constraints.

Previous Models of Election Fraud

Several authors have put forward rational choice based models regarding election fraud. The most broad of these look at the decisions made by ruling regimes regarding fraud. Magaloni (2006) uses a sequential game to explain why authoritarian regimes may hold elections in the first place and why they may create and abide by election reform laws aimed at curbing fraud. Magaloni argues that regimes will hold elections to create illusions of legitimacy, but will engage in fraud when they know that they will be accused of it regardless of the amount of fraud that actually takes place. In contrast, they may actually "tie their hands" by agreeing to antifraud reforms when they believe they can win elections legitimately and when the opposition is united enough to lobby for those reforms. Fearon (2011) creates a similarly macro-level model of the behavior of authoritarian regimes. Fraud is not given such detailed treatment in Fearon's model, however. Instead, he assumes that regimes will engage in as much fraud as they believe they can get away with. Punishment, either from international pressure or domestic opposition, is the sole constraint.

Other models focus more on decisions regarding the nature of fraud itself. Kalinin and Mebane (2011) look at the effect of internal politics on election fraud in Russia. They find that local elites had incentives that changed over time. During the decentralization of the 1990s, local elites had less accountability to the established party and therefore did not have to deliver in terms of fraud. When this pattern shifted, the threat of punishment from the central government provided an incentive for local elites to engage in fraud. Chaves et al. (2009) analyze the 1922 presidential election in Colombia. They develop predictions based on their model of the decisions made by local elites. In this model, officials attempt to maximize the margin of victory for the conservative party. They are constrained by the cost of fraud in terms of public backlash, but this cost is mitigated by the presence of proconservative clergy. Since the presence of clergy is correlated with higher legitimate vote counts for the conservative party, they predict a nonlinear relationship between the number of clergy in a region and the amount of fraud that takes place. More priests make fraud less costly, but also reduce the amount of fraud that can occur since more of the ballots were cast for conservatives in the first place.

While Kalinin and Mebane (2011) and Chaves et al. (2009) examine specific nations, other authors have developed generalizable models to explain the choices made regarding the type of fraud employed. Nichter (2008) and Gans-Morse et al. (2009) discuss the variety of tools available to those who wish to manipulate elections. Given different characteristics of potential voters, different methods may be more appropriate. For example, if a particular election features a high number of opposition voters who are disinclined to vote, it may be cost-effective to pay them to ensure that they do not vote. If, on the other hand, voting is not very costly, it may be better to pay them more to change their vote.

In this chapter I rely on the models developed by Magaloni (2006) and Chaves et al. (2009). Drawing from these models allows me to provide pictures of fraud from two different directions: top-down and bottom-up. Table 3.1 presents the models of election fraud discussed above.

Table 3.1. Models of election fraud

Model	Type	Players	Key Assumptions/Features	Key Parameters	Findings
Chaves et al.	Choice Theoretic	Regional Election Authority	-Vote inflation is achieved by ballot box stuffing -Vote destruction is achieved through coercion -Regional authorities seek to maximize the margin of victory for the conservative party, subject to costs of fraud	p – priests s – ballot stuffing c – coercion	-Model to predicts the effect of p on s and c. Essentially, p affects the costs of each strategy -Measuring s and c, the test empirically for the effect of p -Explains the effect of regional variables on the degree or likelihood of fraud
Gans-Morse, Mazzuca, Nichter	Game Theoretic	Party Machine, Voters	-Party knows voters preferences and propensity to vote -Election fraud happens through direct interaction with voters	x – level of support for party machine c – cost of voting	-Parties will use a variety of methods to alter vote counts -The rate at which methods are employed depends on voter preferences and institutional factors -They compare with empirical evidence from Brazil, Russia, and Argentina -Explains how machines allocate resources to alter voter behavior
Mebane	Game Theoretic	National party, local authorities	-During de-centralization, bargaining between regional authorities and national party takes place before elections -During centralization, national party may reward or punish regional governors postelection	d – value of punishment to gov. λ – prob. a gov is loyal b – value of rewards to governors for fraud	-Looks at Russian elections and explains the difference in fraud in the 90's versus 00's as a change in the structure of the game, as punishment became more viable later -Explains the decision of local authorities to commit fraud in the face of rewards/punishment from national party
Magaloni	Game Theoretic	Party Machine, Opposition parties, voters	-Divided opposition -Imperfect information, opposition does not know if there was or was not fraud	Many	-Established parties are likely to “tie their hands” when they think they can win anyway, when opposition is united -They will engage in fraud when they believe they will face some backlash either way, and when the military is on their side -*Explains* why authoritarian regimes decisions regarding fraud and institutional “hand-tying”

Micro-level Model

Chaves et al. (2009) develop a model of the behavior of local election officials incentivized to deliver wide margins of victory to the machine party. In an attempt to curry favor with the regime, local authorities seek to maximize the returns for the machine party. At the national level, the goal may be to maximize the chances of reaching a plurality. When this plurality is likely without fraud, the regime may choose clean elections. In contrast, local officials are not necessarily concerned with a simple plurality. Most elections do not feature winner-take-all systems at the regional level. Officials at the precinct and county levels have an incentive to make the margin of victory as large as possible in their sphere of influence in order to maximize the machine party's chances at success. They are constrained by two factors: the number of eligible voters and the cost associated with engaging in election fraud. As before, these costs come in the form of civil unrest or reporting by election monitors.

The objective function of the regional election authorities is then expressed as follows:

(1)

$$\max_{s,c} \rho(v^c + s - v^L) - \frac{\varphi}{2}s^2 - \frac{\psi}{2p}c^2$$

Where v^c is the number of votes for the machine (conservative) party, v^L is the number of votes for the challenging (liberal) party. The choice variables are s and c: s is the amount of ballot box stuffing in votes inflated for the machine party, c is the amount of vote suppression that takes place, and p is a regional characteristic associated with support for the machine party. In the Chaves example, this regional characteristic is the number of priests in a given region. Regions with more priests tend to be more conservative. The cost associated with vote suppression is lower in these regions, as the opposition has less support and power. ρ is a parameter denoting the value of the machine party's margin of victory, φ is a parameter that captures the cost of vote inflation, and ψ is a parameter that captures the cost of vote suppression.

The constraint that the local authorities face is:

(2) $v^c + v^L + s - d \leq v_{max}$

The reported number of votes must be less than or equal to the number of voters (v_{max}). The Chaves example treats the maximum number of votes as one third of the total population. Reported vote counts higher than this would be treated with suspicion, and come with the risk of having the region's votes disqualified. Furthermore, the number of votes the machine party receives is a function of the regional characteristic p, and the number of votes the opposition party receives is a function of the amount of vote suppression which takes place:

(3) $v^c = f(p)$

where $f'(p) > 0$.

(4) $v^L = g(c)$

where $g' < 0$ and $g'' > 0$.

As Chaves et al. (2009) note, the inequality in equation (2) may not be binding. They argue that, empirically, it usually does bind. When officials engaged in fraud, they did so to a high degree. Often officials reported vote counts that were right on the edge of believability. The value of v^{max} may vary from one case to another. As discussed in Chapter 4, reported turnout values of 100% are fairly common in Russia. Chaves et al. (2009) assume that the inequality binds. So we can write:

(5) $s = v^{max} - h(p) - g(c)$

and substitute this into the objective function. This leads to the first order conditions:

$$(6) \quad -2\rho g'(c) - \frac{\psi}{p} + \phi(v^{max} - h(p) - g(c))g'(c) = 0$$

Chaves, et al. then derive comparative statics for vote inflation and vote suppression. Of key interest here is the relationship between the two types of fraud and the regional characteristic p:

$$(7) \quad \frac{\delta c}{\delta p} = \frac{\frac{\psi}{p^2}c - \phi g'(c)h'(p)}{-\omega} > 0$$

$$(8) \quad \frac{\delta s}{\delta p} = \underbrace{-h'(p)}_{negative} + \underbrace{-g'(c)\frac{\delta c}{\delta p}}_{positive}$$

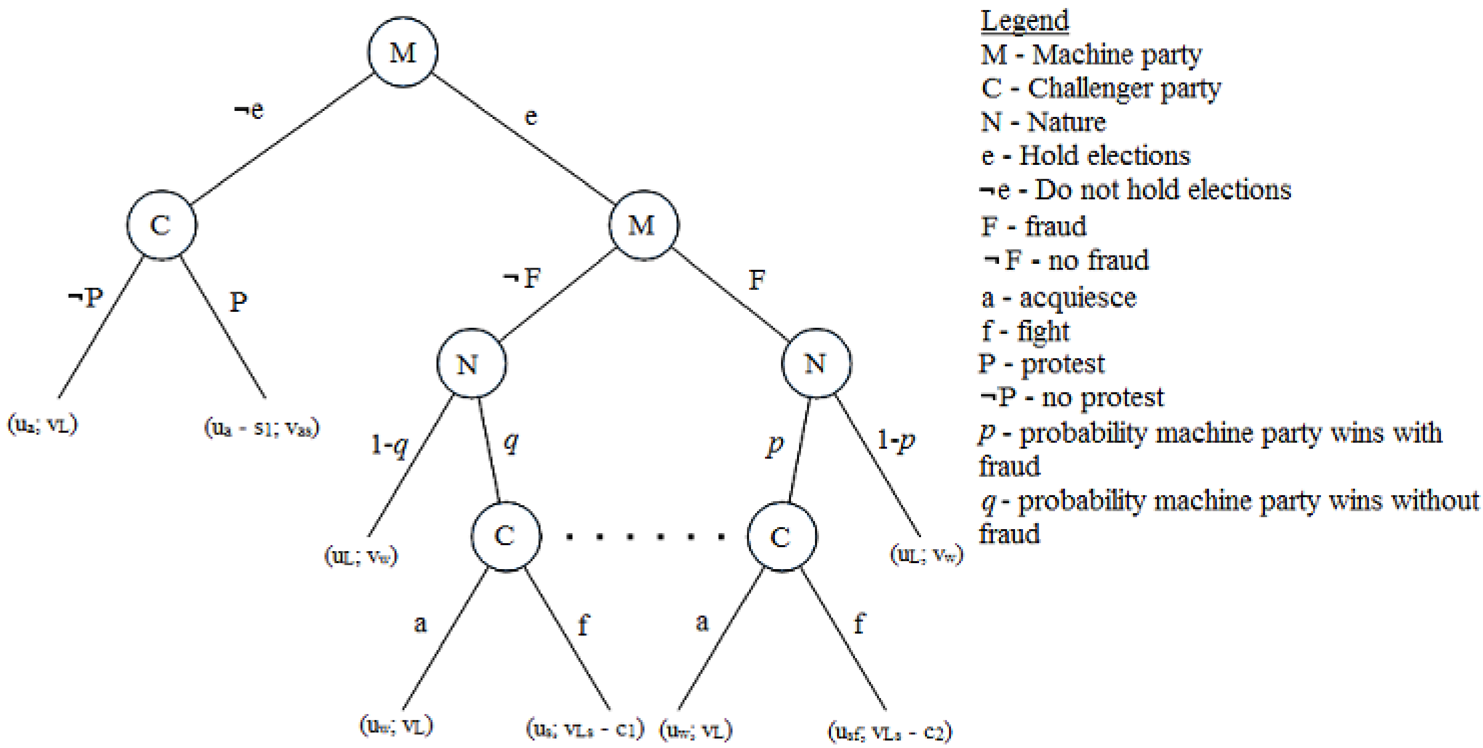
So the amount of vote suppression depends on the regional characteristic. As this characteristic increases, so does the ability of election officials to engage in vote destruction without fear of reprisal. Therefore, we should expect to see more fraud in regions where the machine party enjoys high support. The result for vote inflation implies that the relationship between the amount of vote inflation and the regional characteristic p is ambiguous. At high levels of p, the support for the machine party is high enough to eliminate the need for ballot box stuffing. Turnout cannot be increased any further.

Macro-level Model

An extensive "game tree" representation of my model is found in Figure 3.1. In the first stage of the game—period 1.1—the "machine party" decides whether to hold clean elections, fraudulent elections, or no elections. For simplicity, I assume a first-past-the-post election determines control of government. (We will consider other institutional assumptions below). Note that this is a very majoritiarian assumption. In this system, the machine party seeks to maximize its chances of winning by obtaining a plurality, and minimizing the negative costs of accusations of fraud.

If the machine party does not hold elections, the opposition may stage protests. Whether the party wins or loses the election is based on chance, with the probability of a "win" or loss varying between clean and fraudulent elections. A machine party loss ends the game. If the machine party wins, the opposition chooses whether to fight the results or acquiesce.

Figure 3.1. Game tree for the macro-level model.



Theoretical results

I make the following assumptions about the parameters:

$u_W > u_L$: the machine party would rather win than lose the election.

$p > q$: When the machine party engages in fraud, it increases its likelihood of winning the election.

Hypothesis 1: Regimes will hold elections and engage in fraud when:

1.1) $p(u_{sf}) + (1\text{-}p)(u_L) > u_a$, $p(u_{sf}) + (1\text{-}p)(u_L) > q(u_s) + (1\text{-}q)(u_L)$, *and* $v_{Ls} - c_2 > v_L > v_{as}$ (case where the opposition will fight fraud but not protest);

1.2) $p(u_{sf}) + (1\text{-}p)(u_L) > u_a - s_1$, $p(u_{sf}) + (1\text{-}p)(u_L) > q(u_s) + (1\text{-}q)(u_L)$, *and* $v_{Ls} - c_2 > v_{as} > v_L$ (case where the opposition will fight fraud and protest);

1.3) $p(u_w) + (1 - p)(u_L) > u_a$, $v_L > v_{as}$ *and* $v_L > v_{Ls} - c_2$ (case where the opposition will acquiesce to both);

or

1.4) $p(u_w) + (1 - p)(u_L) > u_a - s_1$, *and* $v_{as} > v_L > v_{Ls} - c_2$ (case where the opposition will protest but not fight fraud).

In 1.1 and 1.2 (these are labeled in Fig. 3.1), "f" dominates "a" for the opposition. In these cases, the machine party will engage in fraud if the increased chance of winning offsets the additional cost of sanctions when there is fraud (the difference between u_s and u_{sf}). In 1.3 and 1.4, the opposition will not fight fraud as "a" dominates "f." As a result, the machine party will engage in fraud since $u_w > u_L$ and fraud increases its chances of winning.

Hypothesis 2: Regimes will hold elections and not engage in fraud when:

2.1) $p(u_{sf}) + (1\text{-}p)(u_L) > u_a$, $q(u_s) + (1\text{-}q)(u_L) > p(u_{sf}) + (1\text{-}p)(u_L)$, *and*

$v_{Ls} - c_2 > v_L > v_{as}$ (opposition fights and does not protest);

2.2) $p(u_{sf}) + (1\text{-}p)(u_L) > u_a - s_1$, $q(u_s) + (1\text{-}q)(u_L) > p(u_{sf}) + (1\text{-}p)(u_L)$, *and*

$v_{Ls} - c_2 > v_{as} > v_L$ (opposition fights and protests);

2.3) $q(u_w) + (1\text{-}q)(u_L) > p(u_w) + (1\text{-}p)(u_L)$, $q(u_w) + (1\text{-}q)(u_L) > u_a - s_1$, *and*

$v_{as} > v_L > v_{Ls} - c_2$ (opposition protests and does not fight);

2.4) $q(u_w) + (1\text{-}q)(u_L) > p(u_w) + (1\text{-}p)(u_L)$, $q(u_w) + (1\text{-}q)(u_L) > u_a$, $v_L > v_{as}$ *and*

$v_L > v_{Ls} - c_2$ (opposition neither fights nor protests).

2.3 and 2.4 are in violation of the assumption that $u_W > u_L$. Since the opposition will not fight fraud, the machine party will always engage in fraud in equilibrium. In 1.1 and 1.2, the opposition fights fraud. In these cases, the machine party will not engage in fraud when the sanctions are worse ($u_s > u_{sf}$) and this difference is not offset by an increase in the probability that the machine party will win the election. Essentially, regimes will not engage in fraud when the sanctions associated with fraud are high and the utility of ruling as a dictator is a lower than expected utility of running a legitimate election.

Hypothesis 3: Regimes will hold elections and may or may not engage in fraud when:

3.1) $v_{Ls} - c_2 > v_L > v_{Ls} - c_1$

In this case, the opposition would prefer to fight when there is fraud, and acquiesce when there is not. Because the opposition is uncertain as to the machine party's move, it may be willing to play mixed strategies in equilibrium if the machine party is also playing mixed strategies.

The machine party will play F with probability m and -F with probability 1-m. The challenger will play "a" with probability n and f with probability 1-n. In equilibrium:

3.2) $m = (-qv_L + qv_{LS} - qc_1) / (pv_L - qv_L - pv_{LS} + pc_2 + qv_{Ls} - qc_1);$

3.3) $n = (-pu_{sf} + pu_L + qu_s - qu_L) / (pu_w - pu_{sf} - qu_w + qu_s).$

Proposition 1: The more competitive the election, the more likely the machine party is to engage in fraud.

Proof:

From 3.2:

$$\frac{\delta m}{\delta p} = \frac{q(v_L - (v_{LS} - c_1))(v_L - (v_{CS} - c_2))}{(pv_L - qv_l - pv_{LS} + pc_2 + qv_{LS} - qc_1)^2}$$

By inspection, the numerator is negative since q is positive, $v_L - (v_{LS} - c_1)$ is positive as in 3.1, and $v_L - (v_{CS} - c_2)$ is negative as in 3.1. When the machine party adopts mixed strategies in response to fraud, it is less likely to engage in fraud, as the likelihood that it will win without fraud increases.

This result is notably different than the results obtained from the micro-level model. In the micro-level model, fraud aims to increase that margin of victory, not increase the odds of success. Therefore, support for the machine party means more vote suppression and potentially more fraud. From the perspective of national parties concerned with pluralities or even a given margin of victory, more support for the machine party will generally mean less fraud.

Conclusion

The models developed in this chapter expand on those developed by Magaloni

(2006) and by Chaves et al. (2009). Macro-level models capture the decisions of national level machine parties. Micro-level models capture the decisions of local elites.

Within these two systems we also have empirical predictions that can be tested. At the micro-level, I predict that fraud will have a non-linear relationship with factors that are correlated with increased support for the perpetrator of fraud. At low levels of support, it may be difficult for elites to get away with fraud due to the potential for unrest. At high levels of support, it is not possible to manipulate the vote to an extreme degree since the returns are already so favorable. Fraud is more likely to occur in regions with mixed levels of support. At the macro-level, I predict that fraud will be more common in more competitive elections. Electoral authoritarian regimes are typically concerned with balancing a healthy margin of victory with the appearance of legitimacy. To this extent, they will attempt to engage in only as much fraud as necessary to obtain their electoral goals.

Individual cases are better represented by one model or the other. For example, election fraud in Mexico appears to be a top-down phenomenon. The 1988 election is infamous for the fraud committed at the national level. When it appeared that the PRI candidate was poised to lose the election, the computers at the central election agency "crashed," and the establishment won the election once the computers "came back online." Mexico fits the predictions found in our macro-level model: Fraud was severe in the competitive 1988 election but was reduced in 1994 when the PRI candidate enjoyed much more popularity. In contrast, Russia may be an example of more bottom-up fraud. Local electoral officials have had an incentive to provide election returns favorable to

United Russia and Vladimir Putin. In the most recent presidential election, Putin appeared to take steps to reduce fraud by installing webcams in virtually every polling place. Due to his popularity, Putin would win elections in Russia even without fraud. While large victories may give him more political clout, given international concerns about legitimacy may urge him to genuinely want fraud decreased. Therefore, the different predictions generated by these two models may enable us to determine the driving forces behind fraud in particular elections. When fraud is more prevalent in competitive elections, efforts at manipulating the vote may come from centralized and well organized national parties. When fraud is determined more by local characteristics than by competitiveness, local elites may be the driving force behind electoral manipulation.

CHAPTER IV

ELECTION FRAUD IN RUSSIA

Historically elections in Russia have been notorious for fraud. The traditions of ballot box stuffing and voter intimidation that defined Soviet era elections have segued into the “democratization” of the early 1990s. In the modern era, Russia has failed to transform into a truly competitive democracy. Elections in Russia have seldom been truly competitive, for a number of reasons. Given the nation’s limitations on speech and demonstration, political debates rarely feature the establishment candidate and are often aired at low-audience times of day. Additionally, some political parties exist only to create the illusion of competition and to take votes from other opposition parties (Wilson 2005). Ultimately, a high degree of electoral fraud prevails.

Elections are plagued by reports of ballot box stuffing, miscounting of ballots, and destruction of votes for the opposition (e.g., Bader 2012). In this chapter, I look at elections to the national Duma and Presidential elections from 2007 through 2012. Outright election fraud was less prevalent during the 1990s and early 2000s (Myagkov et al. 2009; Wilson 2005, p. xv;). Instead of increasing legitimacy and democratization following the collapse of the Soviet Union, conditions point to a steady decline in legitimacy and increase in election fraud (Bader 2012). Kalinin and Mebane (2011) find patterns of growing fraud throughout this period. Beginning in 2007, it appears that election fraud became a problem not only in the ethnic republics and rural areas, but across the nation. Public outrage at this level of fraud increased dramatically following the 2011 elections. This outrage prompted the administration to investigate claims and begin to incorporate some reforms, but such efforts appear to have had little to no effect on the degree of election fraud taking place. The Central Election Commission remains a partisan entity and local elites still have incentives to falsify vote counts. As a result, no significant improvements have been made. Russia does not appear any closer to a democracy following the 2012 election than it was in 2007.

Recent elections

In this section I analyze the Duma elections of 2007 and 2011 and the Presidential elections of 2008 and 2012. I find that fraud may have decreased in the presidential elections of 2012 but is still a persistent and pervasive problem. Election data are available from the Central Election Commission (www.cikrf.ru). Population density data are available from ArcticStat (www.arcticstat.org).

2007 Legislative Election

The State Duma is the lower house in Russia's legislative body. The 2007 Duma election took place during Vladimir Putin's second term in office. Given Putin's popularity and success in the 2004 Presidential election, it was anticipated that his United Russia party would do well. United Russia ended up winning nearly two thirds of the seats in the legislature, giving the Putin administration an almost complete concentration of power (Myagkov et al. 2009, p. 119; Kryshtanovskaya and White, 2009).

In addition to stifling opposition during the campaign, outright election fraud appears to have played a significant role in the election. Russia is divided into 83 federal subjects, including oblasts (administrative subdivisions, or regions, of a republic), republics (which are ethnic and nominally autonomous), “krais” (territories, functionally the same as oblasts), and two federal cities (Moscow and Saint Petersburg). In earlier elections, fraud was more likely limited to ethnic republics and a few other oblasts (Myagkov et al. 2009, p. 122). Beginning with the 2007 Duma election, however, it appears that a regime of electoral fraud became thoroughly entrenched.

Evidence exists of extreme levels of ballot box stuffing across all oblasts (see Figure 4.1). Not only does the distribution of turnout have a wide right-hand tail, but reveals a large spike at 100% turnout. A more detailed picture can be obtained by looking at individual oblasts and republics. Chechnya is undoubtedly one of the most egregious examples of irregular data. Turnout is extremely high—close to 100% across precincts. The slope coefficient for the regression of absolute support for United Russia on turnout is 1.13, while the relative support for United Russia (the expected value for the coefficient) is 99.4%. This means that, although United Russia received almost all of the votes in Chechnya, that party did better in those precincts where turnout was reported closer to 100% (see Figure 4.2).

Figure 4.1. Turnout across oblasts in 2007.

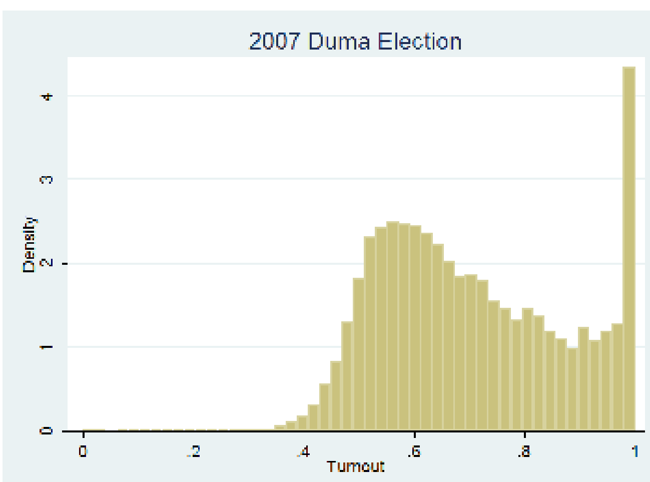
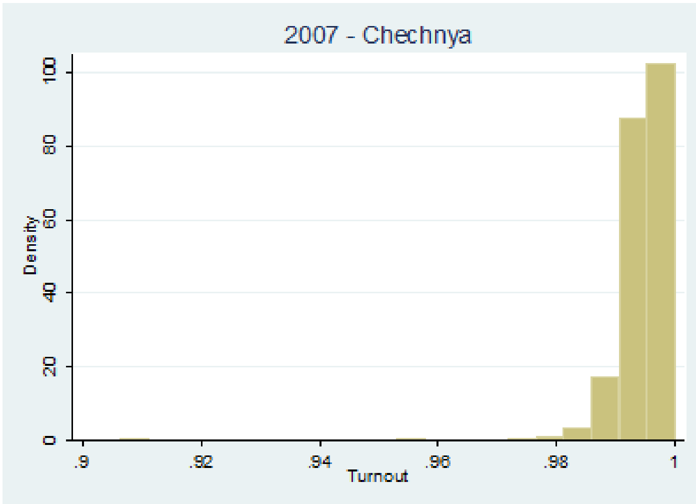


Figure 4.2. Turnout in Chechnya in 2007.



As Mygakov et al. (2009) note, the "usual suspects" of Bashkortostan, Dagestan, and Tatarstan all look suspicious, as well. However, not all of the irregularities are found in ethnic republics. As shown in Figure 4.3, Tyumen Oblast, one of the wealthiest regions in Russia, also displays patterns in the data consistent with ballot box stuffing in favor of United Russia (UR). Using the "jaws" method, Figure 4.4 shows that the irregularities in absolute support and turnout do not appear to be natural. The residuals have a mean that is greater than zero at greater than the .01 level of statistical significance, and the distribution of residuals shows that rather than a normal distribution, the distribution is consistent with ballot box stuffing.

Figure 4.3. Turnout and absolute support in Tyumen, 2007.

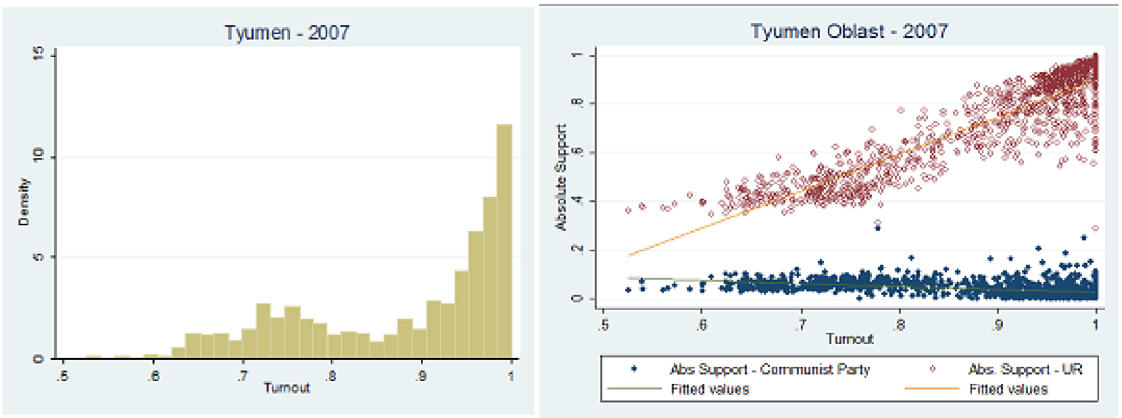
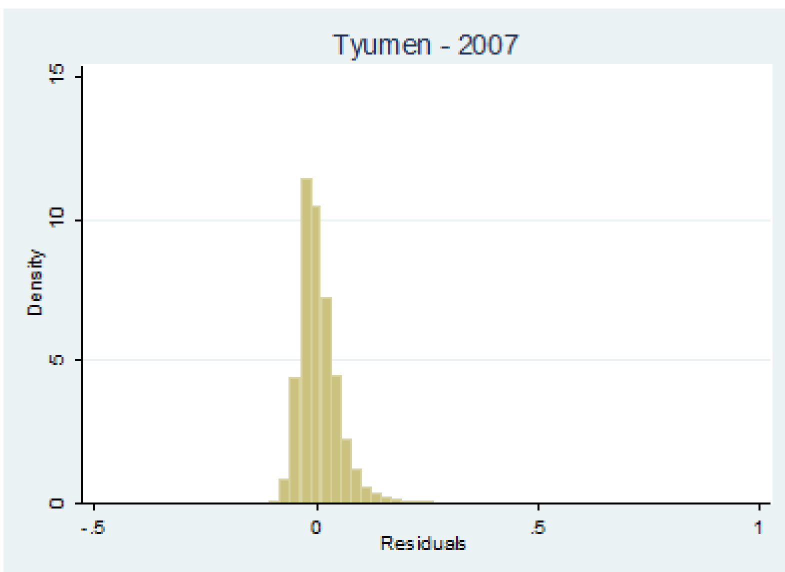
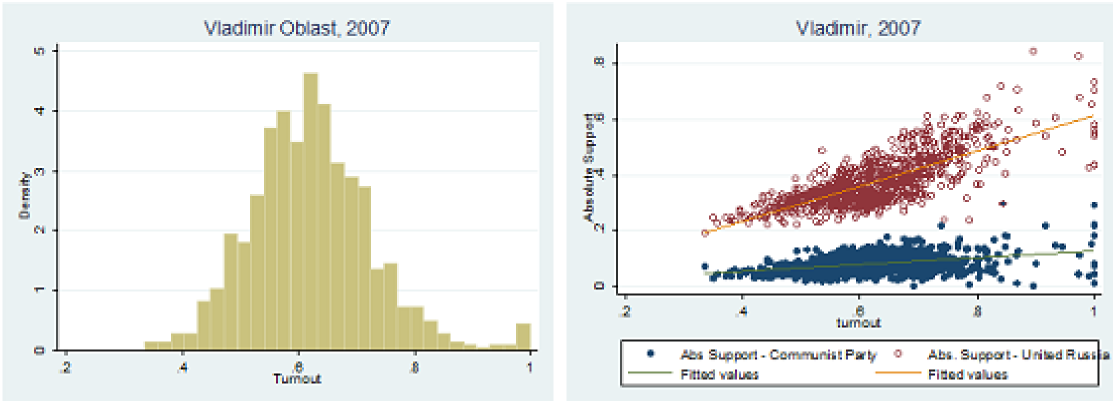


Figure 4.4. Residuals from regressing absolute support for UR on turnout.



Certainly not every federal subject contains irregularities. Vladimir Oblast, for example (see Figure 4.5), appears to have a relatively normal distribution of turnout. The SSD (sum of squared differences) for each party is close to zero.

Figure 4.5. Turnout and absolute support in Vladimir, 2007.

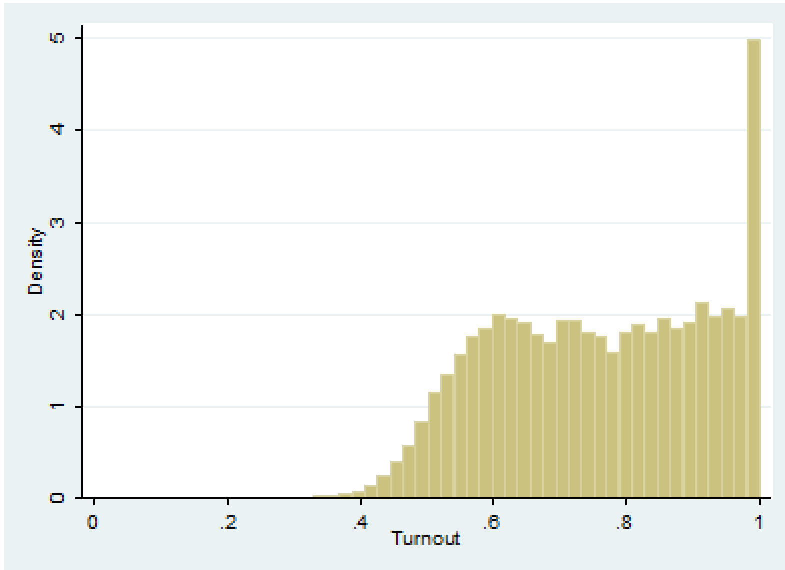


2008 Presidential Election

Term limits prevented Putin from running for president again in 2008. Instead, the UR ticket featured his hand-picked successor Dmitry Medvedev. Criticisms regarding the uneven playing field seen in the 2007 Duma election resurfaced. One example includes the UR candidate refusing to debate the opposition, and debates being aired at odd hours.

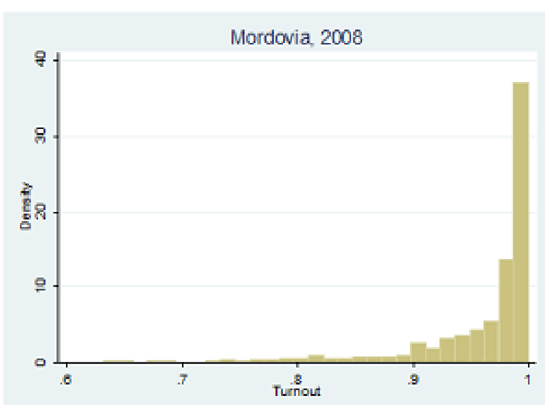
Medvedev won the election handily. The irregularities in the data from the 2008 Presidential election are very similar to those seen in the 2007 Duma election. While fraud may not have altered the outcome of the election, it was certainly pervasive. Figure 4.6 shows that the distribution of turnout at the national level is consistent with an extreme degree of ballot box stuffing. My findings here are consistent with those found by Lukinova et al. (2011), namely that “the level of fraud ... is incontrovertible and that the concept of 'Russian Democracy' remains an oxymoron” (p. 620). The authors find that fraud is more severe than in 2004, and that fraud continues to spread.

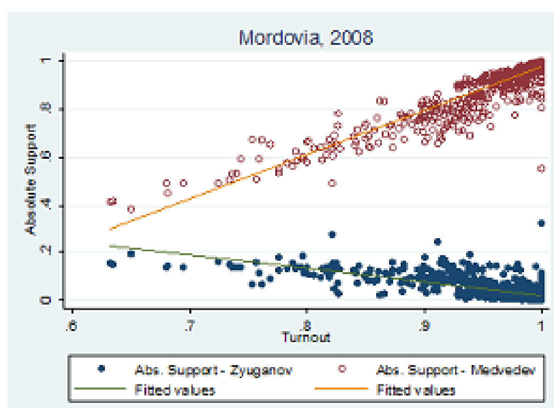
Figure 4.6. Turnout in Russia's 2008 presidential election. All oblasts.



Looking at individual oblasts, we see patterns consistent with a high degree of fraud in favor of the UR candidate Dmitry Medvedev. One of the most egregious examples is found in the ethnic republic of Mordovia (see Figure 4.7).

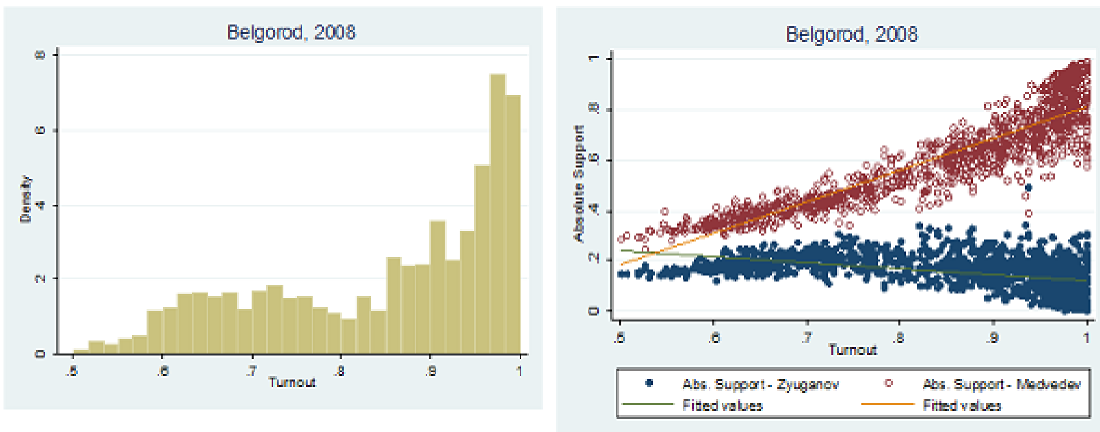
Figure 4.7. Turnout and absolute support in Mordovia.





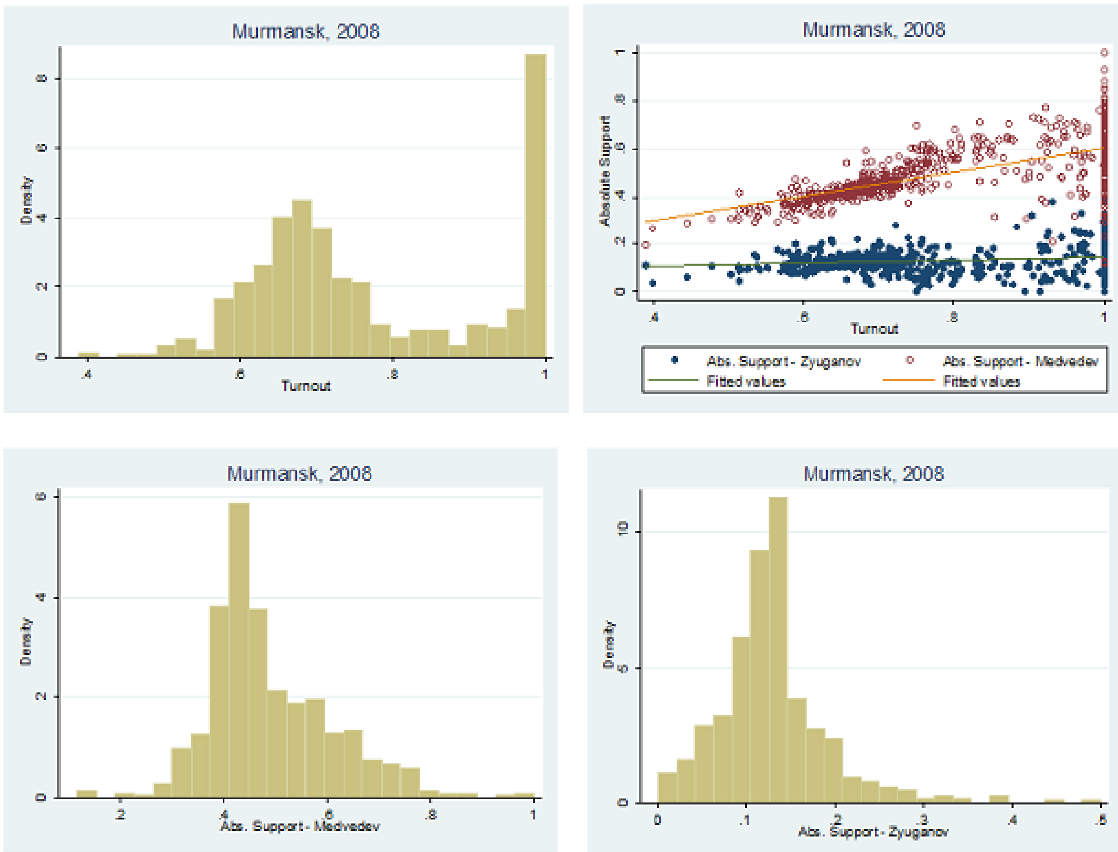
Similar irregularities are found in many of the ethnic republics, but also in other regions. Tyumen again appears particularly egregious with masses of districts reporting near 100% turnout, and near 100% support for Medvedev. Belgorod (see Figure 4.8) is another example of an oblast with very suspicious irregularities.

Figure 4.8. Turnout and absolute support in Belgorod.



Yet not all regions appear so irregular. Murmansk is a particularly interesting case (see Figure 4.9). Despite the fact that the distribution of turnout appears consistent with ballot box stuffing, the SSD for each party is close to zero. Furthermore, the absolute support for each candidate appears normally distributed, suggesting that significant ballot box stuffing did not occur. A more reasonable explanation for the high number of precincts reporting 100% turnout may be the high number of precincts with only a handful of registered voters, where achieving complete turnout is more likely.

Figure 4.9. Turnout and absolute support in Murmansk.

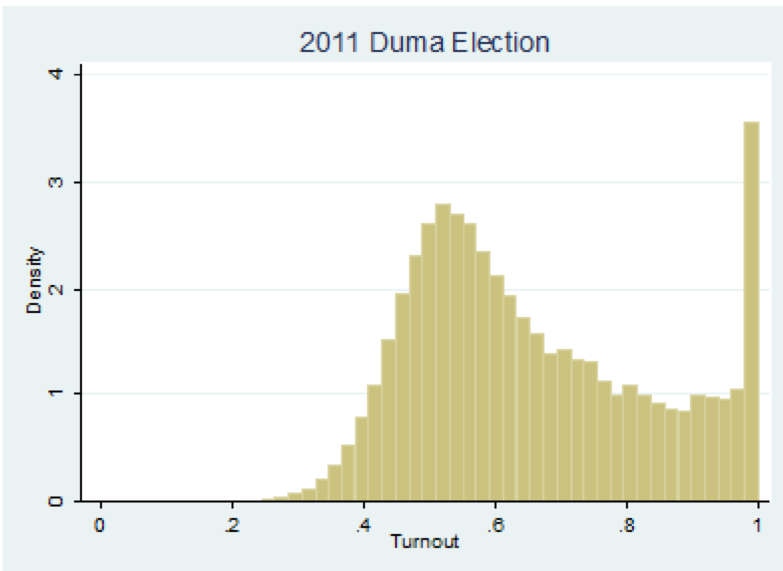


2011 Legislative Election

Despite the fact that United Russia lost its two-thirds majority in the Duma in the 2011 elections, evidence of election fraud was abundant. Perhaps most importantly, several videos featuring alleged fraud were displayed online. These amateur videos showed individuals loading multiple ballots into ballot boxes. The appearance of now familiar issues of fraud in the 2011 Duma election sparked a widespread protest movement. The UR's declining popularity and further evidence of fraud may have contributed to the initiation of these rallies.

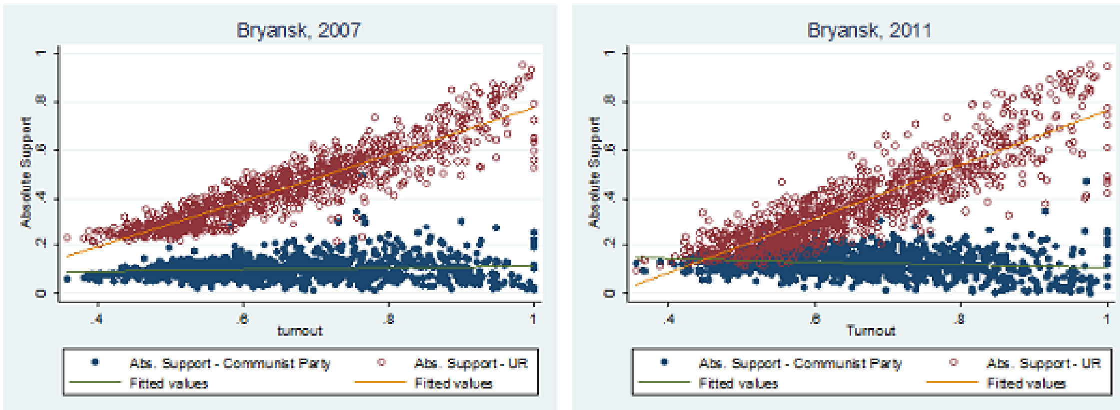
The Medvedev administration was quick to downplay the impact of fraud, but called for investigation and, ultimately, for reform. Although establishment officials argued that fraud had little or no impact on the elections, the data shows irregularities similar to those seen in the previous Duma elections (see Figure 4.10).

Figure 4.10. Turnout across all regions.



In terms of the SSD values, the top five regions are among the usual suspects: Mordovia, Tyumen, Mari El, Belgorod, and Tamal-Nenets. The patterns across these two elections are strikingly similar. The correlation coefficient for the SSD for United Russia is .683. While many of the ethnic republics have irregularities in their data, the patterns are not confined to ethnic republics or to more rural oblasts. In some oblasts, such as Bryansk, the 2011 election appears to be slightly more fraudulent than the Duma election in 2007 (see Figure 4.11). The mean value of the SSD for United Russia in 2007 was . 495; in 2011 it was .596. A *t* test for the difference of means indicates that they are not different at the .1 level of significance, suggesting that increases in fraud, if any, were slight. This outcome also suggests that the lower vote total for the UR party was not due to any significant decrease in fraud, but due to legitimately lower levels of support.

Figure 4.11. Bryansk, 2007 and 2011.

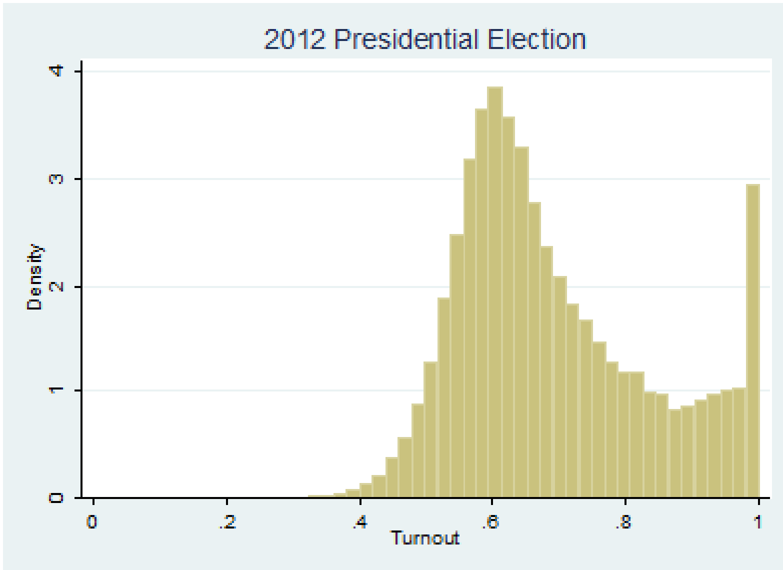


2012 Presidential Election

Following the public outcry in response to fraud in the 2011 Duma election and the ballot box rebuke of United Russia, the stage was set for potential change in the 2012 Presidential election. Although the returning Putin remained the unquestioned favorite, measures were taken to improve transparency. Most notably, webcams were installed in virtually every polling place at a cost of several hundred million dollars (Bratersky 2012).

Despite the public outrage over the rampant fraud and these new efforts at reform, it appears that fraud remained entrenched in Russia's electoral system (see Figure 4.12). Observers continued to report brazen ballot box stuffing and other illegitimate practices (Barry and Schwirtz 2012). The data are consistent with this account. The irregular patterns in the data are similar to those seen in the previous fraudulent elections.

Figure 4.12. Turnout across all regions.

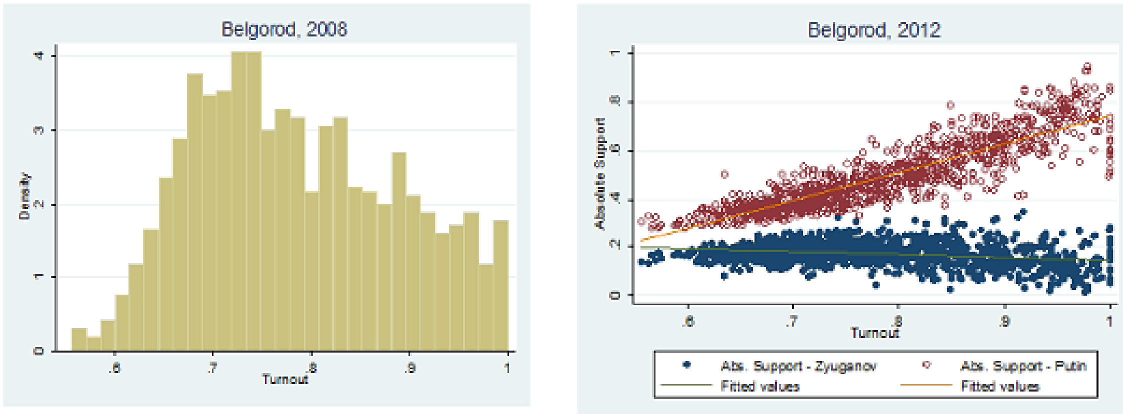


The regions which appear most problematic by SSD value include many of the same regions that had irregularities in previous elections: Chechnya, Mordovia, Tyumen, Yamal-Nenets, Chuckchi, and Tararstan have the highest SSD values for Putin.

While the average SSD for the UR candidate did not decrease, some areas showed improvements. Belgorod, for example, looks much cleaner than it did in 2008 (see Figure 4.13).

Despite this improvement, overall there is no indication that fraud was significantly reduced in the 2012 election. Despite the efforts at curbing practices such as ballot box stuffing, the data still contain many irregularities.

Figure 4.13. Belgorod in 2012.

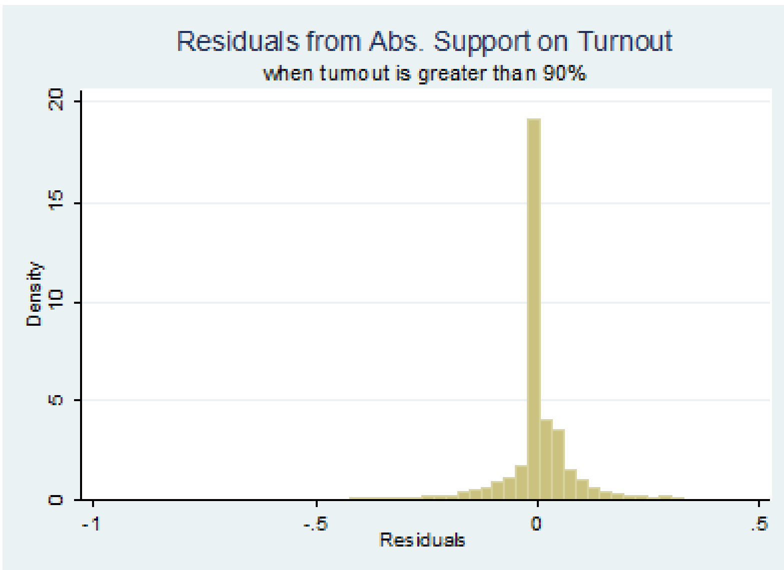


The Urban-Rural Divide

One of the most significant dimensions in Russian politics is the urban-rural divide (Berezkin et al. 1999). Putin and United Russia enjoy their greatest levels of support in small towns and rural areas, whereas the opposition is usually strongest in urban centers (see Figure 4.14). If higher turnout is more likely in those rural areas that support Putin, the irregularities in the data may be naturally occurring rather than the result of fraud.

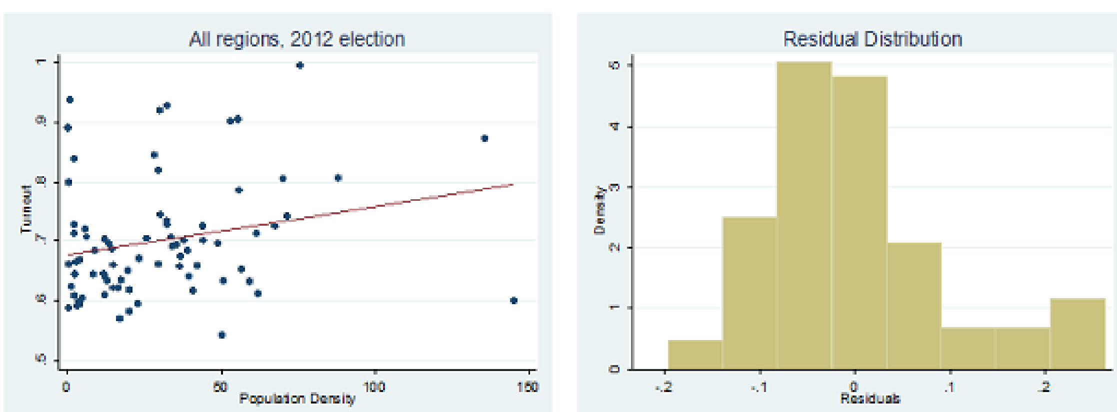
Using both the jaws technique and the residual analysis technique we can determine that the irregularities in the data are not due to natural non-normality caused by the urban-rural divide (see Figure 4.15). Using the jaws method, we notice that the data is shifted away from natural patterns in a way consistent with ballot box stuffing. After regressing absolute support for Putin on turnout within each oblast, analysis of the residuals reveals that they tend to be positive at high levels of turnout. The mean residual is greater than zero at more than the .01 level of significance. The distribution reveals a larger than expected concentration of residuals slightly larger than 0. This is consistent with the results observed at the oblast level previously analyzed in this chapter.

Figure 4.14. Residuals from the absolute support and turnout relationship.



This result is confirmed through residual analysis (see Figure 4.14). Regressing population density on turnout results in residuals that have a wide right-hand tail. This is consistent with ballot box stuffing, controlling for population density.

Figure 4.15. Residual analysis for the 2012 election. Oblast level data.



The urban-rural divide does not explain the irregularities observed in the data, and at the oblast level this divide does not explain the degree of fraud observed, either. Given our predictions in Chapter 3, we might suspect that there is a nonlinear relationship between urbanization and fraud. In very rural oblasts, Putin and UR may enjoy enough support that there is little room or need for fraud. In very urban oblasts, the costs of fraud may be too high. In Moscow, for example, Putin enjoys little support. Attempts at fraud there may be readily caught and seriously protested by the populace and the opposition. Fraud may therefore be more prevalent in regions where urbanization is close to the mean. The data does not support this conclusion. It appears that the level of fraud is independent of the level of urbanization. Regression results are included in Appendix D.

Conclusion

In response to public pressure, political elites in Russia have instituted electoral reforms. Some of these are aimed at systemic factors, such as reducing the potential for "virtual" candidates and increasing access to media. Some reforms have been aimed at reducing outright fraud, such as the installation of webcams. Despite these attempts at reform, it appears that election fraud remains a pervasive problem in Russia. The irregular patterns observed in the data from the 2007 elections are persistent. These patterns can be observed even when controlling for population density, and do not appear to be Natural Non-Normality. Nor do they depend on the characteristics of the region in question—patterns consistent with fraud are found in both urban and rural areas.

While countries like Mexico have been successful in depoliticizing the administration of elections, in Russia the elections are still overseen by partisan elites appointed by those in power. It seems unlikely that election fraud will truly be reduced until election officials lose their incentives to report election returns favorable to the establishment.

CHAPTER V

ELECTION FRAUD AND MEXICO'S DEMOCRATIC TRANSITION

Mexico underwent a transition from single party rule to multiparty democracy beginning in the late 1980s. Many consider the presidential election of 2000 to be the sign that the transition had been completed. This election saw the victory of Vicente Fox, the first person to hold the office outside of the *Partido Revolucionario Institucional* (PRI) since the Mexican Revolution. Emerging democracies can be fragile, however. Fox won with the opposition *Partido Acción Nacional* (PAN), one of the PRI's oldest and most success rivals. Despite the successes of the PAN, many feared that the new political system would be threatened by the old problems of fraud and corruption. Furthermore, the PRI party held a position more moderate than either of the main opposition parties. As the PRI lost power, new opportunities for internal strife arose between the PAN and the *Partido de la Revolución Democrática* (PRD).

Given its political monopoly, the PRI enjoyed success largely due to its popularity (Magaloni, 2006). This popularity was instrumental in sustaining one-party rule until the 1980s, when economic difficulties hit the country along with the rest of Latin America. The PRI party, finding its support waning, had to resort to massive election fraud in order to retain its power. The 1988 presidential election is widely regarded as one of the most fraudulent in the nation's history. When the election returns looked bleak for the PRI, the computers tallying the votes were shut down. Even the sitting president, Miguel de Madrid, has stated that the PRI should have won by a very narrow margin at best, and certainly did not earn the massive margin of victory that was ultimately reported

(Thompson, 2004). Ultimately, this election drove up pressure for electoral reform. This reform was enacted through several new laws and through increases in election monitoring, paving the way for truly competitive elections and the end of the PRI party's dominance in the presidential election of 2000.

While the election of 2000 was heralded as a triumph for clean elections, the election of 2006 created a great deal of controversy. For the early part of election day, election returns showed Andrés Manuel López Obrador of the PRD winning by a substantial margin. Felipe Calderón of the PAN gained ground as more polling stations reported their results, and pulled ahead as the last polling places reported. Calderón was the winner by a very narrow margin. The timing of his comeback win was viewed with suspicion and firsthand accounts of election fraud surfaced. Obrador refused to accept the election results and demanded a recount. A limited recount was authorized as the PRD engaged in large protests. Calderón was ultimately declared the winner of the election by less than 250,000 votes. The PRD remained irate and threats of civil unrest called into question the stability of Mexico's fledgling democracy.

The PAN faced a fair amount of public resentment heading into the 2012 presidential election. The Calderón administration's war on the drug cartels had resulted in massive violence. Economic growth remained sluggish during the global economic downturn. The PAN declined in popularity and finished third in the 2012 presidential election. The PRI emerged as the winners, with Enrique Peña Nieto earning 38% of the vote to Andrés Manuel López Obrador's 32%. Josefina Vázquez Mota of the PAN ended up in a distant third with 25% of the vote.

Again, accusations of fraud plagued the elections. Key among them was "Sorianagate," claims that the PRI purchased gift cards to Soriana grocery stores and distributed them to voters in exchange for their vote (Miroff and Booth 2012). Opponents argued that vote-buying was widespread and that the election had no legitimacy. The Alianza Cívica, a watchdog group, issued a report indicating that vote buying and coercions were major problems in the election (Alianza Cívica, 2012).

Despite allegations of fraud, analysis of the 2006 and 2012 Presidential elections suggests that the elections were relatively clean and free of fraud. It does not appear that the impact of electoral reform has diminished in Mexico.

Election Fraud in Mexico

Election fraud has occurred at varying levels throughout the PRI regime. It was so endemic that the party's candidate in the 1940 election ran on a platform of reducing fraud and overseeing clean elections, only to have his election marked as one of the most fraudulent in the nation's history (Cothran, 1994, p. 44). Fraud likely peaked again in the 1988 election. Fraud continued after this election despite efforts to minimize it. Nearly every strategy of altering vote counts has been tried by "electoral alchemists" at every level of elections.

One of the most prevalent types of election fraud is "vote inflation," increasing the vote count for a particular candidate. This is commonly known as "stuffing the ballot box." In Mexico this is done in two ways. The simplest method is known as the "taco," where extra ballots are just folded together and shoved into the ballot box (Cothran, 1994, pp. 199-200). In a method that is both more complicated and more colorful, establishment party supporters are loaded into a bus and driven from one voting center to another. The "*raton loco*" (or "crazy mouse") drops them off, they vote, and then they board the bus for another location to vote again (Cothran, 1994, p. 200).

An alternative to increasing the votes for one candidate is reducing votes for another. Reports abound of discarded ballots during Mexican elections. Piles of ballots have been found in trash cans and creek beds (Cothran, 1994, p. 202). This has been practiced as recently as the 2006 presidential election, when PRD party supporters have argued that the PRD influenced the outcome of the election. There are other, less direct means of reducing vote counts for a rival candidate. Voter repression can be done by making it more difficult for opposition supporters to vote or, in extreme cases, by threatening violence. The PRI government employed tactics to make registration difficult in regions that had previously supported opposition candidates (Cothran, 1994, p. 198).

A third type of election fraud is the switching of votes from one candidate to another. This could be achieved by "buying votes," by paying voters to vote for a particular candidate. Based on historical evidence, ballot box stuffing and the destruction of votes appear to be much more common practices. Still, reports of vote buying have surfaced as recently as the 2000 election (Frohling and Gallaher 2001, p. 1). Vote buying was believed to be a major factor in the elections of 2009 and 2012 (Alianza Cívica 2012).

Electoral Reforms

Due to the prevalence of fraud in Mexican elections, reform was a major focus during the transitionary period. These reforms took shape in a variety of ways and were spread out over a decade. They were largely inspired by the 1988 presidential election. Following that election, pressure for electoral reform increased dramatically.

In 1990, the Chamber of Deputies passed the Código Federal de Instituciones y Procedimentos Electorales (COFIPE) and formed the *Instituto Federal Electoral* (IFE). Opponents argued that the reform did not go far enough and that it provided more political power to the PRI (Beer 2003, p. 187-189). In spite of these concerns, the bill passed with support from the PAN, and provided a first step towards addressing the problem of electoral fraud. Though the IFE had limited power and independence, there now existed a body for monitoring elections. Critics argued that the law left too much power in the hands of the government. The PRI maintained a strong presence within the IFE. The president of the IFE was the Secretary of the Interior, and membership on the board was based on representation in congress and on Presidential appointment (Klesner 1997). Further reforms in 1994 addressed this issue by granting independence to the IFE (Levy and Bruhn 1999, p. 545; Magaloni 2005, p. 124). Magaloni argues that the PRI had good reason to believe that it would win the election handily, and that it therefore offered this critical piece of election reform in order to appease the Zapatista rebellion (Magaloni, 2006). The 1994 reforms saw the board members representing the political parties lose their votes and the Presidential appointments replaced by "citizen councilors" who were selected by all three major parties (Klesner 1997). When the time came for the 1994 elections, several organizations came together to form the *Alianza Cívica* (Hellinger, 2011, pp. 267-268). This group organized to monitor elections beginning in 1994, which were won by the PRI party candidate Ernesto Zedillo.

These electoral reforms led to the perception that the 1994 presidential election was largely free from fraud. Most believe that Zedillo would have won the election regardless of the reported incidents of fraud (Peeler, 2004, p. 76). Reform continued during the Zedillo administration. A 1995 pact between the PRI and PAN parties was another step towards making elections more legitimate (Peeler, 2004, p. 77). Key among the reforms during Zedillo's presidency was the continued independence of the IFE. In 1996 the IFE was granted even more independence (Levy and Bruhn 1999, p. 545)

It is arguably these reforms, combined with the PRI's emphasis on maintaining popular support through elections, that led to the emergence of competitive democracy in Mexico. The election of Vicente Fox in 2000 was heralded as the cleanest ever in modern Mexico, and was the crowning moment in Mexico's transition from one-party hegemony to a competitive democracy. Throughout the reform era, however, critics argued that the new rules did not go far enough or that they maintained the establishment's grip on power. Following the 2006 presidential election, multiparty democracy appeared to be fragile as the results were decried as illegitimate by the opposition and civil unrest ran high. The legitimacy of the 2012 Presidential election was likewise questioned despite efforts at further reform.

Election Analysis

Data for Mexico’s presidential elections is available through the IFE. Income data come from *Instituto Nacional de Estadística y Geografía* (INEGI, www.inegi.org.mx). The data illustrates the effectiveness of electoral reform. Turnout data is available at the district level for the 2000, 2006, and 2012 elections. These turnout distributions show that while the 2000 election appears relatively clean, the 2006 election showed a marked improvement (see Figure 5.1). Regarding the relative support for the PAN, the most successful opposition party, patterns in its relative support appear which are consistent with vote destruction as recently as 1994. This shift is evident from the large spike near 0% relative support (see Figure 5.2). This irregularity suggests that there may have been a large number of polling places where all votes for the PAN were discarded. As reforms took place, this phenomenon decreases and then disappears, illustrated by the leveling off of this spike in the 2000 election and disappearing by 2006.

Figure 5.1. Turnout 2000, 2006

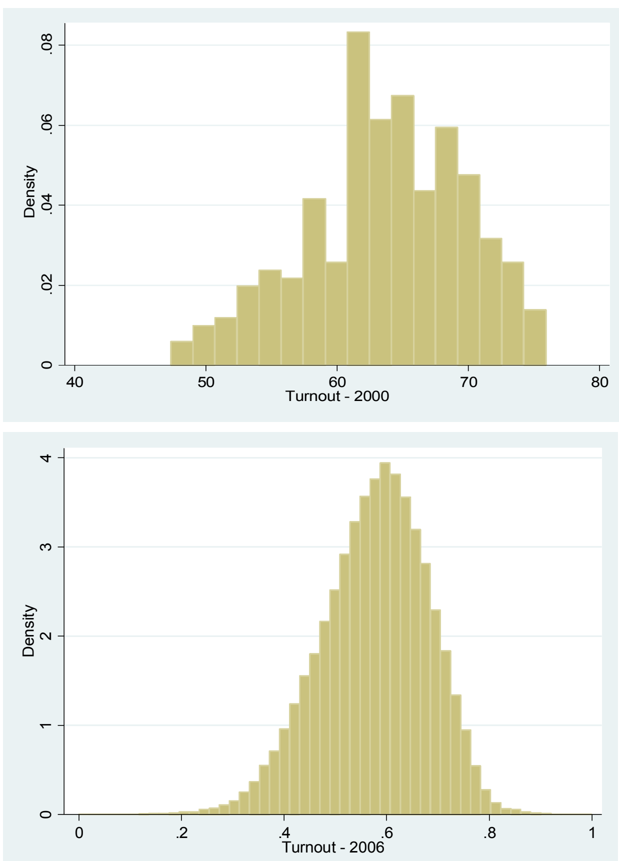
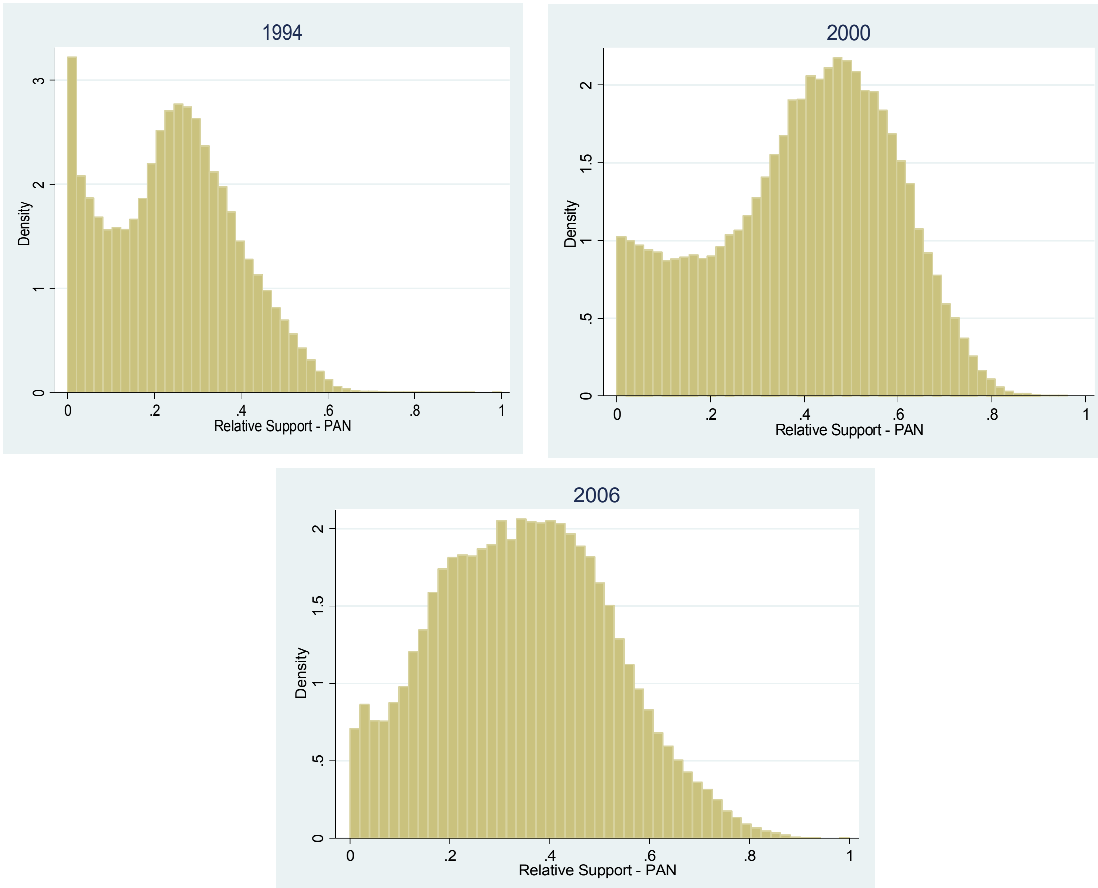


Figure 5.2. Relative support for the PAN, 1994-2006



Analysis of the election data for 2006 in closer detail reveals few irregularities. Turnout appears to be normally distributed in all but one of the 32 federal entities. Absolute support also appears to be normal, suggesting that there were no large-scale occurrences of fraud. I perform statistical tests for skewness on a sample of 100 polling places, examining turnout and absolute support for each candidate. The results of these tests are reported in Table 5.1. The *p* value indicates the degree of confidence with which we can reject the null hypothesis of a symmetric distribution. If the *p* value is very small, the distribution is weighted in one direction or another, and we have an irregular pattern.

If the *p* value is .01 or greater, we cannot reject the null and the data is consistent with data from a clean election. Basically, a small *p* value (.01 or less) means increased likelihood of fraud. Also in the table is the sum of squared differences (SSD) between a candidate's relative support and the slope coefficient from the regression of absolute support on turnout. As noted earlier, we expect that the relationship between those two variables should be equal to the candidate's relative support. In order to measure how large the difference is between our expectations given a clean election and the actual vote returns, I take the difference between the two and square it. The values for the SSDs are fairly small. These results illustrate that increased turnout did not typically help or hurt any given candidate, suggesting that there was not widespread vote inflation or vote destruction.

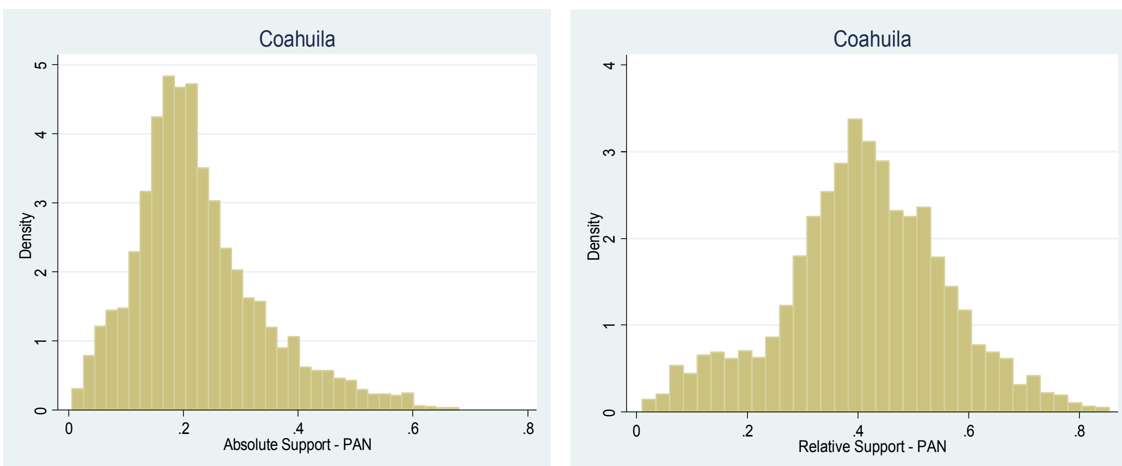
Despite the fact that the patterns in the 2006 Mexican presidential election appear much more regular than those found in Russia and analyzed in chapter 4, Several areas show interesting irregularities requiring further analysis. Several states fail the tests for skewness in each of three key areas: turnout, absolute support for the PAN, and absolute support for the PRD. In order to determine if the patterns are consistent with serious levels of fraud, I will analyze those areas in more detail.

Table 5.1. Results of tests for skewness and sum-of-squared-differences (SSD). 2006 General election

State	P-value, Turnout	P-value for PAN	P-value for PRI	P-value for PRD	SSD PRI	SSD PAN	SSD PRD
Aguascalientes	0.1748695	0.0001034	0.7990565	0.0167008	0.0849004	0.1062451	2.09e-06
Baja California	0.1087076	0.0076445	2.30e-06	0.0329218	0.0117786	0.0353563	0.0061756
Baja California Sur	0.4129414	2.39E-009	0.9604654	0.0695551	0.0001218	0.0021644	0.0008079
Campeche	0.3359162	0.8518066	0.0001253	0.6811587	0.0114953	0.0001476	0.0350149
Coahuila	0.1036673	0.0042776	6.42e-13	1.23e-07	0.1012917	0.1353589	0.0008468
Colima	0.7315931	0.1238657	1.88e-09	3.00e-07	0.0208428	0.0384577	0.0017638
Chiapas	0.3084353	6.65e-06	0.0011187	0.8681015	0.0110855	0.0004673	0.0090823
Chihuahua	8.10e-14***	0.0003674	0	9.83e-06	0.0317414	0.075949	0.0057144
Mexico City	0.3291038	1.30e-07	0.9590139	1.15e-07	0.0006708	0.5197712	0.4597674
Durango	0.0494891*	0.0478002	8.88e-15	0.0537876	0.0465403	0.030296	0.0034391
Guanajuato	0.4373647	0.9729945	0.0011679	0.0017837	0.024521	0.1241726	0.0259871
Guerrero	0.0023338**	2.48e-10	7.98e-06	2.21e-09	0.0017103	0.0039799	0.0003978
Hidalgo	0.777334	0.0492698	3.45e-12	0.7039589	0.0088734	0.013465	8.82e-07
Jalisco	0.1630925	0.004917	1.60e-07	0.8186808	0.0661156	0.1059861	0.0013112
Chimalhuacan	0.0001775***	7.67e-08	0.0000157	0.111313	0.0539918	0.38506	0.1295589
Cuernavaca	0.0408236	0.0004489	2.35e-13	0.0044462	0.0220182	0.1690082	0.0637571
Morelos	0.9079623	0.001828	4.57e-07	0.0330354	0.030644	0.0785001	0.0029434
Nayarit	0.0109863*	0.000036	1.80e-09	0.0323849	0.0002807	0.0018362	0.0030118
Nuevo Leon	0.6532347	0.0311044	0.0002182	0.0000203	0.2747871	0.4758982	0.0083774
Oaxaca	0.0610253	4.00e-15	2.61e-08	0.1705573	0.0150221	0.0013634	0.0019563
Puebla	0.2901594	0.2881085	0.0002446	0.0615274	0.0734063	0.1273675	0.0062859
Querétaro	5.77e-11***	0.3812663	7.40e-10	0.0088344	0.1222339	0.1190585	0.0009588
Quintana Roo	0.0670164	0.0013132	5.87e-08	0.0000278	0.0792577	0.0003326	0.056732
San Luis Potosi	0.1702349	0.8167412	0.0000205	0.0066832	0.0010293	0.0013615	0.0006317
Sinaloa	0.5461032	0.2010081	0.0002741	0.2486984	0.0063389	0.0088434	0.0001392
Sonora	0.5461068	0.2198139	9.11e-09	0.010997	0.0040777	0.0107616	0.0190659
Tabasco	0.7425913	1.61e-08	0.2161954	0.7105193	0.0101278	0.0007969	0.0257461
Tamaulipas	0.0325431*	7.78e-06	6.17e-13	0.005429	0.0066738	0.0000422	0.0031127
Tlaxcala	0.2121896	0.00208	0.0066369	0.2894216	1.18e-06	0.0079512	0.006335
Veracruz	0.5573177	0.046949	0.001589	0.095877	0.0012619	0.0001254	0.0010132
Yucatán	0.5327845	0.0444302	1.83e-07	4.58e-07	0.00084	0.0252077	0.0072643
Zacatecas	0.0201637*	0.8877282	1.54e-09	1.74e-07	0.0009617	0.0008189	0.0000348

Several of the states which failed a skewness test illustrate an interesting problem with this measure. Since absolute support cannot fall below zero, the distribution of absolute support may be truncated. In this case the distribution is asymmetric due to the left-hand tail being lopped off at zero rather than a shift in the data caused by vote manipulation. This can be seen in the cases of Coahuila, Guerrero, and Cuernavaca (see Figure 5.3). In these cases, we may look at relative support instead of absolute support. Relative support will be higher than absolute support, since nonvoters are not counted. This will move the entire distribution to the right, avoiding truncation problems.

Figure 5.3. Truncated absolute support versus relative support.

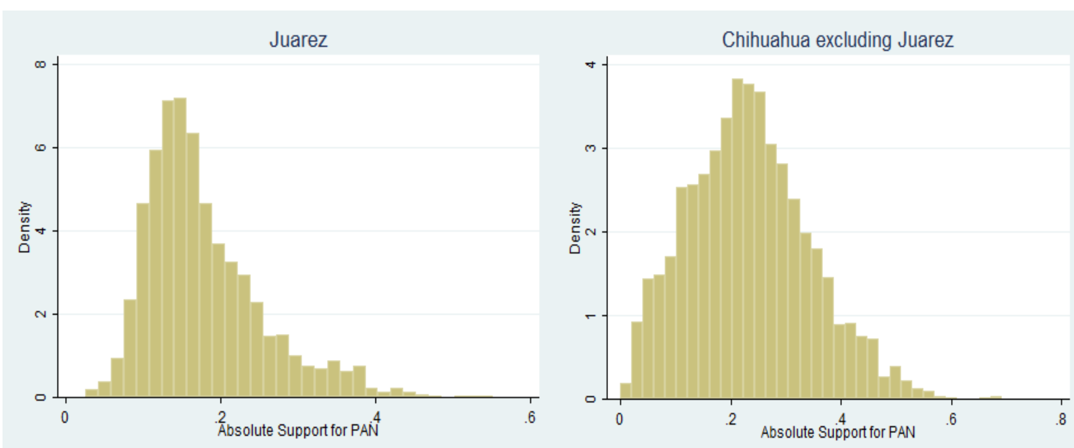


Of the four other states which have particularly suspicious results on the skewness tests, Tamaulipas has distributions which actually appear fairly normal to the eye. In this case, the skewness test might be too sensitive, as the distributions may not be perfectly symmetric but are certainly far from what we see when fraud is widespread. Quintana Roo has distributions of turnout and absolute support that are somewhat irregular, but are not clearly shifted in one direction or another. The relationship between absolute support and turnout does not indicate vote inflation or destruction. Chihuahua and Nayarit, however, are interesting cases.

In Chihuahua (see Figure 5.4), the distribution of absolute support for the PAN has a wide right-hand tail. Absolute support for the PRD appears normal. Turnout appears roughly normal, but does have a slight weighting on the right-hand tail, which would lead us to expect vote inflation for the PAN, but no alteration of PRD votes. This appears to be confirmed by regressing absolute support on turnout in the region. The coefficient for the PAN is .765, while its relative support in the state is only .427.

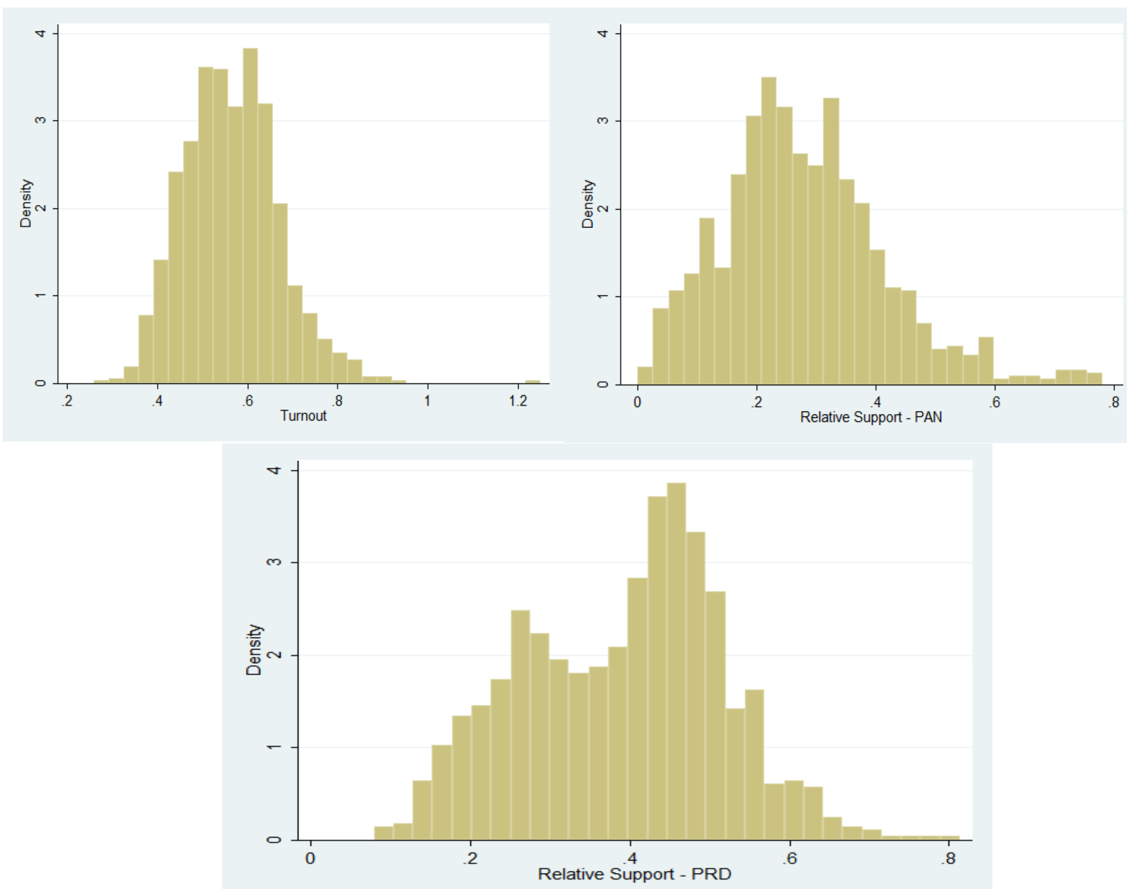
A clearer picture can be obtained by looking at the individual districts within the state. When looking at the distribution of support for the PAN, the city of Juarez displays an irregular pattern where the rest of the state appears fairly normal.

Figure 5.4. Juarez and Chihuahua state, 2006



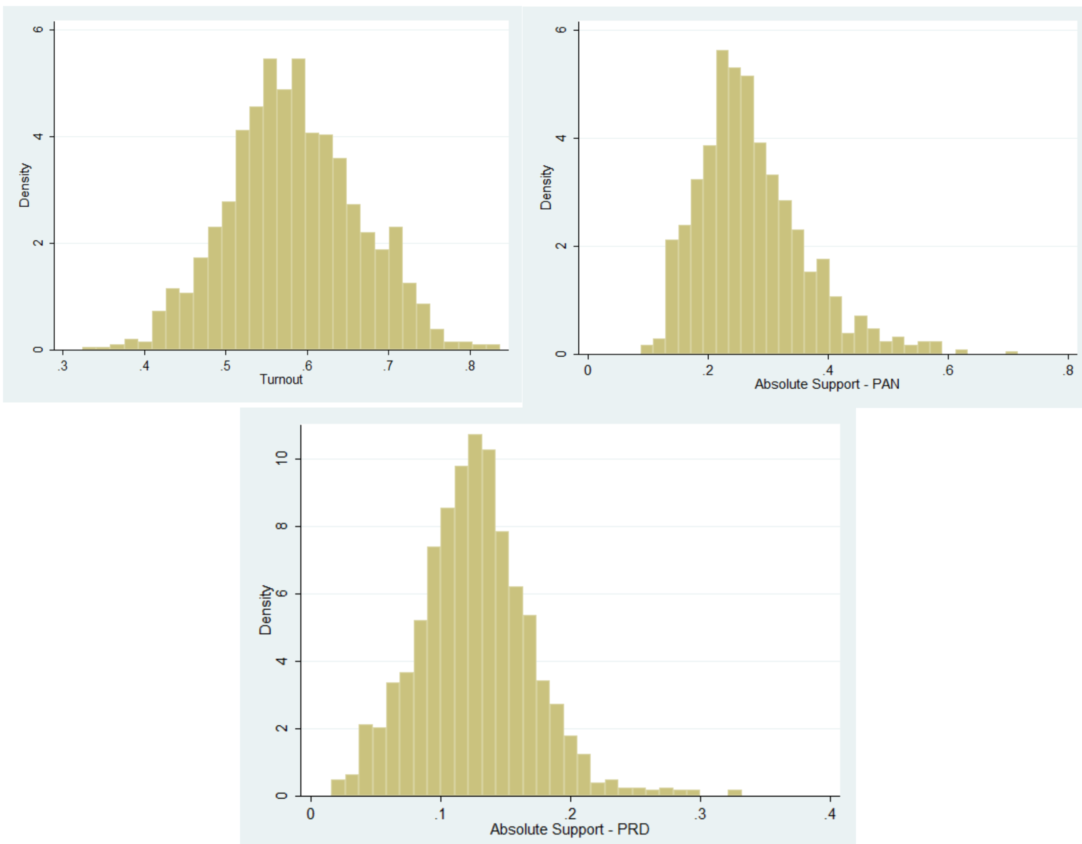
In Nayarit (see Figure 5.5), turnout appears to be relatively normal. Nonetheless, the right-hand tail on the distributions of support for the PAN is wide as is the left-hand tail for the distributions of support for the PRD. This outcome is consistent with minor vote switching from the PRD to the PAN. These irregularities are fairly minor, however, and do not suggest widespread fraud. Still, they are enough to raise alarm bells. Future research should examine this region more closely.

Figure 5.5. Turnout and support in Nayarit, 2006



Another state of particular note is the state of Aguascalientes (see Figure 5.6), which was treated with particular suspicion during the recount. Despite scrutiny, the data does not present any irregularities in this state. The distributions for turnout and absolute support appear fairly normal. Absolute support for the PAN does have a little extra width on the right side, and absolute support for the PRD appears more normal. This distribution might suggest vote inflation, except that turnout appears normal. Widespread vote inflation is therefore unlikely in this case.

Figure 5.6. Turnout and absolute support in Aguascalientes, 2006



The results of the 2012 presidential election likewise display few irregularities. Turnout for the nation as a whole appears normally distributed. Absolute and relative support also appear normal. Skewness measures for each state are low in magnitude. If vote buying or voter intimidation occurred, they were likely at fairly low levels. This suggests that, counter to the Alianza Civica's report, the problems in Mexico's elections are likely declining rather than intensifying.

The Alianza Civica reports that voter intimidation and violation of secret ballot laws may have occurred in Mexico state, Chihuahua, Sinaloa, and Jalisco (Alianza Civica 2012, p. 3). In these states, Alianza Civica reported that children were used to spy on voters. Media reports indicated that the PRI had given voters gift cards to the Soriana grocery chain in an effort to buy their votes. These reports appear most common in Mexico state, the Federal District, San Luis Potosi, and Monterrey (see Figure 5.7). The skewness measures for support for the PRI are presented in Table 5.2. Despite the allegations, few irregularities are present in these states. While the distribution of support in Chihuahua appears suspicious, the other states have much less skewness. This data suggests that reported vote stealing or intimidation were either limited in scope or unsuccessful.

Figure 5.7. Relative support for the PRI in Mexico state and the Federal District.

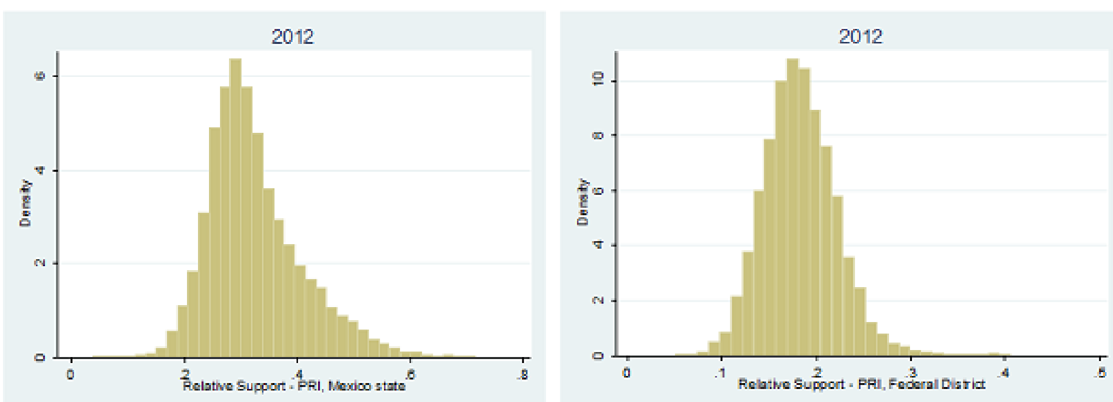
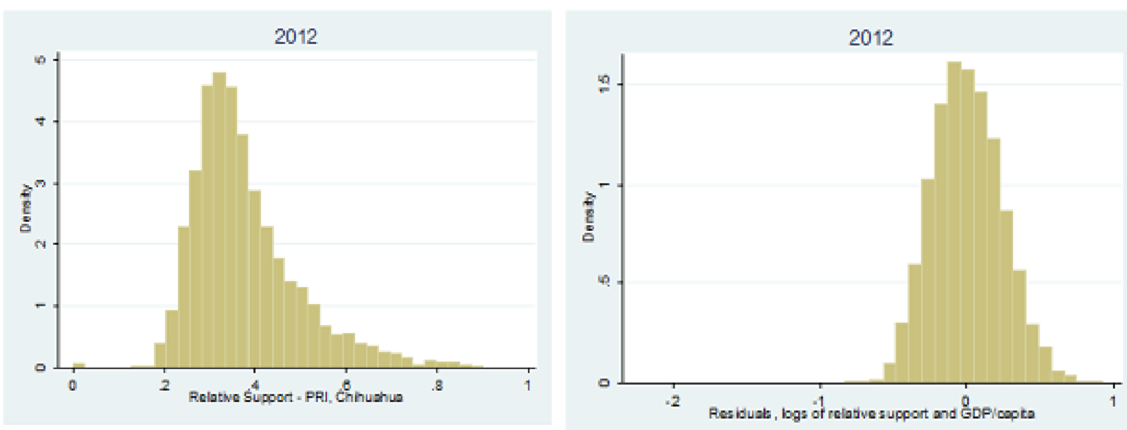


Table 5.2. Relative support in regions with reported vote buying

State	Mean, relative support for PRI	Skewness, Relative support for PRI
Chihuahua	.377	1.185
Federal District	.184	.765
Jalisco	.314	.712
Mexico State	.325	.904
Nuevo Leon	.280	.459
San Luis Potosi	.296	.878
Sinaloa	.382	1.086

Although Chihuahua appears to have a skewed distribution for the PRI's relative support, this distribution is mitigated to some extent when controlling for income (see Figure 5.8). Data on GDP per capita is available at the municipal level.

Figure 5.8. Relative support for the PRI in Chihuahua, residuals from regressing on GDP per capita.



Conclusion

For most of the twentieth century, democracy in Mexico was characterized by single-party rule. Elections were used by the PRI as a means of legitimizing and cementing their power. Popular support and the ability to deliver favorable election returns by controlling the electoral process contributed to the regime's stability. The problems of the 1980s ultimately forced the PRI to change its approach. Multiple factors led to the establishment of clean elections. After the 1988 elections, public pressure for electoral reform was very high (Cothran, 1994, p. 187; Hellinger, 2011, p. 267). Leading up to the 1994 elections, the Zapatista uprising put additional pressure on the PRI. Confident that it could win the election, and wanting to stem violence, the party agreed to further electoral reforms that included granting independence to the IFE (Magaloni, 2006). Given the emphasis on elections as a foundation of the party's power and legitimacy, the PRI had “raised the cost of blatant electoral fraud” (Levy and Bruhn 1999, p. 546). The sweeping electoral reforms of the 1990s resulted in a dramatic reduction in fraud and paved the way for an opposition candidate to win the presidency. When Vicente Fox won the office in 2000, it appeared that Mexico's transition to a competitive democracy had been successful.

Still, democracy in Mexico has undergone many challenges. Although presidential elections have become much more legitimate, corruption and fraud remain in many areas and at lower levels of government (Lawson 2000). Additionally, the competitiveness of elections is still limited by access to resources, and campaign finance reform has been difficult (Eisenstadt and Poiré 2005). Perhaps more ominously, the 2006 election was marred by controversy. When PRD candidate Andrés Manuel López Obrador lost his large lead in the closing hours of the election, suspicions were raised. Allegations and firsthand reports of election fraud fueled tensions, and the legitimacy and stability of Mexican democracy came into doubt. While massive civil unrest was avoided, the fledgling system clearly remained fragile.

The 2012 elections resulted in similar accusations. The outrage following these elections was less intense than that which followed the 2006 elections, likely due to the fact that the PRI won by a large number of votes. Still, given the PRI's history of vote fraud, suspicions were high.

Despite the allegations, it appears that both the 2006 and 2012 elections were clean and that electoral reform has been successful. No evidence exists of widespread election fraud. In comparison to places where fraud is pervasive and ongoing, such as in Russia, all of the critical distributions appear normal. The 2006 elections appear to be even more legitimate than the vaunted elections of 2000 that first established multiparty democracy. Given the sensitivity of the forensic techniques and the closeness of the election, however, the possibility that minor amounts of election fraud were present and affected the outcome cannot be ruled out. In contrast, the margin of victory for the PRI in 2012 suggests that any minor election fraud which did take place did not alter the result. Doubts about Mexico's democracy following these elections are unfounded, and efforts for further reform should focus on problems that still exist.

While Mexico still faces serious obstacles in its effort to develop democracy, there does not appear to be any “backsliding.” The data from the last three elections should lead to increased confidence. Given the damage done to the electoral system's credibility, the stakes for the next election are very high. To further complicate the situation, Andrés Manuel López Obrador is making another run, and public opinion polls indicate that the race may be close once again.

CHAPTER VI

ELECTION FRAUD AND DISENFRANCHISEMENT DURING RECONSTRUCTION AND REDEMPTION IN THE U.S. SOUTH

Historically, election fraud has played a large role in American politics. In this chapter I analyze presidential elections during the Reconstruction and Redemption eras in the U.S. South. The consensus in the literature is that this period was marked by widespread fraud in virtually every form.

I find evidence that the historical accounts are accurate. During Reconstruction (1867-1877) which followed the U.S. Civil War, Democrats engaged in fraud, likely at local levels where they held more power. Nearly a decade after Reconstruction had ceased, Democrats in the South were able to institute legal measures, such as poll taxes and literacy tests, to suppress the vote for the opposition. The irregularities in the data are primarily found in the Deep South states of Louisiana, Mississippi, Alabama, Georgia, and South Carolina. Unsurprisingly, these are also the states where restrictive voting laws were first adopted and enforced (Kousser 1974). While other forms of fraud were no doubt common, vote destruction appears to have been the primary means for the Democrats to engage in electoral manipulation in the South.

The efficacy of election fraud and voter intimidation tactics in the South is unclear. Key (1949) argues that the disenfranchisement of black voters and the establishment of one-party rule predated the codification of suffrage restricting laws. Terry (2013) points out that the effects of lynching as a means of violent vote suppression are not so clear. Empirical research is ambiguous, and the author finds only weak evidence that lynching increased Democratic vote shares (Terry 2013, p. 33). Kousser (1974) takes direct issue with Key's thesis, arguing that the establishment of antisuffrage laws such as poll taxes and literacy tests played a vital role in suppressing the black vote and creating the solid South.

This chapter uses election forensics to contribute to this debate. I attempt to measure the impact of vote manipulation before and after the institution of suffrage restricting laws and find that vote suppression was pervasive prior to the establishment of these laws. While the patterns consistent with vote destruction become stronger as these laws come into effect, Key appears to be correct in his assertion that black voters were suppressed by other means well in advance of tactics such as poll taxes and literacy tests.

In addition, the case of the U.S. South allows me to test the tools developed in Chapter 2 on a case where fraud may have been even more pervasive than in the cases of Russia or Mexico. The legalization of electoral malpractice found in the U.S. South may have been more extreme than the electoral authoritarian regimes analyzed elsewhere in this dissertation.

Previous Research

Campbell (2005) provides accounts of election fraud throughout American history. The most visible form of election fraud in the South during the Reconstruction and Redemption eras was voter intimidation (Campbell 2005, pp. 58-72). The Ku Klux Klan regularly engaged in both threats and actual violence. Other more direct methods were also employed to suppress the black vote. Illegal poll taxes were enforced against blacks, and votes cast for Republicans were discarded (Campbell 2005, p. 62).

Election fraud in the South was extremely pervasive. In the 1872 presidential election, the electoral votes for both Louisiana and Arkansas were thrown out due to irregularities. The infamous 1876 presidential election was likewise marred by extreme election fraud (Campbell 2005, pp. 66-71). The Republican vote was suppressed or stolen across the South, and the Republican party responded by using its control of the election boards to alter the result.

Over time, systemic changes caused the nature of fraud to change. Once Reconstruction had ended, Southern states fought off attempts to have elections monitored by the federal government (Campbell 2005, pp. 92-93). As federal oversight decreased, Southern states began to add the strategy of disenfranchisement to the toolbox of electoral manipulation. Literacy tests and poll taxes began to be used to suppress the African-American vote in a form of vote destruction. At the same time, vote buying was a common practice in the late 1800s, in which voters were paid in an effort to steal votes from the opposition and inflate vote totals (Argersinger 1985-1986, p. 673). Key (1949) argues that the methods employed in the 1880's successfully solidified the South as a one-party regime.

Kousser (1974) offers an in-depth look at the system of disenfranchisement which came to dominate the South in the final decades of the 19th century. Once Reconstruction ended and Democrats were able to regain more control of their state governments, they began to introduce legislation explicitly aimed at reducing votes for the Republican party, whether those came from white or black voters. Kousser asserts that it was these laws that had the greatest impact in disenfranchising the Southern Democrats' opposition.

Analysis

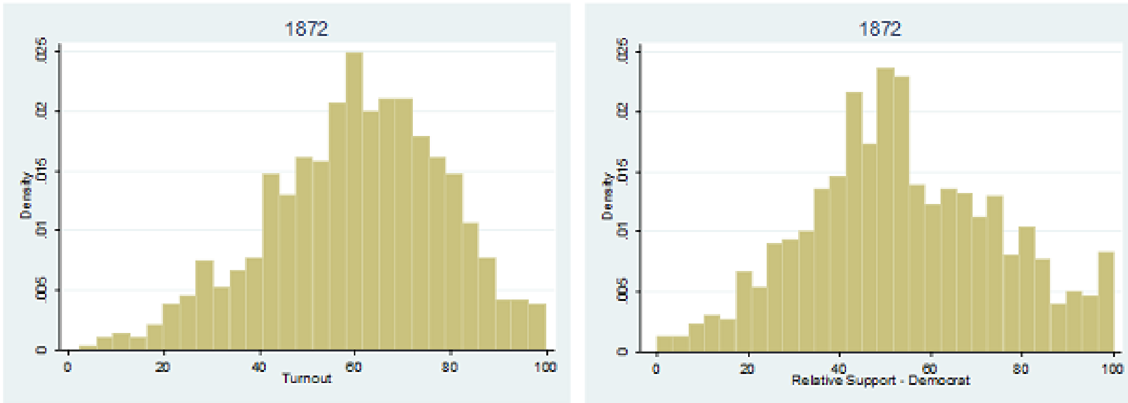
In order to examine the nature of electoral manipulation in the South, I look at county level data on turnout and vote shares in presidential elections. This data is available from the ICPSR (Clubb 2006). I also include demographic information from the U.S. Census Bureau, also available from the ICPSR (Haines 2010).

Reconstruction Era Elections

According to Foner (1988), "1872 witnessed the most peaceful election of the entire Reconstruction period" (p. 508). Republicans performed well in the South, prompting the party to claim that "in a peaceful election, they constituted the South's natural voting majority" (Foner 1988, p. 508). To the extent that there had been fraud, it possibly benefited the Republicans rather than the Democrats. Louisiana and Arkansas both sent electors for Grant, but their electoral votes were nullified by Congress.

Metrics indicate that this election had few irregularities (see Figure 6.1). Both turnout and relative support for the Democratic candidate are relatively normal. Turnout for Grant has a mean of 62.5% with skewness of -.29. Mean relative support for Greeley was 51.7% with skewness of .159.

Figure 6.1. Turnout and relative support for Greeley (D) in Southern states in 1872

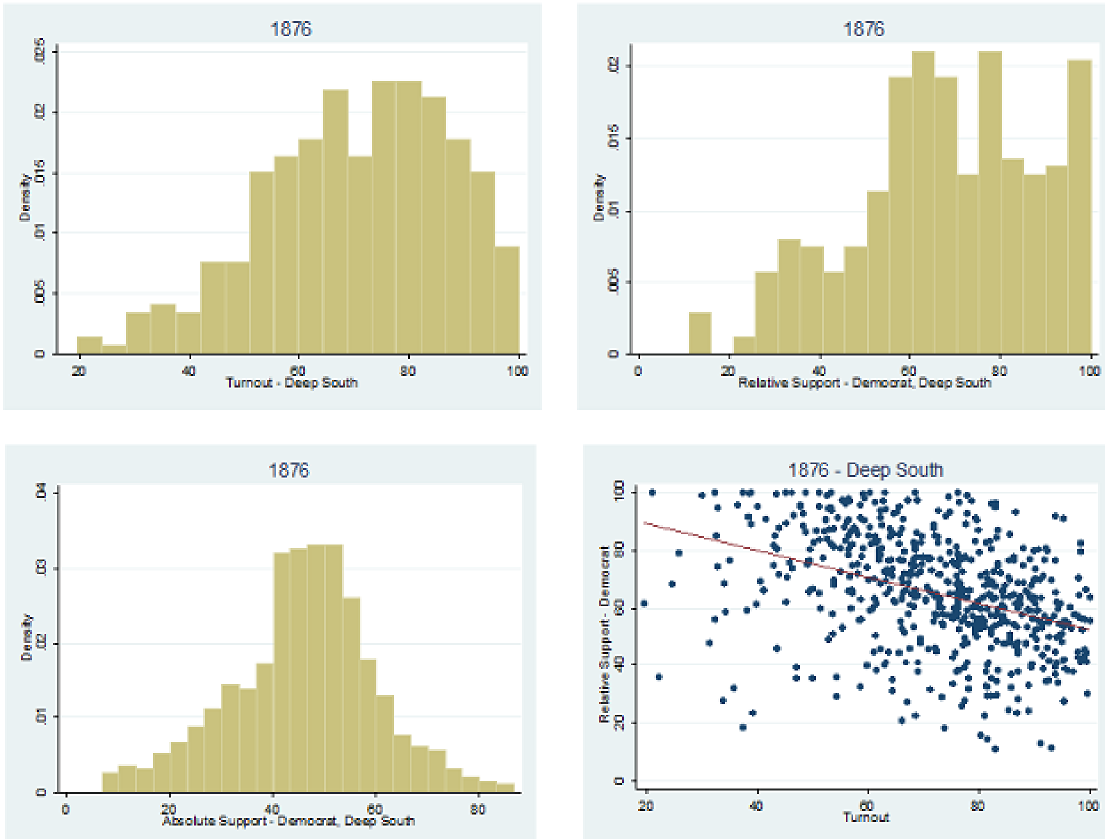


In contrast to the election of 1872, 1876 saw one of the most controversial elections in U.S. history. Events of 1875 had left the Republican party much weaker in the South (Foner 1988, p. 569). Democrats increasingly turned to intimidation in an effort to suppress the black vote (Foner 1988, p. 570). Evidence suggests that this intimidation was effective at the local level. Monroe County, Mississippi, went for Grant in 1868 but switched amid "over a thousand affidavits of men who had been driven away from the polls in Monroe County by force and threats" (Campbell 2005, p. 67). Still, the Republican side had the potential to alter the vote counts and the outcome. As early returns showed Tilden as the likely winner, Hayes could be the winner if he were able to take South Carolina, Florida, and Louisiana, states "where the [Republican] party controlled the voting machinery" (Foner 1988, p. 575). The Republican party chairman sent a telegram to party leaders in those states, asking them to "hold their state." The end result was electoral manipulation in both directions. Republicans felt that violence, intimidation, and fraud at local levels cost Hayes the Southern states, while Democrats felt that they had been "cheated out of office by Republican office-holders at the state level and in Congress" (Campbell 2005, p. 73). Historians have disagreed over who would have won the election had it been conducted legitimately (Campbell 2005, p. 77). In the end, the election resulted in the infamous Compromise of 1877 which brought an end to Reconstruction.

For this election, the metrics point to some degree of vote destruction in favor of the Democrat Tilden. While turnout appears fairly normal, some evidence suggests a wide left-hand tail in the distribution (see Figure 6.2). Skewness decreased from that observed for the South in the 1872 election, with a skewness of -.47. The change in relative support for the Democrat is more noticeable. There is a large spike near 100% relative support for Tilden (D). This spike does not appear in the distribution of absolute support (which appears fairly normal), suggesting that artificial increases in the Democrat vote counts were rare. The irregularity in relative support, therefore, would have been caused by vote destruction rather than vote inflation. Support for this hypothesis can be seen in the correlation between relative support for Tilden and turnout. The correlation coefficient is -.39, suggesting that those counties in which Tilden had high levels of support also had low levels of turnout.

The data appears to corroborate the Republican claims that the vote count was affected by efforts to intimidate and suppress the Republican vote. Democrats argued that Tilden may have won Louisiana, South Carolina, or Florida. This argument is more difficult to evaluate due to a low number of observations. Data is available for just 20 counties in South Carolina, 38 counties in Florida, and 54 parishes in Louisiana. No clear patterns emerge in those data.

Figure 6.2. Turnout, relative and absolute support for Democrats in the Deep South (Alabama, Mississippi, Louisiana, Georgia, and South Carolina)



Post-Reconstruction Elections

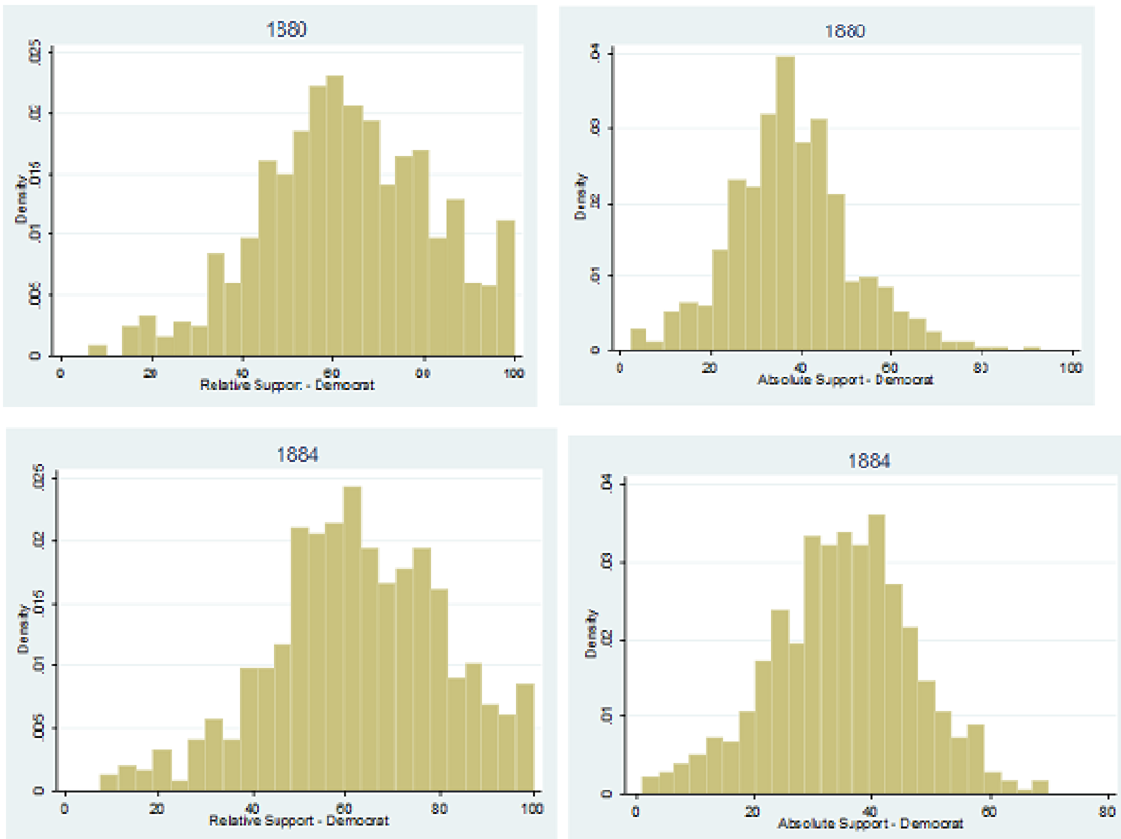
Following the 1876 election, Democrats in the South were poised to reclaim power thanks to the withdrawal of federal troops. Prior to the absolute death of Reconstruction, Democrats were forced to rely on relatively subtle or indirect methods of fraud and disenfranchisement. After it was clear that Reconstruction would not recommence, Southern Democrats were able to be more bold (Kousser 1974, pp. 45-46). Ballot box stuffing and vote stealing became commonplace during the 1880s (Campbell 2005, p. 89; Kousser 1974, pp. 46-47).

In 1890, the Lodge Bill aimed to provide more federal oversight and ensure fair elections. When this bill failed, the strategy of disenfranchisement became cemented in the South. By 1901 the entire Deep South had laws aimed at restricting the franchise (Kousser 1974, p. 239). This systematic disenfranchisement, along with the advent of the secret ballot, had serious effects on turnout and votes for the Republican party. Kousser estimates that African-American turnout was reduced by an average of 62% and the vote shares of opposition parties reduced by an average of 45% in these states during that period (Kousser 1974, p. 241).

Turnout decreases steadily and dramatically following the 1876 election. While this is consistent with vote destruction, it may also be part of a larger trend of decreasing turnout, due to several potential causes. Burnham (1970) argues that the election of 1900 was a critical realignment in American politics. This trend in the Deep South predates that election. Heckelman (1995) finds that turnout decreased when the secret ballot was adopted, even controlling for overall trends in turnout (Heckelman 1995, p. 115). Poll taxes and literacy tests had similar effects. These factors may explain more of the radical decline we see beginning with the election in 1892. The secret ballot is a means of suppressing vote stealing and vote inflation. Poll taxes and literacy tests, on the other hand, are essentially means of vote destruction.

Looking at the metrics for fraud, the elections of 1880 and 1884 appear fairly regular (see figure 6.3). This would support Kousser's hypothesis that Southern Democrats waited to ensure that Reconstruction was truly finished before beginning the work of disenfranchising their opposition. Beginning in 1888, we again see a spike in the distribution of relative support for the Democratic party at 100% (see Figure 6.4). This spike persists even when using the residual analysis method to control for the percentage of the population which is black. This suggests that vote suppression and not race is the underlying cause of the irregularity.

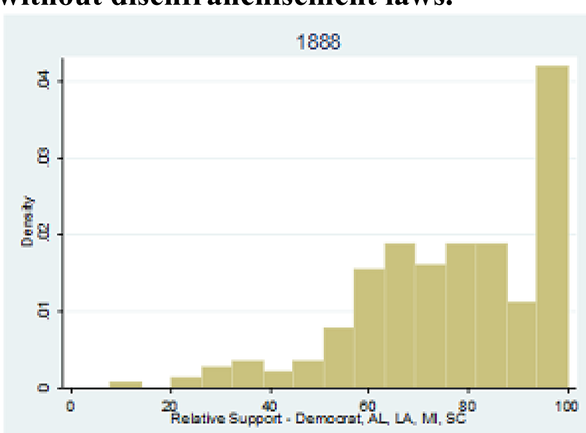
Figure 6.3. Absolute and relative support for the Democratic candidate, 1880 and 1884



In 1888, Georgia was the only Deep South state to have a poll tax (Kousser 1974, p. 239). Literacy tests had not been implemented. This lends support to Key's (1949) thesis that disenfranchisement was pervasive despite the fact that the legal framework for it was not yet in place. Figure 6.5 shows the distribution of relative support for the Democratic candidate in the four Deep South states that did not yet have poll taxes. The spike at 100% is pronounced, despite the lack of restrictive voting laws.

Figure 6.4. 1888 Presidential election. Turnout, relative and absolute support for Cleveland (D).

Figure 6.5. Relative support for Cleveland (D) in Deep South states without disenfranchisement laws.

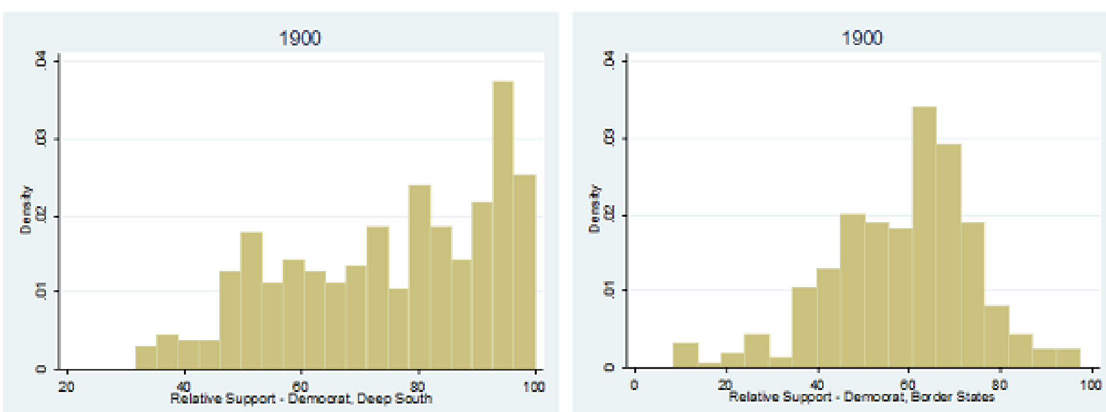


In 1890, Mississippi instituted both a poll tax and literacy tests (Kousser 1974, p. 239). Despite the fact that Mississippi and Georgia now had laws restricting suffrage, relative support for the Democratic candidate in these two states appears more normal than in those Deep South states without such laws (see Figure 6.6). This trend lends further support to Key's theory. Even in the absence of these laws, vote destruction appears significant. The irregularities in the data for Alabama, Louisiana, and South Carolina remain when controlling for race.

The irregularities in the data support the conclusion that vote destruction was a significant factor in election returns following Reconstruction. The patterns observed in the 1876 election disappeared in the elections of 1880 and 1884. They reemerge beginning in the 1888 election, when turnout shrinks steadily over time, becoming increasingly deflated. Relative support for the Democrats maintains spikes near 100%, while absolute support appears unchanged. These increases in relative support are correlated with lower levels of turnout and controlling for racial demographics. Differences between the Deep South and the other Southern states are also readily apparent in the data. In the border states, relative support for the Democratic party appears fairly normal (see Figure 6.7). In the Deep South states, however, the suspicious spike near 100% is prevalent.

Figure 6.6. Relative support (D) in Deep South states without disenfranchisement laws and with disenfranchisement laws

Figure 6.7. Relative support (D) in the Deep South and the rest of the Confederate States in 1900.



Vote fraud likely took several forms during the Redemption era. While ballot box stuffing and other vote inflation and vote stealing tactics were employed, vote destruction appears to be the strongest factor in altering election outcomes. Table 6.1 shows metrics for vote destruction over time. The patterns associated with vote destruction become increasingly apparent in the years spanning 1872 to 1904. Again, the election of 1888 reveals the beginning of strong patterns of vote suppression. This suppression occurred in advance of the widespread adoption of restrictive voting laws. The pattern gets stronger as these laws are implemented, suggesting that Key (1949) is correct in asserting that disenfranchisement took place even before it was codified into law. Given the continuing strength of these patterns, however, Kousser (1974) clearly is correct in asserting that these laws did have a serious impact.

Table 6.1. Mean and skewness for turnout and relative support (D) in the Deep South. 1872-1904.

Year	Mean – Turnout	Skewness – Turnout	Mean – Relative Support (D)	Skewness – Relative Support (D)
1872	60.68	-.23	51.135	.141
1876	70.534	-.47	67.618	-.401
1880	52.761	.320	67.63	-.48
1884	47.11	.035	67.557	-.584
1888	45.425	.218	74.984	-.658
1892	46.641	.009*	73.425	-.556
1896	36.578	.307	67.683	-.280*
1900	25.623	1.043	75.056	-.442
1904	22.175	1.369	79.061	-.845

*Symmetric due to bimodality.

Conclusion

Anecdotal evidence suggests that fraud was rampant in the South during and after Reconstruction. During Reconstruction, Democrats in the South were limited in the tools available to them, partly because they had very limited power. When Reconstruction ended, they were able to disenfranchise their opposition. This disenfranchisement took many forms, including extralegal violence and legal codification. Key (1949) argues that disenfranchisement existed through intimidation and fraud prior to legal suffrage restriction.

Using election forensics, I find that the irregularities in the data support this narrative. Following a fairly clean election in 1872, the fiasco of 1876 led to a period when Democrats were able to slowly regain power but were hesitant to engage in drastic electoral manipulation. Beginning with the election of 1888, Democrats in the Deep South were able to suppress the vote for the Republican party and begin to artificially increase their vote shares. Although this occurred before the expansion of suffrage restricting laws, the effects become even stronger once these laws spread. I therefore find support for the theories of both Key and Kousser. While disenfranchisement was pervasive even in the absence of poll taxes or literacy tests, these types of restrictions still had a serious impact.

Vote fraud in the U.S. Deep South during the Reconstruction and Redemption eras not only carries historical importance, but provides a unique case for election forensics. While vote inflation in Russia appears to be the primary means of fraud, the U.S. South during Redemption shows us a very different pattern. By employing analysis of support, we can determine that vote destruction was the primary means of electoral fraud. By contrasting the U.S. case with those of Russia and Mexico the flexibility of forensic tools becomes apparent.

CHAPTER VII

CONCLUSION

Election forensics is concerned with uncovering irregular patterns in aggregate election data. One possible explanation for these observed irregularities is election fraud. Just as in diagnosing the cause of a medical condition, several possible “diagnoses” can be made concerning symptoms observed in election data. This dissertation aims to improve the ability to determine the cause of these observed vote irregularities.

Tools for Detecting Fraud

In this dissertation I have expanded on the tools available in election forensics in two main ways. First, I have developed two simple methods for controlling for other variables. These tools allow us to rule out other causes for the symptoms of vote fraud that we see. Second, I discuss metrics that can provide insight into the type of vote fraud that may have occurred.

One major challenge in election forensics has been that of confounding variables. Often clean elections are assumed to have normally distributed turnout or to have particular digits in the vote count which follow the distribution of Benford's Law. While fraud can cause irregularities in these patterns, so can a number of other factors. Any process which alters votes in some areas and not in others could have this effect. Increased campaigning, racial demographics, income, or the degree of urbanization are all possible examples. The observed non-normality in distributions of turnout or support may be natural or legitimate rather than due to fraud.

Two methods for dealing with this issue have been put forth in this dissertation: the “jaws” method and the residual analysis method. The jaws method analyzes the relationship between absolute support and turnout. Specifically, this method examines the residuals obtained by regressing the former on the latter. I hypothesize that, in a clean election, these residuals will be normally distributed. If vote inflation is present, however, the data will be shifted in a way that causes the residuals at the high end of turnout to be positive for the beneficiary of extra votes, and negative for the victim of vote inflation. The benefit of this approach is that it requires no additional data, but is based on patterns consistent with ballot box stuffing.

The residual analysis method may provide a clearer picture when additional information is available. This method relies on regressing turnout or support on a control variable. If the residuals obtained in that regression are normally distributed in the absence of fraud, manipulation will cause a shift in these residuals. This shift is due to the fact that turnout or support is increased in some districts while the value for the control variable is not.

While these two new approaches allow us to rule out alternative explanations for irregularities, using new metrics can aid in determining the type of fraud that may drive these irregularities. Existing tools are well suited to uncover vote inflation. While they can detect vote stealing and vote destruction, other metrics can explore these cases. Irregularities in absolute support and relative support can be consistent with any of the three types of fraud. Vote manipulation will shift relative support for both candidates in much the same way it shifts turnout. In the case of vote stealing, turnout will not be affected at all, so analyzing support is critical for these cases. If the vote counts are altered for only one candidate, absolute support for the other will not be altered at all. By analyzing levels of support together with turnout, we can determine possible explanations for irregularities. Not only can we tell that fraud has occurred, but we can also gain information about the nature of that fraud.

Lastly, I adopt two formal models of the behavior of machine parties and election officials in regards to vote manipulation. These models are extensions of existing models of fraud behavior that offer contrasting views based on the goals of the decision makers. First, the macro-level model captures the behavior of national level decision makers. The party or a particular candidate may want to engage in only enough fraud to win the election in an effort to maintain legitimacy. While there is an advantage to winning an election with a solid majority, this tactic acknowledges the need to keep election returns within the realm of reason. Second, the micro-level model is taken from Chaves (2009) and looks at the decisions of local election officials. While the goal at the national level may be to maximize the probability of getting a plurality, local level officials seek to maximize the margin of victory for the machine party. They may fear backlash from the local populace, but otherwise may engage in as much fraud as possible. This results in more fraud in competitive districts, but potentially more fraud overall.

By using these new methods, election forensics can gain a much more complete understanding of the patterns in electoral data. As an emerging field, election forensics has been primarily concerned with uncovering symptoms. This project aims to help explain the root causes of those symptoms.

Three Stories of Fraud

The second part of this dissertation examined three distinct cases. Russia, Mexico, and the United States all have long histories of election fraud. Together they offer a view of electoral authoritarian regimes at every stage.

Russia has been unable to shake its issues with election fraud, even after the nation's democratic transition. It appears that Putin and United Russia have successfully developed an electoral authoritarian regime. Despite attempts at reform, extreme irregularities in the data persist. While Putin and his party would likely win elections even in the absence of fraud, local elites may have incentives to produce returns that show even stronger support. Despite the installation of webcams in every polling place, fraud has persisted.

This Russian profile contrasts sharply with Mexico. Although the *Partido Revolucionario Institucional* (PRI) maintained an electoral authoritarian regime for 70 years, the transition to democracy apparently has not been impeded by widespread fraud as in the case of Russia. The reforms in Mexico have been much more successful than those in Russia. The key difference between these two nations is the independence of election administrators. One of the most important reforms in the Mexican case made the central election authority (the *Instituto Federal Electoral*, IFE) independent of the ruling party. Where the board members were appointed by the PRI in the past, beginning in the mid 1990s the organization was made increasingly apolitical. Initially, opposition parties were given some power in determining the board's membership. As the board gained increasing independence, electoral irregularities became less common. In Russia, however, the election officials at every level are still political appointments.

The United States during the eras of Reconstruction and Redemption) offers a unique situation in terms of election fraud. While the vote manipulation that went on was certainly pervasive, it was also perversely legitimate. An analysis of the transition from Reconstruction to Redemption shows the effect of the steady and vast efforts of vote destruction. By legally disenfranchising the opposition, Democrats in the South were able to skew electoral returns to a much greater degree than any of the ballot box stuffing or vote stealing that happened in the region. This approach was made possible only when federal oversight disappeared, and federal intervention was required to reverse it.

These cases provide excellent tests of the methods developed in the first part of the dissertation. In addition, they provide a theory regarding the genesis and conclusion of regimes built on election fraud. Fraud occurs when those counting the ballots are political actors. Combating fraud requires changing the incentives for those who count the votes. Technologies like webcams, voting machines, and secret ballots have likely had some positive effect, but those who wish to manipulate vote counts have always found ways around these restrictions. Reform attempts may be more likely to succeed when they address the motive rather than the means.

Concluding Remarks

The primary goal of this project has been to make election forensic techniques more accurate and applicable. To that end, the major contribution of this dissertation is the introduction of tools to control for the presence of confounding variables. Such controls are imperative for the veracity of election forensics. In addition, I have aimed to develop a deeper understanding of fraud through the use of formal models and empirical cases. Future research should continue to refine these techniques, develop more accurate measures of fraud, and evaluate legitimacy in more cases.

The field of election forensics shows great promise. As emerging democracies continue to struggle with problems of electoral legitimacy, election forensics can provide an additional tool for election observers and policy makers. While having observers operating on the ground level is certainly important, election forensics can offer a broader view at a low cost. International pressure can have a tremendous and positive impact on electoral legitimacy. Providing a more clear picture of the degree of election fraud present in an election may encourage the development of democracy.

APPENDIX A

SIMULATION PARAMETER VALUES

Voter i's ideal point (x_i, y_i) is given by:

$b_{1x} = -2$

$b_{2x} = 4$ to simulate natural non-normality, 0 otherwise

$b_{1y} = -1$

$\sigma_x^2 = 2$

$\sigma_y^2 = 2$

Voters decide to vote if

$b_{1v} = 1$

$b_{2v} = 4$

$T = 8$

$\sigma_1^2 = 2$

$\sigma_2^2 = 2$

$\sigma_3^2 = 2$

$\sigma_4^2 = 2$

$M_1 = 3$

$M_2 = 2$

$M_3 = 4$

$M_4 = 3$

Candidate 1's x position was 16, with a y position varying between -10, -6, -4. This created elections that were relatively competitive or uncompetitive.

Candidate 2's x position was 11, with a y position of 2.

The probability that fraud occurred for Candidate 1 varied between .2, .35, and .5. The probability of fraud was allowed to differ based on whether or not the candidate received a majority of the votes cast. All permutations were simulated.

APPENDIX B

SIMULATION RESULTS

Run	Prob Fraud – Low Support	Rate of Fraud	Type of Fraud	TO Residual Skewness	TO Residual Skew Postfraud	Absolute Support Residual Skewness	Absolute Support Residual Skew post-fraud	NNN
1	0.2	0.2	Proportional Stealing	0.48	0.48	0.12	0.18	yes
2	0.2	0.2	Proportional Inflation	0.66	0.65	0.01	0.01	yes
3	0.2	0.2	Proportional Stealing	0.41	0.41	0.09	0.71	yes
4	0.2	0.2	Proportional Inflation	0.41	0.41	0.05	0.05	yes
5	0.2	0.2	Proportional Stealing	0.45	0.45	0.06	0.1	yes
6	0.2	0.2	Proportional Inflation	0.45	0.46	0.06	0.04	yes
7	0.2	0.2	Proportional Stealing	0.63	0.63	0.13	0.16	yes
8	0.2	0.2	Proportional Inflation	0.31	0.31	0.2	0.21	yes
9	0.2	0.2	Proportional Stealing	0.6	0.6	0.2	0.22	yes
10	0.2	0.2	Proportional Inflation	0.42	0.42	0.12	0.12	yes
11	0.2	0.2	Proportional Stealing	0.57	0.57	0.18	0.25	yes
12	0.2	0.2	Proportional Inflation	0.6	0.59	0.21	0.19	yes
13	0.2	0.2	Proportional Stealing	0.66	0.66	0.08	0.01	yes
14	0.2	0.2	Proportional Inflation	0.39	0.39	0.08	0.1	yes
15	0.2	0.2	Proportional Stealing	0.6	0.6	0.07	0.14	yes
16	0.2	0.2	Proportional Inflation	0.58	0.57	0.1	0.12	yes
17	0.2	0.2	Proportional Stealing	0.41	0.41	0.17	0.07	yes
18	0.2	0.2	Proportional Inflation	0.32	0.32	0.07	0.06	yes
19	0.2	0.35	Proportional Stealing	0.53	0.53	0.07	0.01	yes
20	0.2	0.35	Proportional Inflation	0.61	0.57	0.06	0	yes
21	0.2	0.35	Proportional Stealing	0.6	0.6	0.05	0.03	yes
22	0.2	0.35	Proportional Inflation	0.51	0.52	0.08	0.03	yes
23	0.2	0.35	Proportional Stealing	0.62	0.62	0.13	0.03	yes

24	0.2	0.35	Proportional Inflation	0.52	0.48	0.01	0.02	yes
25	0.2	0.35	Proportional Stealing	0.68	0.68	0.17	0.05	yes
26	0.2	0.35	Proportional Inflation	0.44	0.43	0.16	0.12	yes
27	0.2	0.35	Proportional Stealing	0.71	0.71	0.17	0.17	yes
28	0.2	0.35	Proportional Inflation	0.68	0.68	0.21	0.17	yes
29	0.2	0.35	Proportional Stealing	0.79	0.79	0.17	0.08	yes
30	0.2	0.35	Proportional Inflation	0.25	0.25	0.16	0.11	yes
31	0.2	0.35	Proportional Stealing	0.43	0.43	0.17	0.24	yes
32	0.2	0.35	Proportional Inflation	0.68	0.68	0.05	0	yes
33	0.2	0.35	Proportional Stealing	0.58	0.58	0.08	0.12	yes
34	0.2	0.35	Proportional Inflation	0.59	0.58	0.01	0.02	yes
35	0.2	0.35	Proportional Stealing	0.49	0.49	0.08	0.05	yes
36	0.2	0.35	Proportional Inflation	0.43	0.42	0.04	0.07	yes
37	0.2	0.5	Proportional Stealing	0.58	0.58	0.1	0.2	yes
38	0.2	0.5	Proportional Inflation	0.65	0.63	0.07	0.06	yes
39	0.2	0.5	Proportional Stealing	0.48	0.48	0.02	0.12	yes
40	0.2	0.5	Proportional Inflation	0.49	0.49	0.08	0.08	yes
41	0.2	0.5	Proportional Stealing	0.44	0.44	0.05	0.2	yes
42	0.2	0.5	Proportional Inflation	0.4	0.37	0.07	0.03	yes
43	0.2	0.5	Proportional Stealing	0.44	0.44	0.2	0.01	yes
44	0.2	0.5	Proportional Inflation	0.44	0.43	0.15	0.13	yes
45	0.2	0.5	Proportional Stealing	0.57	0.57	0.17	0.02	yes
46	0.2	0.5	Proportional Inflation	0.33	0.28	0.25	0.15	yes
47	0.2	0.5	Proportional Stealing	0.49	0.49	0.1	0.03	yes
48	0.2	0.5	Proportional Inflation	0.46	0.44	0.1	0.08	yes
49	0.2	0.5	Proportional Stealing	0.59	0.59	0.09	0.23	yes
50	0.2	0.5	Proportional Inflation	0.55	0.54	0.12	0.13	yes
51	0.2	0.5	Proportional Stealing	0.47	0.47	0.11	0.22	yes
52	0.2	0.5	Proportional Inflation	0.63	0.63	0.1	0.08	yes
53	0.2	0.5	Proportional Stealing	0.52	0.52	0.1	0.3	yes

54	0.2	0.5	Proportional Inflation	0.63	0.64	0.06	0.05	yes
55	0.35	0.2	Proportional Stealing	0.59	0.59	0.07	0.3	yes
56	0.35	0.2	Proportional Inflation	0.3	0.31	0.11	0.13	yes
57	0.35	0.2	Proportional Stealing	0.76	0.76	0.04	0.41	yes
58	0.35	0.2	Proportional Inflation	0.83	0.82	0.13	0.13	yes
59	0.35	0.2	Proportional Stealing	0.46	0.46	0.1	0.31	yes
60	0.35	0.2	Proportional Inflation	0.51	0.51	0.05	0.05	yes
61	0.35	0.2	Proportional Stealing	0.3	0.3	0.15	0.6	yes
62	0.35	0.2	Proportional Inflation	0.38	0.37	0.09	0.07	yes
63	0.35	0.2	Proportional Stealing	0.63	0.63	0.25	0.53	yes
64	0.35	0.2	Proportional Inflation	0.53	0.52	0.25	0.25	yes
65	0.35	0.2	Proportional Stealing	0.72	0.72	0.19	0.43	yes
66	0.35	0.2	Proportional Inflation	0.38	0.36	0.18	0.2	yes
67	0.35	0.2	Proportional Stealing	0.52	0.52	0.11	0.21	yes
68	0.35	0.2	Proportional Inflation	0.55	0.55	0.05	0.07	yes
69	0.35	0.2	Proportional Stealing	0.43	0.43	0.11	0.24	yes
70	0.35	0.2	Proportional Inflation	0.76	0.76	0.13	0.12	yes
71	0.35	0.2	Proportional Stealing	0.47	0.47	0.11	0.25	yes
72	0.35	0.2	Proportional Inflation	0.55	0.54	0.12	0.13	yes
73	0.35	0.35	Proportional Stealing	0.71	0.71	0.03	0.22	yes
74	0.35	0.35	Proportional Inflation	0.98	0.96	0.2	0.18	yes
75	0.35	0.35	Proportional Stealing	0.42	0.42	0.04	0.28	yes
76	0.35	0.35	Proportional Inflation	0.33	0.3	0.03	0.01	yes
77	0.35	0.35	Proportional Stealing	0.85	0.85	0.01	0.23	yes
78	0.35	0.35	Proportional Inflation	0.48	0.48	0.09	0.09	yes
79	0.35	0.35	Proportional Stealing	0.61	0.61	0.18	0.43	yes
80	0.35	0.35	Proportional Inflation	0.5	0.45	0.23	0.21	yes
81	0.35	0.35	Proportional Stealing	0.46	0.46	0.16	0.36	yes
82	0.35	0.35	Proportional Inflation	0.4	0.4	0.18	0.13	yes
83	0.35	0.35	Proportional Stealing	0.47	0.47	0.12	0.42	yes

84	0.35	0.35	Proportional Inflation	0.57	0.52	0.16	0.08	yes
85	0.35	0.35	Proportional Stealing	0.43	0.43	0.04	0.19	yes
86	0.35	0.35	Proportional Inflation	0.37	0.38	0.11	0.16	yes
87	0.35	0.35	Proportional Stealing	0.48	0.48	0.06	0.2	yes
88	0.35	0.35	Proportional Inflation	0.62	0.61	0.21	0.2	yes
89	0.35	0.35	Proportional Stealing	0.44	0.44	0.08	0.21	yes
90	0.35	0.35	Proportional Inflation	0.44	0.42	0.07	0.09	yes
91	0.35	0.5	Proportional Stealing	0.5	0.5	0.12	0.25	yes
92	0.35	0.5	Proportional Inflation	0.49	0.45	0.05	0.02	yes
93	0.35	0.5	Proportional Stealing	0.49	0.49	0.03	0.18	yes
94	0.35	0.5	Proportional Inflation	0.5	0.47	0.11	0.11	yes
95	0.35	0.5	Proportional Stealing	1.1	1.1	0.01	0.15	yes
96	0.35	0.5	Proportional Inflation	0.52	0.48	0.03	0.01	yes
97	0.35	0.5	Proportional Stealing	0.29	0.29	0.23	0.21	yes
98	0.35	0.5	Proportional Inflation	0.51	0.4	0.2	0.09	yes
99	0.35	0.5	Proportional Stealing	0.63	0.63	0.31	0.34	yes
100	0.35	0.5	Proportional Inflation	0.41	0.35	0.21	0.2	yes
101	0.35	0.5	Proportional Stealing	0.44	0.44	0.15	0.25	yes
102	0.35	0.5	Proportional Inflation	0.48	0.44	0.21	0.12	yes
103	0.35	0.5	Proportional Stealing	0.53	0.53	0.07	0.13	yes
104	0.35	0.5	Proportional Inflation	0.48	0.49	0.14	0.08	yes
105	0.35	0.5	Proportional Stealing	0.26	0.26	0.13	0.08	yes
106	0.35	0.5	Proportional Inflation	0.44	0.44	0.12	0.16	yes
107	0.35	0.5	Proportional Stealing	0.28	0.28	0.08	0.11	yes
108	0.35	0.5	Proportional Inflation	0.54	0.52	0.19	0.29	yes
109	0.5	0.2	Proportional Stealing	0.62	0.62	0.02	0.48	yes
110	0.5	0.2	Proportional Inflation	0.57	0.57	0.06	0.04	yes
111	0.5	0.2	Proportional Stealing	0.46	0.46	0.2	0.7	yes
112	0.5	0.2	Proportional Inflation	0.64	0.64	0.12	0.12	yes
113	0.5	0.2	Proportional Stealing	0.49	0.49	0.1	0.64	yes

114	0.5	0.2	Proportional Inflation	0.29	0.3	0.03	0.02 yes
115	0.5	0.2	Proportional Stealing	0.61	0.62	0.23	0.69 yes
116	0.5	0.2	Proportional Inflation	0.34	0.33	0.23	0.22 yes
117	0.5	0.2	Proportional Stealing	0.72	0.72	0.25	0.6 yes
118	0.5	0.2	Proportional Inflation	0.47	0.47	0.15	0.14 yes
119	0.5	0.2	Proportional Stealing	0.59	0.59	0.22	0.58 yes
120	0.5	0.2	Proportional Inflation	0.55	0.53	0.09	0.09 yes
121	0.5	0.2	Proportional Stealing	0.58	0.58	0.18	0.8 yes
122	0.5	0.2	Proportional Inflation	0.67	0.68	0.11	0.12 yes
123	0.5	0.2	Proportional Stealing	0.39	0.39	0.02	0.66 yes
124	0.5	0.2	Proportional Inflation	0.53	0.52	0.09	0.1 yes
125	0.5	0.2	Proportional Stealing	0.59	0.59	0.03	0.77 yes
126	0.5	0.2	Proportional Inflation	0.67	0.69	0.13	0.12 yes
127	0.5	0.35	Proportional Stealing	0.46	0.46	0.08	0.63 yes
128	0.5	0.35	Proportional Inflation	0.43	0.4	0.15	0.16 yes
129	0.5	0.35	Proportional Stealing	0.32	0.32	0.02	0.59 yes
130	0.5	0.35	Proportional Inflation	0.6	0.58	0.05	0.04 yes
131	0.5	0.35	Proportional Stealing	0.4	0.4	0.07	0.6 yes
132	0.5	0.35	Proportional Inflation	0.67	0.66	0.04	0.03 yes
133	0.5	0.35	Proportional Stealing	0.41	0.41	0.14	0.6 yes
134	0.5	0.35	Proportional Inflation	0.33	0.33	0.18	0.15 yes
135	0.5	0.35	Proportional Stealing	0.37	0.37	0.16	0.6 yes
136	0.5	0.35	Proportional Inflation	0.49	0.49	0.24	0.2 yes
137	0.5	0.35	Proportional Stealing	0.53	0.53	0.14	0.67 yes
138	0.5	0.35	Proportional Inflation	0.39	0.38	0.24	0.21 yes
139	0.5	0.35	Proportional Stealing	0.64	0.64	0.07	0.56 yes
140	0.5	0.35	Proportional Inflation	0.51	0.48	0.08	0.11 yes
141	0.5	0.35	Proportional Stealing	0.4	0.4	0.12	0.44 yes
142	0.5	0.35	Proportional Inflation	0.62	0.53	0.09	0.09 yes
143	0.5	0.35	Proportional Stealing	0.49	0.49	0.05	0.61 yes

144	0.5	0.35	Proportional Inflation	0.6	0.56	0.08	0.08 yes
145	0.5	0.5	Proportional Stealing	0.52	0.52	0	0.56 yes
146	0.5	0.5	Proportional Inflation	0.44	0.41	0.04	0.06 yes
147	0.5	0.5	Proportional Stealing	0.68	0.68	0.04	0.6 yes
148	0.5	0.5	Proportional Inflation	0.54	0.51	0.15	0.07 yes
149	0.5	0.5	Proportional Stealing	0.59	0.59	0.19	0.64 yes
150	0.5	0.5	Proportional Inflation	0.58	0.57	0.01	0.01 yes
151	0.5	0.5	Proportional Stealing	0.46	0.46	0.18	0.55 yes
152	0.5	0.5	Proportional Inflation	0.51	0.49	0.21	0.14 yes
153	0.5	0.5	Proportional Stealing	0.8	0.8	0.13	0.57 yes
154	0.5	0.5	Proportional Inflation	0.43	0.37	0.14	0.17 yes
155	0.5	0.5	Proportional Stealing	0.47	0.47	0.11	0.68 yes
156	0.5	0.5	Proportional Inflation	0.7	0.7	0.26	0.27 yes
157	0.5	0.5	Proportional Stealing	0.6	0.6	0.03	0.55 yes
158	0.5	0.5	Proportional Inflation	0.63	0.56	0.01	0.03 yes
159	0.5	0.5	Proportional Stealing	0.56	0.56	0.07	0.67 yes
160	0.5	0.5	Proportional Inflation	0.41	0.41	0.11	0.09 yes
161	0.5	0.5	Proportional Stealing	0.48	0.48	0.11	0.53 yes
162	0.5	0.5	Proportional Inflation	0.37	0.35	0.2	0.22 yes
163	0.2	0.2	Proportional Stealing	0.01	0.01	0.61	0.53 no
164	0.2	0.2	Proportional Stealing	0.08	0.08	0.47	0.57 no
165	0.2	0.2	Proportional Stealing	0.08	0.08	0.55	0.66 no
166	0.2	0.2	Proportional Inflation	0.16	0.88	0.39	0.22 no
167	0.2	0.2	Proportional Inflation	0.09	0.87	0.33	0.18 no
168	0.2	0.2	Proportional Inflation	0.09	1.11	0.5	0.31 no
169	0.2	0.2	Proportional Stealing	0.22	0.22	0.49	0.44 no
170	0.2	0.2	Proportional Stealing	0.18	0.18	0.4	0.42 no
171	0.2	0.2	Proportional Stealing	0.24	0.24	0.37	0.46 no
172	0.2	0.2	Proportional Inflation	0.14	0.71	0.49	0.26 no
173	0.2	0.2	Proportional Inflation	0.09	0.76	0.48	0.26 no

174	0.2	0.2 Proportional Inflation	0.21	0.95	0.36	0.2	no
175	0.2	0.2 Proportional Stealing	0.14	0.14	0.46	0.31	no
176	0.2	0.2 Proportional Stealing	0.12	0.12	0.46	0.49	no
177	0.2	0.2 Proportional Stealing	0.17	0.17	0.58	0.74	no
178	0.2	0.2 Proportional Inflation	0.09	0.55	0.51	0.26	no
179	0.2	0.2 Proportional Inflation	0.09	0.72	0.42	0.21	no
180	0.2	0.2 Proportional Inflation	0.12	0.58	0.58	0.42	no
181	0.2	0.35 Proportional Stealing	0.03	0.03	0.51	0.42	no
182	0.2	0.35 Proportional Stealing	0.09	0.09	0.45	0.43	no
183	0.2	0.35 Proportional Stealing	0.07	0.07	0.34	0.46	no
184	0.2	0.35 Proportional Inflation	0.12	1.04	0.44	0.21	no
185	0.2	0.35 Proportional Inflation	0.27	1.2	0.33	0.04	no
186	0.2	0.35 Proportional Inflation	0.13	1.17	0.37	0.13	no
187	0.2	0.35 Proportional Stealing	0.22	0.22	0.5	0.33	no
188	0.2	0.35 Proportional Stealing	0.06	0.06	0.44	0.42	no
189	0.2	0.35 Proportional Stealing	0.2	0.2	0.41	0.55	no
190	0.2	0.35 Proportional Inflation	0.11	1.11	0.42	0.03	no
191	0.2	0.35 Proportional Inflation	0.12	1.11	0.26	0.09	no
192	0.2	0.35 Proportional Inflation	0.1	0.89	0.43	0.16	no
193	0.2	0.35 Proportional Stealing	0.23	0.23	0.54	0.32	no
194	0.2	0.35 Proportional Stealing	0.11	0.11	0.36	0.38	no
195	0.2	0.35 Proportional Stealing	0.28	0.28	0.44	0.69	no
196	0.2	0.35 Proportional Inflation	0.19	0.97	0.6	0.1	no
197	0.2	0.35 Proportional Inflation	0.07	0.71	0.33	0	no
198	0.2	0.35 Proportional Inflation	0.27	0.53	0.49	0.26	no
199	0.2	0.5 Proportional Stealing	0.07	0.07	0.47	0.22	no
200	0.2	0.5 Proportional Stealing	0.11	0.11	0.36	0.35	no
201	0.2	0.5 Proportional Stealing	0.11	0.11	0.41	0.47	no
202	0.2	0.5 Proportional Inflation	0.07	1.37	0.58	0.06	no
203	0.2	0.5 Proportional Inflation	0.04	1.43	0.33	0.15	no

204	0.2	0.5	Proportional Inflation	0.02	1.49	0.68	0.01	no
205	0.2	0.5	Proportional Stealing	0.13	0.13	0.5	0.14	no
206	0.2	0.5	Proportional Stealing	0.26	0.26	0.42	0.31	no
207	0.2	0.5	Proportional Stealing	0.17	0.17	0.5	0.58	no
208	0.2	0.5	Proportional Inflation	0.18	1.19	0.4	0.16	no
209	0.2	0.5	Proportional Inflation	0.14	0.96	0.37	0.08	no
210	0.2	0.5	Proportional Inflation	0.18	0.89	0.37	0.11	no
211	0.2	0.5	Proportional Stealing	0.08	0.08	0.44	0.07	no
212	0.2	0.5	Proportional Stealing	0.13	0.13	0.37	0.32	no
213	0.2	0.5	Proportional Stealing	0.1	0.1	0.37	0.65	no
214	0.2	0.5	Proportional Inflation	0.24	1	0.36	0.22	no
215	0.2	0.5	Proportional Inflation	0.1	0.72	0.46	0.06	no
216	0.2	0.5	Proportional Inflation	0.04	0.41	0.48	0.22	no
217	0.35	0.2	Proportional Stealing	0.17	0.17	0.39	0.59	no
218	0.35	0.2	Proportional Stealing	0.21	0.21	0.37	0.61	no
219	0.35	0.2	Proportional Stealing	0.21	0.21	0.49	0.6	no
220	0.35	0.2	Proportional Inflation	0.09	0.54	0.42	0.45	no
221	0.35	0.2	Proportional Inflation	0.13	0.83	0.25	0.24	no
222	0.35	0.2	Proportional Inflation	0.2	0.94	0.54	0.37	no
223	0.35	0.2	Proportional Stealing	0.11	0.11	0.4	0.53	no
224	0.35	0.2	Proportional Stealing	0.14	0.14	0.42	0.58	no
225	0.35	0.2	Proportional Stealing	0.26	0.26	0.54	0.66	no
226	0.35	0.2	Proportional Inflation	0.15	0.55	0.51	0.31	no
227	0.35	0.2	Proportional Inflation	0.03	0.76	0.47	0.35	no
228	0.35	0.2	Proportional Inflation	0.16	0.71	0.45	0.33	no
229	0.35	0.2	Proportional Stealing	0.01	0.01	0.41	0.5	no
230	0.35	0.2	Proportional Stealing	0.22	0.22	0.35	0.51	no
231	0.35	0.2	Proportional Stealing	0.22	0.22	0.42	0.69	no
232	0.35	0.2	Proportional Inflation	0.19	0.42	0.45	0.31	no
233	0.35	0.2	Proportional Inflation	0.06	0.64	0.51	0.3	no

234	0.35	0.2	Proportional Inflation	0.1	0.6	0.46	0.33 no
235	0.35	0.35	Proportional Stealing	0.21	0.21	0.49	0.62 no
236	0.35	0.35	Proportional Stealing	0.1	0.1	0.32	0.51 no
237	0.35	0.35	Proportional Stealing	0.27	0.27	0.54	0.74 no
238	0.35	0.35	Proportional Inflation	0.24	0.8	0.45	0.38 no
239	0.35	0.35	Proportional Inflation	0.14	0.99	0.29	0.24 no
240	0.35	0.35	Proportional Inflation	0.08	1.13	0.43	0.12 no
241	0.35	0.35	Proportional Stealing	0.22	0.22	0.38	0.46 no
242	0.35	0.35	Proportional Stealing	0.01	0.01	0.32	0.52 no
243	0.35	0.35	Proportional Stealing	0.21	0.21	0.49	0.7 no
244	0.35	0.35	Proportional Inflation	0.09	0.75	0.33	0.12 no
245	0.35	0.35	Proportional Inflation	0.13	0.77	0.34	0.16 no
246	0.35	0.35	Proportional Inflation	0.16	0.68	0.39	0.06 no
247	0.35	0.35	Proportional Stealing	0.11	0.11	0.55	0.5 no
248	0.35	0.35	Proportional Stealing	0.36	0.36	0.34	0.57 no
249	0.35	0.35	Proportional Stealing	0.09	0.09	0.46	0.78 no
250	0.35	0.35	Proportional Inflation	0.09	0.6	0.37	0.11 no
251	0.35	0.35	Proportional Inflation	0.25	0.62	0.45	0.16 no
252	0.35	0.35	Proportional Inflation	0.02	0.5	0.47	0.25 no
253	0.35	0.5	Proportional Stealing	0.18	0.18	0.47	0.4 no
254	0.35	0.5	Proportional Stealing	0.08	0.08	0.36	0.46 no
255	0.35	0.5	Proportional Stealing	0.27	0.27	0.58	0.56 no
256	0.35	0.5	Proportional Inflation	0.03	0.89	0.43	0.24 no
257	0.35	0.5	Proportional Inflation	0	1.08	0.45	0.01 no
258	0.35	0.5	Proportional Inflation	0.08	1.04	0.54	0.15 no
259	0.35	0.5	Proportional Stealing	0.08	0.08	0.42	0.36 no
260	0.35	0.5	Proportional Stealing	0.05	0.05	0.45	0.54 no
261	0.35	0.5	Proportional Stealing	0.17	0.17	0.42	0.61 no
262	0.35	0.5	Proportional Inflation	0.12	0.69	0.5	0.05 no
263	0.35	0.5	Proportional Inflation	0.16	0.81	0.34	0.02 no

264	0.35	0.5	Proportional Inflation	0.01	0.86	0.43	0.06 no
265	0.35	0.5	Proportional Stealing	0.22	0.22	0.32	0.35 no
266	0.35	0.5	Proportional Stealing	0.04	0.04	0.26	0.47 no
267	0.35	0.5	Proportional Stealing	0.1	0.1	0.55	0.87 no
268	0.35	0.5	Proportional Inflation	0.12	0.62	0.52	0.03 no
269	0.35	0.5	Proportional Inflation	0.15	0.55	0.41	0.06 no
270	0.35	0.5	Proportional Inflation	0.14	0.42	0.49	0.14 no
271	0.5	0.2	Proportional Stealing	0.07	0.07	0.39	0.7 no
272	0.5	0.2	Proportional Stealing	0.27	0.27	0.46	0.58 no
273	0.5	0.2	Proportional Stealing	0.11	0.11	0.47	0.6 no
274	0.5	0.2	Proportional Inflation	0.13	0.6	0.46	0.53 no
275	0.5	0.2	Proportional Inflation	0.23	0.64	0.39	0.38 no
276	0.5	0.2	Proportional Inflation	0.22	0.99	0.53	0.3 no
277	0.5	0.2	Proportional Stealing	0.03	0.03	0.46	0.68 no
278	0.5	0.2	Proportional Stealing	0.05	0.05	0.3	0.56 no
279	0.5	0.2	Proportional Stealing	0.17	0.17	0.44	0.67 no
280	0.5	0.2	Proportional Inflation	0.1	0.5	0.53	0.45 no
281	0.5	0.2	Proportional Inflation	0.07	0.44	0.36	0.26 no
282	0.5	0.2	Proportional Inflation	0.12	0.74	0.54	0.27 no
283	0.5	0.2	Proportional Stealing	0.05	0.05	0.45	0.59 no
284	0.5	0.2	Proportional Stealing	0.11	0.11	0.49	0.65 no
285	0.5	0.2	Proportional Stealing	0.23	0.23	0.6	0.84 no
286	0.5	0.2	Proportional Inflation	0.25	0.4	0.54	0.38 no
287	0.5	0.2	Proportional Inflation	0.14	0.42	0.29	0.23 no
288	0.5	0.2	Proportional Inflation	0.11	0.46	0.43	0.32 no
289	0.5	0.35	Proportional Stealing	0.21	0.21	0.33	0.73 no
290	0.5	0.35	Proportional Stealing	0.11	0.11	0.34	0.77 no
291	0.5	0.35	Proportional Stealing	0.08	0.08	0.43	0.65 no
292	0.5	0.35	Proportional Inflation	0.14	0.47	0.46	0.4 no
293	0.5	0.35	Proportional Inflation	0.15	0.78	0.38	0.22 no

294	0.5	0.35	Proportional Inflation	0.04	1.09	0.53	0.08	no
295	0.5	0.35	Proportional Stealing	0.07	0.07	0.52	0.8	no
296	0.5	0.35	Proportional Stealing	0.15	0.15	0.39	0.62	no
297	0.5	0.35	Proportional Stealing	0.15	0.15	0.41	0.66	no
298	0.5	0.35	Proportional Inflation	0.18	0.4	0.47	0.35	no
299	0.5	0.35	Proportional Inflation	0.1	0.5	0.4	0.31	no
300	0.5	0.35	Proportional Inflation	0.24	0.68	0.48	0.02	no
301	0.5	0.35	Proportional Stealing	0.07	0.07	0.43	0.68	no
302	0.5	0.35	Proportional Stealing	0.11	0.11	0.4	0.89	no
303	0.5	0.35	Proportional Stealing	0.22	0.22	0.35	0.77	no
304	0.5	0.35	Proportional Inflation	0.06	0.31	0.55	0.35	no
305	0.5	0.35	Proportional Inflation	0.05	0.36	0.27	0.07	no
306	0.5	0.35	Proportional Inflation	0.1	0.43	0.43	0.14	no
307	0.5	0.5	Proportional Stealing	0.04	0.04	0.42	0.62	no
308	0.5	0.5	Proportional Stealing	0.01	0.01	0.25	0.73	no
309	0.5	0.5	Proportional Stealing	0	0	0.52	0.64	no
310	0.5	0.5	Proportional Inflation	0.15	0.39	0.48	0.38	no
311	0.5	0.5	Proportional Inflation	0.02	0.76	0.44	0.19	no
312	0.5	0.5	Proportional Inflation	0.15	1.08	0.58	0.12	no
313	0.5	0.5	Proportional Stealing	0.09	0.09	0.53	0.76	no
314	0.5	0.5	Proportional Stealing	0.03	0.03	0.48	0.76	no
315	0.5	0.5	Proportional Stealing	0.06	0.06	0.37	0.79	no
316	0.5	0.5	Proportional Inflation	0.2	0.51	0.4	0.17	no
317	0.5	0.5	Proportional Inflation	0.05	0.56	0.35	0.03	no
318	0.5	0.5	Proportional Inflation	0.02	0.67	0.4	0.1	no
319	0.5	0.5	Proportional Stealing	0.18	0.18	0.51	0.7	no
320	0.5	0.5	Proportional Stealing	0.04	0.04	0.27	0.71	no
321	0.5	0.5	Proportional Stealing	0.06	0.06	0.34	0.83	no
322	0.5	0.5	Proportional Inflation	0.06	0.18	0.31	0.22	no
323	0.5	0.5	Proportional Inflation	0.17	0.37	0.4	0.18	no

324	0.5	0.5	Proportional Inflation	0.1	0.36	0.53	0.11 no

APPENDIX C

METRICS OF FRAUD IN RUSSIAN ELECTIONS

Region	2007 Turnout Skew	2007 SSD Comm Party	2007 SSD UR	2008 Turnout Skew	2008 SSD Comm Party	2008 SSD UR	2011 Turnout Skew	2011 SSD Comm Party	2011 SSD UR	2012 Turnout skew	2012 SSD Comm Party	2012 SSD UR
Adygea	0.61	-0.01	0.8	0.23	-0.04	0.77	-0.05	-0.04	0.89	0.8	-0.03	0.31
Agin-Buryat	-1.59	0	0.58									
Altay	0.15	0	0.55	-0.19	-0.02	0.61	0.08	-0.04	0.79	0.26	-0.02	0.38
Altay Krai	0.42	-0.02	0.13	0.35	-0.02	0.04	0.92	-0.03	0.13	0.75	-0.03	0.14
Amur Oblast	-0.15	-0.01	0.4	-0.3	-0.03	0.28	0.59	-0.03	0.38	0.76	-0.02	0.23
Arkhangelsk	1.25	-0.01	-0.06	0.94	-0.02	-0.04	1.68	-0.01	0.1	1.66	0.01	-0.12
Astrakhan	1.08	-0.01	0.33	0.31	-0.03	0.34	0.88	-0.02	0.54	1.22	-0.02	0.28
Bashkortostan	-1.74	0.09	2.09	-1.75	0.07	1.35	-1.18	0	1.29	-0.84	0	1.03
Belgorod	-0.25	0	1.06	-0.66	0.01	1.06	-0.35	0	1.7	0.25	-0.04	1.02
Bryansk	0.41	-0.02	0.49	-0.05	-0.06	0.5	0.51	-0.04	0.97	0.51	-0.03	0.77
Buryatia	0.32	-0.01	0.43				0.29	-0.04	0.57	0.28	-0.03	0.48
Chechenya	-8.34	0	0.3	-1.42	0.03	0.63	-5.98	0	-0.32	-17.23	0.22	5.07
Chelabynsk	0.21	-0.01	0.21	0.34	-0.03	0.29	0.78	-0.02	0.8	0.59	-0.02	0.25
Chita	0.11	-0.01	0.19									
Chukchi	-0.57	0	0.42	-1	0.02	1.18	-0.77	0	0.86	-0.46	0.05	1.61
Chuvashia	0.04	-0.01	1.07	-0.5	-0.02	0.84	0.55	-0.03	0.85	-0.19	-0.04	0.68
Dagestan	-3.01	-0.01	0.3	-3.1	0.01	0.47	-2.77	0	0.16	-2.62	0.03	0.66
Ingushetia	-2.63	0	-0.11	0.05	0	0.85	0.27	0	-0.22	0.79	0	0.01
Irkutsk	0.64	-0.01	0.35	0.02	-0.04	0.34	0.86	-0.05	0.38	1.11	-0.04	0.32
Ivanovo	0.81	-0.01	0.18	0.58	-0.02	0.08	0.99	-0.04	0.4	0.96	-0.03	0.4
Jewish Autonomous	0.23	-0.02	0.54	-0.09	-0.03	0.55	0.58	-0.03	0.34	0.58	-0.03	0.3

Oblast												
Kabardino-Balkaria	-2.68	0.07	1.83	-0.13	0.05	0.75	-5.34	1.05	-0.66	0.99	-0.02	0.49
Kaliningrad	1.44	0	-0.26	1.15	-0.04	-0.06	1.52	-0.03	-0.02	1.78	0.02	-0.25
Kalmikia	0.53	-0.01	0.86	0.07	-0.03	0.78	0.21	-0.02	1.16	0.63	-0.02	0.96
Kaluga	0.47	-0.01	0.56	-0.09	-0.04	0.42	0.31	-0.04	0.7	0.67	-0.03	0.58
Kamchatka	-0.01	0	-0.02	-0.29	-0.01	-0.01	-0.04	-0.02	0.11	-0.04	-0.02	-0.01
Karachaevo-Cherkessia	-3.08	0.02	0.93	-2.26	0.07	1.21	-4.89	-0.01	0.64	-3.68	0.01	0.78
Karelia	1.25	-0.01	-0.16	1.16	-0.02	-0.09	1.74	0	-0.06	1.61	0	-0.21
Kemerov	-0.37	0	0.71	-0.88	-0.01	0.38	-0.19	-0.01	0.89	-0.54	0	0.88
Khabarovsk	0.93	-0.01	0.19	0.36	-0.03	0.17	1.12	-0.03	0.2	1.12	-0.02	0.08
Khakassia	0.75	0	0.21	0.41	-0.01	0.14	0.58	-0.03	0.4	0.72	-0.03	0.47
Khanty-Mansii	0.67	0	0.18	0.03	-0.02	0.22	0.85	-0.02	0.33	0.9	-0.02	0.46
Kirov	0.15	-0.01	0.34	-0.36	-0.02	0.42	0.76	-0.03	0.23	0.97	-0.02	0.19
Komi	0.35	-0.01	0.21	-0.21	-0.02	0.33	-0.4	-0.02	0.67	0.25	-0.02	0.45
Kostroma	0.38	-0.01	0.22	0.14	-0.03	0.13	0.69	-0.06	0.36	1.19	-0.05	0.35
Krasnodarsk Krai	0.42	-0.02	0.23	-0.93	-0.01	0.66	0.12	-0.03	0.72	0.37	-0.03	0.6
Krasnoyarsk Krai	0.49	0	0.19	0.36	-0.01	0.07	0.85	-0.04	0.36	0.78	-0.02	0.28
Kurgan	0.15	-0.01	0.28	-0.41	-0.01	0.1	0.26	-0.02	0.31	0.31	-0.02	0.3
Kursk	0.29	-0.01	0.57	-0.19	-0.04	0.41	0.48	-0.03	0.57	0.77	-0.03	0.4
Leningrad	1.59	-0.01	0.27	0.76	-0.03	0.21	1.84	-0.03	0.4	1.01	-0.02	0.37
Lipetsk	0.01	-0.01	0.63	-0.18	0	1.1	0.61	-0.04	0.93	0.61	-0.04	0.74
Magadan	0.43	-0.01	0.21	0.55	-0.02	0.22	0.74	-0.02	1.1	1.08	-0.03	0.23
Mari El	-0.45	0	1.32	-0.84	0.14	1.99	0.05	-0.01	2.07	0.29	-0.04	0.94
Mordovia	-3.35	0.14	1.96	-2.6	0.32	2.45	-3.17	0.15	2.28	-1.76	0.11	1.78
Moscow	1.05	-0.02	0.38	0.16	-0.03	0.42	1.71	-0.05	0.28	1.24	-0.04	0.34
Moscow City	2.1	-0.01	0.29	0.4	-0.03	0.41	0.93	-0.04	0.57	2.85	-0.03	0.34
Murmansk	0.65	0	-0.15	0.35	-0.02	-0.15	0.7	-0.02	0.06	0.46	-0.01	0.06
Nenets	1.02	0	0.38	1.09	-0.01	0.16	1.1	0	0.02	0.4	-0.02	0.08
Nizhegorod	0.53	-0.01	0.41	0.17	-0.05	0.32	0.27	-0.07	0.67	0.38	-0.03	0.77

North Osetia	0.37	-0.01	0.18	-0.14	-0.04	0.32	-2.01	-0.01	0.32	-1.2	-0.02	0.18
Novgorod	0.26	-0.01	0.61	0.23	-0.03	0.33	0.52	-0.03	0.6	0.5	-0.02	0.35
Novosibirsk	0.54	-0.02	0.29	0.01	-0.04	0.1	0.52	-0.06	0.42	0.78	-0.03	0.23
Omsk	-0.41	-0.02	0.56	-0.76	-0.04	0.37	0.11	-0.04	0.58	0.54	-0.04	0.38
Orel	-0.25	0	0.99	-0.8	0.02	1.02	0.41	-0.06	1.1	0.75	-0.07	0.76
Orenburg	0.38	-0.01	0.32	-0.05	-0.04	0.23	0.75	-0.04	0.37	0.82	-0.04	0.29
Penza	-0.31	0	0.9	-0.97	-0.02	0.69	-0.09	-0.02	1.13	0.29	-0.01	1.15
Perm Krai	1.13	-0.01	0.33	0.84	-0.02	0.19	1.65	-0.02	0.19	1.64	-0.01	0.12
Primorsky Krai	0.51	-0.01	0.33	-0.22	-0.02	0.65	1.07	-0.04	0.56	0.59	-0.04	0.57
Pskov	0.38	-0.02	0.57	-0.32	-0.02	0.73	0.66	-0.05	0.59	0.98	-0.03	0.4
Rostov	0.12	0	1.04	-0.2	0	0.8	0.57	-0.03	1.02	0.76	-0.03	0.75
Ryazan	0.66	-0.02	0.47	0.77	-0.05	0.34	0.86	-0.04	0.64	0.82	-0.04	0.46
Sakha Yakutia	-0.27	-0.01	0.19	-0.26	0	0.23	-0.04	-0.02	0.47	-0.15	-0.01	0.49
Sakhalin	0.37	-0.01	-0.16	0.45	-0.03	-0.05	0.48	-0.03	0.01	0.5	-0.03	0.06
Samara	0.66	-0.02	0.48	0.17	-0.04	0.33	0.78	-0.05	0.57	0.72	-0.03	0.33
Saratov	0.26	-0.01	0.74	-0.64	-0.01	0.74	-0.14	-0.01	1.02	0.29	-0.02	0.76
Smolensk	0.46	-0.02	0.52	-0.09	-0.05	0.45	0.49	-0.04	0.63	0.76	-0.04	0.43
St Petersburg City	2.56	-0.01	0.12	0.64	-0.03	0.15	1.77	-0.01	0.14	1.69	-0.02	0.37
Stavropolsky Krai	1.32	-0.01	0.22	0.37	-0.05	0.36	1.14	-0.02	0.26	0.84	-0.03	0.24
Sverdlovsk	1.36	0	0.19	1.34	-0.01	0.17	1.84	-0.02	0.15	1.72	-0.01	0.01
Tamboy	0.24	-0.03	0.51	-0.71	-0.02	0.57	-0.31	-0.02	0.55	-0.07	-0.03	0.48
Tatarstan	-1.2	0.05	2.1	-1.11	0.16	2.15	-1.3	0.04	1.58	-1.25	0.05	1.48
Tomsk	0.18	-0.01	0.44	0.32	-0.03	0.32	0.98	-0.04	0.46	1.07	-0.03	0.18
Tula	1.12	-0.01	0.37	0.08	-0.04	0.33	-0.42	-0.02	0.53	0.18	-0.03	0.42
Tuva	-0.55	0	0.53	-0.79	0.01	0.58	-1.4	0	0.84	-1.57	0.01	1
Tver	0.65	-0.01	0.44	0.17	-0.03	0.45	0.8	-0.04	0.64	1.01	-0.03	0.41
Tyumen	-0.73	0.01	1.66	-0.88	0.09	2.34	-0.73	0.03	2.1	0	0.03	1.64
Udmurtia	0.49	-0.01	0.61	0.21	-0.02	0.53	0.73	-0.03	0.68	0.76	-0.02	0.54
Uliyanovsk	-0.08	-0.01	0.76	-0.25	-0.04	0.47	0.1	-0.03	1.12	0.46	-0.05	0.64

Ust-Orda Buryat	-0.72	0	0.53									
Vladimir	0.56	0	0.04	0.66	-0.03	0.14	1.53	-0.04	0.73	1.8	0	0.02
Volgograd	0.44	-0.01	0.21	-0.05	-0.01	0.05	0.93	-0.04	0.5	0.67	-0.03	0.5
Vologda	0.41	-0.01	0.3	0.22	-0.02	0.26	1.22	-0.01	0.24	1.34	-0.02	0.24
Voronezh	-0.02	-0.01	0.8	-0.21	-0.04	0.71	-0.1	-0.02	1.22	0.39	-0.03	0.99
Yamal-Nenets	-0.44	0	0.94	-0.75	0.1	1.75	-0.22	0.01	1.6	-1.53	0.09	1.62
Yaroslav	0.75	-0.01	0.35	0.84	-0.03	0.21	1.26	-0.04	0.26	1.34	-0.03	0.23
Zabaykalsky				0.06	-0.03	0.25	0.5	-0.03	0.31	0.7	-0.02	0.23

APPENDIX D

REGRESSION RESULTS FOR URBAN/RURAL DIVIDE IN RUSSIA

2012

			Number of obs = 78		
SSD Putin	Coef.	Std. Err. t	P>t	[95% Conf. Interval]	
Population Density	-.0000252	.0000667 -0.38	0.706	-.0001581	.0001076
Constant	.5618519	.0764115 7.35	0.000	.4096652	.7140386

2011

			Number of obs = 78		
SSD United Russia	Coef.	Std. Err. t	P>t	[95% Conf.	Interval]
Population Density	-.0000209	.0000529 -0.39	0.694	-.0001262	.0000844
Constant	.6194174	.0605865 10.22	0.000	.4987488	.7400859

REFERENCES CITED

Chapter I

Beer, C. C. (2003). *Electoral competition and institutional change in Mexico*. Notre Dame, IN, University of Notre Dame Press.

Campbell, T. (2005). *Deliver the vote: A history of election fraud*. New York, Carroll & Graf.

Carothers, T. (1997). "The observers observed." *Journal of Democracy* **8**(3): 17-31.

Chaves, I. N., L. Fergusson, et al. (2009). He who counts elects: Determinants of fraud in the 1922 Colombian Presidential Election. *NBER Working Paper Series.* Cambridge, National Bureau of Economic Research**:** 36.

Cothran, D. A. (1994). *Political stability and democracy in Mexico*. Westport, CT, Praeger.

Deckert, J., M. Myagkov, et al. (2011). "Benford's Law and the Detection of Election Fraud." *Political Analysis* **19**(3): 245-268.

Fearon, J. D. (2011). "Self-enforcing democracy." *The Quarterly Journal of Economics* **126**(4): 1661-1708.

Foner, E. (1988). *Reconstruction, 1863-1877*. New York, Harper & Row.

Huntington, S. P. (1991). "Democracy's third wave." *Journal of Democracy* **2**(2): 12-34.

Hyde, S. D. (2008). How International Election Observers Detect and Deter Fraud. *Election Fraud*. R. M. Alvarez, T. E. Hall and S. D. Hyde. Washington, D.C., Brookings**:** 201-215.

Kelley, J. (2009). "The more the merrier? The effects of having multiple international election monitoring organizations." *Perspectives on Politics* **7**(1): 59-64.

Key, V. O. (1949). *Southern politics in State and Nation*. New York, Vintage Books.

Kousser, J. M. (1974). *The Shaping of Southern Politics*. New Haven and London, Yale University Press.

Lehoucq, F. E. and I. Molina (2002). *Stuffing the ballot box: fraud, electoral reform, and democratization in Costa Rica*. Cambridge, Cambridge University Press.

Magaloni, B. (2006). *Voting for autocracy.* Cambridge, Cambridge University Press.

Mebane, W. R. (2011). "Comment on “Benford's Law and the Detection of Election Fraud”." *Political Analysis* **19**(3): 269-272.

Myagkov, M. G., P. C. Ordeshook, et al. (2009). *The forensics of election fraud : Russia and Ukraine*. Cambridge ; New York, NY, Cambridge University Press.

Schedler, A. (2002). "The menu of manipulation." *Journal of Democracy* **13**(2): 35-50.

Chapter II

Beber, B. and A. Scacco (2008). What the numbers say: A digit-based test for election fraud using new data from Nigeria. *104th Annual Meeting of the American Political Science Association.* Boston, MA.

Berezkin, A., M. Myagkov, et al. (1999). "A note on the urban-rural divide in the 1996 Russian electorate and the effect of distance from urban centers." *Post-Soviet Geography and Economics* **40**(6): 395-406.

Canada Research Chair on Comparative Aboriginal Condition of Université Laval. ArcticStat Socioeconomic Circumpolar Database, www.arcticstat.org. Accessed 6/10/2012.

Cantu, F. and S. M. Saiegh (2011). "Fraudulent democracy? An analysis of Argentina's *Infamous Decade* using supervised machine learning." *Political Analysis* **19**(4): 409-433.

Deckert, J., M. Myagkov, et al. (2011). "Benford's Law and the Detection of Election Fraud." *Political Analysis* **19**(3): 245-268.

Gehlbach, S. (2000). "Shifting electoral geography in Russia's 1991 and 1996 presidential elections." *Post-Soviet Geography and Economics* **41**(5): 379-387.

Herron, E. S. (2010). "The effect of passive observation methods on Azerbaijan’s 2008 presidential election and 2009 referendum." *Electoral Studies* **29**: 417-424.

McDonald, M. P. and M. Altman. "Public Mapping Project." 2012, from http://www.publicmapping.org/.

Mebane, W. R. (2006). Election forensics: Vote counts and Benford's Law. *2006 Summer Meeting of the Political Methodology Society.* Davis, CA.

Myagkov, M., P. C. Ordeshook, et al. (2009). *The forensics of election fraud : Russia and*

Ukraine. Cambridge, Cambridge University Press.

Philpot, T. S., D. R. Shaw, et al. (2009). "Winning the Race: Black Voter Turnout in the 2008 Presidential Election." *Public Opinion Quarterly* **73**(5): 995-1022.

Chapter III

Chaves, I. N., L. Fergusson, et al. (2009). He who counts elects: Determinants of fraud in the 1922 Colombian Presidential Election.

Fearon, J. D. (2011). "Self-enforcing democracy." *The Quarterly Journal of Economics* **126**(4): 1661-1708.

Gans-Morse, J., S. Mazzuca, et al. (2010). Who gets bought? Vote buying, turnout buying, and other strategies. Cambridge, MA, Weatherhead Center for International Affairs.

Kalinin, K. and W. R. M. Jr. (2011). Understanding electoral frauds through evolution of Russian federalism: from “bargaining loyalty” to “signaling loyalty”. *2011 Annual Meeting of the Midwest Political Science Association.* Chicago, IL.

Magaloni, B. (2006). *Voting for autocracy.* Cambridge, Cambridge University Press.

Nichter, S. (2008). "Vote buying or turnout buying? Machine politics and the secret ballot." *American Political Science Review* **102**(1): 19-31.

Chapter IV

Bader, M. (2012). "Trends and patterns in electoral malpractice in post-Soviet Eurasia." *Journal of Eurasian Studies* **3**(1): 49-57.

Barry, E. and M. Schwirtz (2012). After Election, Putin Faces Challenges to Legitimacy. *The New York Times Online*. March 5, 2012.

Berezkin, A., M. Myagkov, et al. (1999). "A note on the urban-rural divide in the 1996 Russian electorate and the effect of distance from urban centers." *Post-Soviet Geography and Economics* **40**(6): 395-406.

Bratersky, A. (2012). Election Webcam Installation Begins. *The Moscow Times Online*. Moscow, Sanoma. **4807**. January 23, 2012.

Kalinin, K. and W. R. Mebane Jr. (2011). Understanding electoral frauds through evolution of Russian federalism: from “bargaining loyalty” to “signaling loyalty”. *2011 Annual Meeting of the Midwest Political Science Association*. Chicago, IL.

Kryshtanovskaya, O. G., & White, S. (2009). The Sovietization of Russian Politics. *Post-Soviet Affairs*, *25*(4), 283-309.

Lukinova, E., M. Myagkov, et al. (2011). "Metastatised fraud in Russia's 2008 Presidential election." *Europe-Asia Studies* **63**(4): 603-621.

Myagkov, M., P. C. Ordeshook, et al. (2009). *The forensics of election fraud: Russia and Ukraine*. Cambridge, Cambridge University Press.

Wilson, A. (2005). *Virtual politics: faking democracy in the post-Soviet world*. Yale University Press.

Chapter V

Alianza Cívica. (2012). Boletin de prensa. T. Camacho. México D.F. Web. http://www.alianzacivica.org.mx/archivos/pub/9557Informe%203%20de%20julio%20.pdf

Beer, C. C. (2003). *Electoral competition and institutional change in Mexico*. Notre Dame, Indiana, University of Notre Dame Press.

Cothran, D. A. (1994). *Political stability and democracy in Mexico : the "perfect dictatorship"?* Westport, Conn.: Praeger.

Eisenstadt, T. and A. Poiré (2005) Campaign Finance and Playing Field "Levelness" Issues in the Run-up to Mexico's July 2006 Presidential Election.

Frohling, O., C. Gallaher, et al. (2001). Imagining the Mexican Election. *Antipode* **33**(1): 1-16.

Hellinger, D. (2011). *Comparative politics of Latin America: democracy at last?* New York, Routledge.

Instituto Federal Electoral (2011). "Estadísticas y Resultados Electorales." Retrieved 3/5/2012, 2012, from http://www.ife.org.mx/portal/site/ifev2/Estadisticas_y_Resultados_Electorales/.

Klesner, J. L. (1997). Electoral Reform in Mexico's Hegemonic Party System: Perpetuation of Privilege or Democratic Advance? *Annual Meeting of the American Political Science Association*. Washington, D.C.

Lawson, C. (2000). "Mexico's Unfinished Transition: Democratization and Authoritarian Enclaves in Mexico." *Mexican Studies/Estudios Mexicanos* **16**(2): 267-287.

Levy, D. C. and K. Bruhn (1999). Mexico: Sustained civilian rule and the question of democracy. *Democracy in developing countries*. L. Diamond, J. Hartlyn, J. L. Linz and S. M. Lipset. Boulder, London, Lynne Rienner Publishers**:** 519-574.

Magaloni, B. (2005). The demise of Mexico's one-party dominant regime: elite choices and the masses in the establishment of democracy. *The Third Wave Of Democratization In Latin America: Advances And Setbacks*. F. Hagopian and S. Mainwaring. Cambridge, Cambridge University Press**:** 121-148.

Magaloni, B. (2006). *Voting for autocracy*. Cambridge, Cambridge University Press.

Miroff, N. and W. Booth (2012). Mexico's presidential election tainted by claims of vote buying. *The Washington Post*. July 4, 2012. Web.

Peeler, J. A. (2004). *Building democracy in Latin America/ John Peeler*. Boulder, Colo., Lynne Rienner Publishers.

Thompson, G. (2004) Ex-President in Mexico Casts New Light on Rigged 1988 Election. *The New York Times*

Chapter VI

Argersinger, P. H. (1985-1986). "New perspectives on election fraud in the gilded age." *Political Science Quarterly* **100**(4): 669-687.

Burnham, Walter D. (1970). *Critical Elections and the Mainsprings of American Politics*. New York: Norton, 1970.

Campbell, T. (2005). *Deliver the vote*. New York, Carroll & Graf.

Clubb, Jerome M., William H. Flanigan, and Nancy H. Zingale. Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972. ICPSR08611-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-11-13.

Cox, G. W. and J. M. Kousser (1981). "Turnout and rural corruption: New York as a test case." *American Journal of Political Science* **25**(4): 646-663.

Foner, E. (1988). Reconstruction: America's Unfinished Revolution. New York, Harper and Row.

Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-

2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. doi:10.3886/ICPSR02896.v3

Heckelman, J. C. (1995). "The effect of the secret ballot on voter turnout rates." *Public Choice* **82**: 107-124.

Key, V. O. (1949). *Southern politics in State and Nation*. New York, Vintage Books.

Kousser, J. M. (1974). *The Shaping of Southern Politics*. New Haven and London, Yale University Press.

Ruggles, Steven, and Russell R. Menard. Census of Population, 1880 [United States]: Public Use Sample. ICPSR06460-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1995.

Terry, W. C. (2013). Political competition and African-American lynching in the U.S. South, 1882-1930.