Electoral Crime Under Democracy: Evidence from Brazil

Andre Assumpcao[*]

June 1st, 2019

Abstract

This paper presents the first analysis of the relationship between electoral crimes and electoral performance in large democracies. Using a sample of candidates charged with electoral crimes in the race to municipal office in Brazil between 2004 and 2016 and employing an instrumental variables strategy, I find that being ultimately convicted by the Brazilian Electoral Court reduces the probability of election by 23.1 percentage points and convicted candidates' vote share by 13.3 percentage points. These results are robust to different estimation strategies and are not explained by changes in voter nor candidate behavior once an unfavorable ruling is made public. I further estimate the electoral payoff when crimes are not detected and whether voters impose differential punishment for substantial or procedural rule-breaking; while there is a significant gain in the probability of election and vote share (4.7 and 4.9 percentage points respectively) when candidates deploy substantial illegal tactics, voters impose the same electoral penalty regardless of candidates' charges. This result explains why candidates and parties would still employ illegal tactics while risking detection by the Brazilian Electoral Court.

Keywords: electoral politics; judicial politics; comparative politics; illegal behavior and the enforcement of law; political economy.

JEL classification: D72; K42; P48.

---

[*]Ph.D. Candidate, Department of Public Policy, The University of North Carolina at Chapel Hill. Contact details: aassumpcao@unc.edu.

1 Introduction

In democratic regimes, office-seeking politicians employ various tactics to get elected. They might promise voters more resources to increase the provision of local public goods, such as schools, hospitals, or roads; they can promote their candidacies by running ads on TV and, more recently, on social media; they might even meet with their constituents and ask for their vote based on their personal connection. While these tactics are different, sometimes complementary ways to win an election, they all characterize *play-by-the-rules* strategies, in which individuals follow legal provisions when running for office. Governments allow such electoral practices because they make electoral systems more competitive, and increase access to political office to more groups representing citizens. In this paper, however, I focus on illegal tactics to win elections and produce the first analysis of the effect of electoral crimes on ballot performance in large democratic regimes.

Scholars have not ignored these mechanisms used for winning elections. Lehoucq (2003) offers a comprehensive account of electoral fraud, which takes up a variety of forms, such as procedural rule-breaking, illegal campaigning, violence, and even unequivocal vote buying. In a more recent study, Gans-Morse et al. (2013) design a theoretical framework encompassing four types of clientelism practices (vote, turnout, and abstention buying, and double persuasion) and their adoption under five different institutional designs. They argue that the choice of illegal action is conditional on the design of electoral systems. For instance, in an environment of increased political polarization, we should expect to see more of turnout buying but less of vote buying.

Indeed, most studies looking into illegal electoral tactics have two common characteristics: first, they are largely concerned with coercive threats that prevent free and fair elections, as suggested by Mares and Young (2016); second, they focus heavily on non or partially democratic regimes, evidenced by the vast literature on electoral authoritarianism (Levitsky and Way, 2002; Gandhi and Lust-Okar, 2009; Ichino and Schündeln, 2012; Schedler, 2015; Asunka et al., 2017). This is a rich literature that helps understanding the use of elections for regime consolidation and continuity. Nevertheless, I address two unexplored issues that are supplemental to the established literature investigating electoral fraud.

The first contribution here is uncovering the effect of electoral crimes that are harder to detect or whose relationship with electoral outcomes is less known or well understood. For instance, politicians might use illegal forms of advertising or slush funds to spend beyond their campaign limits in order to win an election. Likewise, candidates and political parties might put forth candidacies for public office even if they do not meet all electoral requirements, a particular feature of Brazilian electoral law,[1] as a means of raising their profiles for future elections. These strategic moves are less easily understood than flat out vote buying, for example. The second contribution is precisely understanding how electoral crimes shape electoral outcomes in large democracies. Brazil is one of the top five largest democracies in the world and as such is an important research setting for understanding the use of illegal electoral tactics. Beyond just size, and despite a recent fallback, the

---

[1]For instance, parties need to file financial records proving they are financially solid and candidates must not have been convicted of crimes at the appellate level either at the state and federal judicial systems.

quality of Brazil's democracy makes it an important case study: since 2006, the country consistently ranks in the top 20 percent countries in the V-Dem Electoral Democracy Index (Coppedge et al., 2018).

Besides the electoral fraud scholarship, the present study contributes to the broader literature of political economy of development. Brazil has an unique institutional design in which the judiciary branch has an entire system of State (TRE) and Federal (TSE) electoral courts resolving electoral claims. Their mandate is to guarantee free and fair competition for public office, enforcing the Brazilian Electoral Code of 1965 and subsequent legislation, and to prevent that candidates not meeting legal requirements join electoral races. To the extent that the Electoral Courts are successful in rooting out this type of wrongdoing, we should expect more electoral accountability from office-holding politicians. Candidates would avoid illegal tactics to preserve their future career prospects. Understanding if electoral systems as such are effective should provide an important takeaway for countries sharing the same institutional design. Mexico, South Africa, and India are but a few of developing countries which also have a dedicated electoral authority similar to Brazil's. In addition, this paper investigates another source of judiciary power beyond settling legal disputes between economic and political agents; since every political candidate in Brazil needs a judicial authorization to run for office, the Electoral Court holds an enormous amount of power in shaping up political representation – an unusual role played by judiciary branches.

Another important contribution in this study is the use of court documents as data. I collect and code judicial rulings from TRE and TSE courts on candidacies for municipal office in Brazilian elections between 2004 and 2016. For a subset of these documents, I implement a support-vector machine classification algorithm to find the exact allegations against candidates that prevent them from running for office. I split such allegations in two categories, substantial and procedural rule-breaking, to identify heterogeneous effects of electoral crime on performance for the 2012 and 2016 elections. This project forms part of a recent wave of studies using court documents to measure economic and political outcomes in development settings (Sanchez-Martinez, 2018; Lambais and Sigstad, 2018).

Using these court documents, I recover the causal effect of electoral crimes adopting an instrumental variables (IV) strategy. Since the judicial ruling on candidacies is composed of up to three instances of review, I can instrument decisions at the trial stage (first instance) for appellate decisions (second or third instance) for a subset of candidates who have untried appeals by election day. In such cases, the Brazilian electoral code establishes that candidates can be voted for, and that their votes should be counted, regardless of the charges brought against their candidacy in the first place. While the trial ruling is endogenous, e.g. potentially correlated with other factors determining electoral chances, appellate rulings issued *after* election day cannot influence electoral outcomes beyond their connection with trial decisions. Thus, for this subsample of candidates running for office who have an untried appeal standing at the time of election, I can identify the causal effect of crimes on performance.

The main IV result shows that a conviction for electoral crime reduces the probability of election

and a candidate's vote share by 23.1 and 13.3 percentage points, respectively. These estimates are statistically significant at the one percent level and significantly differ from OLS point estimates. These results are robust to the inclusion of covariates and fixed-effects, coefficient stability tests (Oster, 2019), Monte Carlo simulations of IV parameters, and reverse causality checks. Unauthorized candidates are also significantly further away from the election threshold in both proportional (city council) and majoritarian (mayor) systems, but this effect is indistinguishable from OLS estimates. These results indicate that an unfavorable candidacy ruling negatively impacts a candidate's chances but the actual effect size is only consistent in the probability of election and vote share estimates. Though we should be careful when comparing these results with studies looking at punishment for corruption in Brazil (Ferraz and Finan, 2008, 2011; Winters and Weitz-Shapiro, 2013), which is a more severe crime prosecuted by various other legal authorities and judicial bodies, the evidence here points to the same negative impact of (detected) illegal behavior on electoral performance.

I proceed further and test whether such illegal tactics, when undetected, have any electoral payoff. I find that candidates accused of a substantial breach of electoral law (e.g., candidates or parties have used illegal campaign strategies, channeled slush funds for campaign ads, have previous outstanding judicial convictions preventing them from running for office) increase their probability of election and vote share by about 4.8 percentage points compared to candidates accused of procedural rule-breaking. This result supports a positive electoral effect of adopting such campaign strategies. Finally, I test whether voters impose heterogeneous penalties in cases of substantial or procedural rule-breaking. I find, however, that voters are not sophisticated and do not differentiate them crime charges. Taken together, these results are an indication that candidates might risk punishment in exchange for the electoral benefit that is realized if they are not detected by the Brazilian Electoral Court.

In the remainder of this paper, I explain the institutional background allowing for causal identification in section 2, present the data in section 3, and discuss the theoretical mechanism underlying the relationship between electoral crimes and performance in section 4. Section 5 discusses the empirical strategy and section 6 presents the main results. Section 7 explores exclusion restriction tests. In section 8, I investigate alternative explanations for the effect of electoral crimes on performance coming from changes in the behavior of voters, candidates, and judges. Section 9 discusses heterogeneous punishment effects. Section 10 concludes and suggests further avenues of research.

2   Institutional Background

The Brazilian Federal (TSE) and State Electoral Court (TRE) systems have existed intermittently since 1932 but only became institutionally relevant after the country's return to democracy in 1985. Since then, electoral courts have a fundamental role in guaranteeing free and fair elections. Their mandate is to enforce the Electoral Code of 1965 and subsequent legislation, particularly the law establishing conditions for ineligibility to public office (1990), the Law of Political Parties (1995),

the Law of Elections (1997), and the Clean Slate Act of 2010.

These courts are have four main responsibilities: (i) electoral rule-making; (ii) judicial consultations clarifying and establishing jurisprudence for conflicting electoral norms; (iii) administration of the electoral process, which consists in publishing the electoral calendar, testing voting machines, distributing voting machines to all districts, counting and publishing electoral results; and, finally, (iv) conflict resolution on claims of breach of electoral law.

In this project, I am mostly interested in the courts' conflict resolution function and its underlying judicial review process. According to Brazilian Law, every individual running for office, at every level, has to submit proper documentation proving that they meet eligibility requirements for the office they are running; for instance, they should be 35 years of age or older to run for president or senator; executive-office holders, if running for any other elected office, must step down from their current post six months before election day. Every electoral cycle, the highest-level electoral court, TSE, establishes a schedule for submission of all these documents, which are reviewed at lower-level courts by electoral judges who issue rulings authorizing every single candidacy in the country. This is the main institutional feature that allows for causal identification of electoral crimes on performance.

An example helps illustrate this point. The most recent municipal elections took place on October 2, 2016. The deadline for submitting all candidacy documents was August 15, 2016. Between August 15 and September 12, electoral courts reviewed and authorized each candidacy for mayor or city councilor. The review process started at the electoral district in which the candidate is running for office, and their trial ruling comes out of the designated electoral judge for that district. These judges are part of the state court system and, when appointed to the electoral bench, are on leave from their original tenured positions at the state system.[2] They serve on two-year mandates, with one reappointment allowed, such that they never oversee the same district for more than one electoral cycle. If either a candidate or someone else, such as opponents or the Office of Electoral Prosecutions (MPE), files an appeal to the trial ruling, the case is presented before a panel of three judges at the state electoral court TRE. There are seven appellate court justices in each state's TRE, serving up to four-year mandates, and they are immune to local politics. In any state, six of these judges are voted in by their fellow tenured judges at the state and federal court systems and the last member is appointed by the President of Brazil. If plaintiffs or defendants are unhappy with the appellate court decision, they can appeal their case before the federal court TSE, which serves as the third and final instance of judicial review for mayor and city councilor candidates.

The September 12[3] date is the key institutional feature that allows observing performance for politicians who violate electoral rules. It is the last day for entering candidate information onto electronic voting (EV) machines distributed at every single polling station in the country.[4] All

---

[2]In Brazil, judges are appointed to the bench in state and federal courts when they are approved in nationally-competitive entrance examinations. They are automatically tenured after a two-year trial period, therefore their entire career is independent of electoral politics.

[3]The exact day varies marginally every cycle. In 2018, for instance, deadline for candidacy submission was Aug 15, last day for loading candidate information was Sep 17, and election day was Oct 7.

[4]Fujiwara (2015) describes this technology in detail.

candidates who have untried appeals by this date will have their information loaded, and thus can be voted for, in the EV machines on election day. Because of this feature, I can observe the electoral performance of candidates who eventually are convicted of electoral crimes and compare to candidates who are cleared of similar charges. If candidates have had a final ruling before September 12, or if they have decided not to appeal their trial sentence, I cannot observed their performance because TSE will not include their information in the EV machines.

Exogenous variation in convictions for electoral crimes comes from the timing according to which appeals sentences are issued by higher-level courts. Often, the high number of candidates running for municipal office, the judicial backlog, or the conditions of a particular electoral race make it difficult for electoral courts to hand out final decisions by Sep 12. Moreover, since candidates with outstanding appeals will have their information loaded onto EV machines regardless, there is no strong incentive for courts to issue decisions between then and Oct 2. In the lead-up to election day, judges and court officials are working around the clock making sure that 540,000+ EV machines are delivered to 450,000+ voting stations across the whole territory of Brazil; judges are ruling on smaller electoral cases that might or might not be appended to candidacy cases; court officials are meeting with political parties and discussing the local electoral situation, so on and so forth. It is not uncommon, therefore, that final decisions are issued only after election day has passed, specially in municipal elections, when stakes are lower than in federal elections.

When candidate appeals are not ruled in time for elections, candidate information (picture, name, voting number) is displayed in the EV machine but their votes are computed *sub judice* – their vote count will be considered valid only when the TRE or TSE publish their final decision on any individual's candidacy. Effectively, thus, the decisions at the appeals stage cannot affect electoral outcomes, since they are issued only after election day has passed, but they bear a strong relationship to the sentence handed out by the trial judge in each electoral district. Decisions at trial are mostly endogenous to electoral outcomes, but the use of appeals as instruments leaves out only their exogenous part – allowing for causal identification.

The primary limitation of this study is that I can only recover causal effects of electoral crimes under restrictive conditions pertaining to municipal elections in Brazil. At any other electoral race, both the trial and appeals stages are handled by the TRE and this might shape the way electoral judges issue rulings in response to the importance of the office for which a candidacy has been presented. For instance, senators are much more influential than city councilors and have a direct channel of communication with the President of Brazil, who is responsible for appointing one judge per TRE. Second, there are a number of candidates who do not appeal their trial ruling and as such do not appear on the EV machine on election day. Thus, I cannot observe their ballot performance. It is likely that these candidates are heterogeneous in many dimensions when compared to candidates who have outstanding appeals, such as their political experience, or their drive to hold elected office. These candidates should be the object of future projects measuring the effect of electoral crime on electoral performance in developing countries, and this paper inaugurates such literature.

3 Data

The main data source for electoral performance is TSE's repository of electoral data. TSE publishes electoral results, vote counts, candidate's individual characteristics, and their candidacy's situation on election day for all elections since 1994. I focus on the municipal elections after the introduction of the EV machine in 2002 for even performance measures across elections and municipalities. My sample is composed of 9,470 candidates for mayor or city councilor who appealed, or had third-parties appealing, the trial ruling on their candidacy authorization. These candidates have been displayed in the EV device and could have been voted for on election day. Their candidacy remained pending after elections and they have only been allowed to take up office once a final ruling was issued. I create three outcome measures from TSE's data: (1) *the probability of election*, which is a binary variable taking up value one when the candidate received enough votes to be elected. For mayor candidates, under majoritarian rule, this means 50 percent plus one of all valid votes. For city council candidates, under proportional rule, this means having received enough votes to rank amongst the most voted candidates within the designated number of vacancies for each municipality; (2) *vote share* as a share of total valid votes; (3) *vote distance to election cutoff*, which is the percentage point distance between a candidate's vote share and the votes necessary for election. Outcomes (1) and (2) are make or break measures of electoral crime: we can use them to estimate whether a convicted candidate is predicted to win or lose an election; conditional on having won (or lost) an election, outcome (3) describes the relative safety (or damage) resulting from employing certain electoral tactic.

I scrape court documents containing the allegations against each candidate from the TSE website, which makes all their rulings public. I have developed a Python program[5] that downloads case file and sentences for all candidates in my sample. Though the information is public, due to data maintenance limitations at the TSE, 99.5 percent of court documents come from candidates in the 2012 and 2016 municipal elections. I match court documents to candidates using an individual identifier provided by the Electoral Court so that I can recover all documents for each candidacy.

Table 1 reports the summary statistics of the sample. The average age is 46.3 years, and the overwhelming majority of candidates facing judicial challenges is male. Nine percent of them have any political experience, captured by whether they held any other elected office in the past. These candidates have reported, on average, campaign spending amounting to R$ 52,555. Using the current exchange rate, this is equivalent to $15,000 per campaign. Sixty-four percent have seen an unfavorable ruling from the trial judge at their electoral district and 53.7 percent have had an unfavorable ruling after appealing their case to higher courts. Notice that all candidates have seen charges brought against them at trial, otherwise they would not have outstanding appeals by election day and would not be part of this sample; the conviction variables here, however, capture unfavorable decisions issued by trial judges. If an electoral judge allowed a candidate to run for office, then either the trial or appeals variables become zero. Though not reported in table

---

[5]For the benefit of research transparency and replication, all programs and analysis scripts are freely available online on GitHub.

1, I also collect information on candidates marital status and education.[6] These are categorical variables, and the most frequent marital status is married (62.6 percent) and education level is high school (30.8 percent). Finally, I report the means for the three outcomes in this analysis. The probability of election is 19.1, while candidate's vote distance to cutoff and vote share are -4.1 and 10.1 percentage points.

4   Theory

Assume there are three representative agents interacting in an election: voter A, candidate B, and judge J. They each have their utility function $f(X_c, \varepsilon_c)$ represented by a matrix of observed candidate characteristics X_c and unobserved characteristics ε_c. The former could be anything from policy positions, age, ethnicity, marital status, or campaign expenditures. A candidate's political ability, the deals they make with parties, supporters, or sponsors are the latter. In this setting, candidates' information is essential for determining the outcome of the election: A chooses candidates that maximize their utility $U_A = f_A(X_c, \varepsilon_c)$. For instance, A might prefer politically-aligned, more educated candidates because they believe these to be the most prepared candidates to take up office. Most importantly, A dislikes candidates who have a criminal record because it signals dishonesty and poor prospective political performance. This is made explicit by separating out criminal charges c_c from matrix X_c in A's utility function and setting the first derivative of f_A with respect to c_c to negative, as follows in equations (1) and (2):

$$U_A = f_A(X_c, c_c, \varepsilon_c) \qquad (1) \qquad\qquad \frac{\partial U_A}{\partial c_c} < 0 \qquad (2)$$

In addition to the representative voter's preferences, I am also interested in candidate's B preferences. B derives utility from holding elected office and is looking to adopt strategies that will help their electoral endeavor. They cannot withhold or control certain characteristics, such as age, gender, ethnicity, but can choose amongst campaign expenditure levels (included in X_c) and electoral strategies that will get them closer to winning an election. Some beneficial strategies are legal while others are not (b_c), and B will adopt a mix of strategies such that the expected electoral payoff remains positive – there is a strictly positive risk of illegal strategy detection by judge J, and candidate B chooses strategies before judges authorize each political campaign.[7]

$$U_B = f_B(X_c, c_c, b_c, \varepsilon_c) \qquad (3)$$

The focus in this paper is identifying how criminal records and illegal strategies, respectively summarized by c_c and b_c, impact a candidate's chances of election. More specifically, I am looking

---

[6]I also have information on each candidate's party and use it as fixed-effects in the empirical sections.

[7]I assume a uniform distribution of detection risk across electoral districts for the reasons laid out in section 2, guaranteeing the independence and quality of electoral judges in Brazil. Note that applications of this simple model to other jurisdictions would likely change this assumption to adjust to features of other judiciary systems.

at the effect of convictions for electoral crimes on electoral performance. The reasoning is straight-forward. Voters dislike crimes and are likely to punish candidates who are found guilty of electoral code violations in the run-up to an election. This mechanism would predict a negative first derivative for c_c in equation $\partial U_B / \partial c_c < 0$: convictions on record hurt a candidate's chances. Some of this effect, however, could be offset by the boost in votes that would come from illegal strategies b_c. Suppose a candidate prints and distributes negative material on their opponents. The information in the advertisement is false, and such ad is not allowed in the jurisdiction where this office race is taking place. Though judges could eventually ban such material, once the information is out it might hurt targeted opponents beyond reparation. In such hypothetical scenario, the strategy was illegal but benefited the candidate running the ad. Thus, I want to identify both effects on electoral performance.

4.1 Application to the Brazilian Context

The majority of the literature in electoral crimes is concerned with more severe actions, such as fraud or vote buying (Lehoucq, 2003). This paper is an important contribution to the scholarship by looking at other, more common, and more nuanced violations to electoral rules that are particular to large democracies, such as Brazil.

As discussed in section 2, Brazil is a particularly interesting research setting because of the structure of its dedicated Electoral Court System. All candidates need an authorization from the electoral judge in their district to run for office. This decision comes out as a judicial sentence, just like in any other legal case. The responsible judge verifies a candidate's application and issues an authorization based on compliance with electoral code. Judges check whether a candidate's party has met all electoral requirements, whether candidates have met all criteria of the office to which they are running, and other legal provisions as established by electoral law.

By coding statues and judicial sentences, I create two sets of reasons why someone is prevented from running for office: (1) *procedural* rule-breaking are cases in which candidates have missing documentation or other trivial action to run for office. For instance, they could have forgotten to include a copy of their ID card in their application or they could have missed a deadline in the application process; (2) *substantial* rule-breaking, which are more severe cases in which either parties or candidates are in breach of more substantial elements of electoral law. Their party might have not kept all financial records from previous elections, they might have an outstanding conviction on appeals for previous crimes, or they might have been convicted for running illegal campaign strategies against opponents. *Substantial* cases are much more likely to be connected to campaign, office, or government crimes disliked by voters, and thus are the group purposely using illegal strategies b_c. I expect the relationship between b_c and electoral performance to be positive in equation $\partial U_B / \partial b_c > 0$.

Moreover, these candidacy cases (called *registro de candidatura*, or candidacy registration in free translation) have standard sentences and penalties. Judges dismiss candidacies when they do not meet all requirements. There is no jail time nor immediate financial penalties for candidates

8

and parties. Though this might not be ideal from a policymaking perspective, it creates a subset of legal cases less susceptible to external influence and relative stability in terms the application of legal statues and convictions.[8]

The final hypothesis I test is whether there are heterogeneous treatment effects by conviction type. If voters are sophisticated, not only they punish candidates with unfavorable trial rulings (*the conviction effect*) but they also differentiate the punishment conditional on the crime (*the sentence type effect*). One can reasonably expect that candidates charged with more severe crimes, such as illegal campaign spending, or convicted for previous crimes, signal a more systematic criminal behavior and should be punished more harshly than candidates missing deadlines or lacking hard copies of certain documents. Though the judicial punishment is the same, the electoral punishment could still reflect the relatively more severe violations. There is substantial evidence in the literature against voter sophistication in other information contexts (Avis et al., 2018; Banerjee et al., 2010; Chong et al., 2015; de Figueiredo et al., 2011; Ferraz and Finan, 2011; Weitz-Shapiro and Winters, 2017; Winters and Weitz-Shapiro, 2013); this paper explores yet another mechanism of providing information to voters (judicial decisions) and investigates how voters react to it.

5   Identification Strategy

In this paper, I adopt an instrumental variables (IV) approach that allows the causal identification of the effect of electoral crime on performance. As described in section 2, I can only recover local average treatment effects (LATE) for the subsample of candidates who are charged with electoral crimes, by the trial electoral judge or third-parties, and have an outstanding appeal on their trial decision by election date. Candidates who break electoral code but are not detected are not part of this study, neither are candidates who have chosen not to appeal their trial sentence. For this sample of candidates with untried appeals on election day, I estimate the following regression model in three ways and using three different measures of electoral performance:

$$y_i = \alpha + \rho \cdot c_i + X\beta + \sum \lambda_{i,k} + \varepsilon_i \tag{4}$$

The dependent variable y_i forms are: (1) the probability of election, taking up value one when either the mayor or city council candidate had enough votes for election in their district; (2) the total vote share of candidate i received in their race; (3) the vote distance to the election cutoff, which is the percentage point margin between candidate i's vote share and that of the single elected candidate (when running for mayor) or last elected candidate (when running for city council). Using outcome (1), I can measure the impact of crime on the most important outcome of any political

---

[8]These cases, however, are often appended to other cases at the electoral court system and can create financial liabilities for candidates and their parties. The analysis of these cases is beyond the scope of this study as they do not meet the criteria for causal identification developed here. There is also growing interest for electoral court reform in Brazil. Some experts criticize the fact that justices do not have fixed appointments and thus do not specialize in electoral crimes; others say they that parties and candidates strategically game harsher punishments by requesting other court systems to move charges to electoral courts.

campaign, i.e. being elected; outcome (2) serves as a measure of the overall impact of crime on candidate popularity if they are found guilty by the electoral judge; the last outcome (3) tells us about the relative benefit of employing an illegal tactic when candidates are trying to secure an electoral lead or narrow in on races in which they are trailing another candidate; X is the matrix of candidate characteristics, such as candidate age, gender, marital and education status, political experience, and campaign espending; $\sum \lambda_{i,k}$ is a set of k fixed-effects to capture any additional unobservable heterogeneity coming from party, election, and municipal factors shared by subsets of candidates.

The main independent variable is the binary indicator for convictions for electoral crime c_c at the electoral court system for candidate i. If a candidacy has been rejected by the trial judge responsible for that electoral district, c becomes one. I use convictions at trial in OLS regressions for benchmarking the effect on electoral performance; in reduced-form regressions, I replace convictions at trial for convictions on appeal – which becomes one when the candidate has seen an unfavorable ruling at higher courts within the electoral system. The reduced-form regressions hint at any potential correlation between instruments and outcomes beyond the channel via the endogenous decision at trial. I lastly estimate model (4) using two-stage least squares (2SLS) regressions, in which I instrument convictions at trial for convictions on appeal. Since I am looking at appellate court decisions issued after election day, the exclusion restriction is straightforward as the instrument is measured only *after* outcomes have been observed.[9] Any effect of appellate decisions only influences electoral performance via their relationship with convictions at trial. I address additional concerns about violations to the exclusion restriction in the following sections, but the baseline instrumental variables (4) and the first-stage regression equations are:

$$y_i = \alpha + \rho \cdot \hat{c}_{i,\text{trial}} + X\beta + \sum \lambda_{i,k} + \varepsilon_i \tag{5}$$

$$c_{i,\text{trial}} = \alpha + \rho \cdot c_{i,\text{appeals}} + X\beta + \sum \lambda_{i,k} + \varepsilon_i \tag{6}$$

For every specification of equations (5) and (6), I estimate versions excluding and including individual characteristics (matrix X) and fixed-effects $\sum \lambda_{i,k} + \varepsilon_i$. In addition to instrument validity tests, I also report coefficient stability tests across different specifications to demonstrate that selection on unobservables is not driving coefficient estimates, as discussed in Altonji et al. (2005); Nunn and Wantchekon (2011); Oster (2019); Pei et al. (2019). I discuss and test other alternative, confounding explanations in following sections and provide the empirical strategy at each stage of analysis.

---

[9]In addition to the temporal effect, the other theoretical arguments discussed in section 2 support the exogeneity of the instrument. Electoral judges are tenured state judges which have no ties to local politicians. Their wages, career prospects, and time on electoral bench are all independent of the action of mayors and city councilors.

5.1 Inclusion Restriction Checks

The first step in this analysis is guaranteeing I have a strong instrument for the endogenous regression of interest (conviction at trial). Table 2 provides us anecdotal evidence on the relationship between convictions at either stage of the judicial review process. The overall reversal rate of trial decisions is 10.9 percent. Reversals come mostly from candidacy cases that had been denied by trial judges (16.6 percent). The unconditional Pearson correlation coefficient between convictions at trial and on appeals is .79. These results make intuitive sense given the presumed quality of judges and standard sentencing (both in format and content) discussed in previous sections.

A more robust test, however, is reported in table 3. I present three first-stage regressions on the relationship between the endogenous variable convictions at trial and instrument convictions on appeal, which are the *post*-election decisions issued by higher courts in the electoral system. Across models progressively including candidate characteristics and municipal, electoral, and party fixed-effects, the coefficient on the instrument is statistically significant (p-value $< .01$). The magnitude remains stable within the .738-.766 range, which means that a conviction on appeals explains three quarters of the outcome at trial. The positive relationship confirms the anecdotal evidence in table 2.

I additionally report each coefficient point estimate, confidence intervals (CIs), and F-statistics for all three regressions in figure 1. The inclusion of covariates and fixed-effects across models marginally shifts down the magnitude of instrument points estimates. In all cases, however, the F-statistic of excluded instruments remains greater than industry standards at $F = 10$ (Bound et al., 1995). This means that the first-stage model is significantly predicting the candidacy outcome at trial and can confidently be used for partialling out the causal effect of convictions on electoral performance.

Finally, in table 4, I present the Hausman tests for OLS consistency. I report the results for bivariate regressions between convictions at trial and on appeals for all outcomes.[10] Each row contains the *F-stat* and *p*-values for the null of OLS consistency across outcomes. I reject consistency for outcomes 1 and 2 (p-value $< .01$) when using the full sample and for outcome 3 when splitting the sample into city council and mayor candidates (also p-value $< .01$). Since the vote distance to election cutoff is much smaller when votes are spread out across many candidates in proportional elections (city council) than in majoritarian elections (mayor), the asymptotic equivalence between OLS and IV parameters in row 3 is entirely plausible (p-value $= .17$; fail to reject H0). Both conviction variables fail to explain electoral outcomes measured in such way and thus carry over low predictive power to their regression models.

These tests confirm the instrument choice and significantly meet the inclusion restriction conditions for causal identification under an IV design. In the following sections, I conduct exclusion restriction tests to provide further support for the effects of crime on performance in this paper.

---

[10]I also run multivariate versions of Hausman tests but there are no changes to p-values. Results are available upon request.

6   Results

Table 5 reports the effect of electoral infractions on the probability of election of each politician. For mayor candidates, this variable turns on when the candidate was the most voted in their election. For city council candidates, this variable turns on when the candidate has received enough votes to finish the election within the number of vacancies in their municipality. For instance, if a municipality has 12 seats in its city council, a candidate who received the same number, or more, votes than the 12th placed candidate has outcome value one.[11] This is the most important outcome, and directly tests the first theoretical claim suggested in section 4, that is, voters would impose electoral penalties when candidates are convicted of electoral crimes and this results in a worse electoral performance than otherwise (the conviction effect).

In columns 1-3 of table 5, I report the OLS estimates of the effect of crimes on outcome. The point estimates start at a 20.8 percentage point reduction on the probability of election but decrease to 16.3 percentage points in model 3, which includes candidate controls and fixed-effects. Thus, there is a significant and negative effect of crimes on the probability of election (p-value $< .01$). Regardless of the specification, there is a negative, baseline relationship between conviction for electoral crimes and performance. The inclusion of covariates and fixed-effects soaks up some of the variation in the conviction variable and controls for observed factors which could be correlated with the convictions.

Unobservable factors, however, could still drive part of the result in columns 1-3. A plausible hypothesis is that some electoral races are more relevant and, as such, there is more competition for seats than otherwise. Candidates might even be more likely not to play by the rules and bring many unfounded claims against their opponents. Or they could just be more skilled or driven. To effectively test such confounding effects, I report the results of instrumental variables regressions in columns 4-6. Note that all IV conviction coefficients have significantly larger magnitudes than their OLS equivalents (again at the one percent level). They range from -27.2 to -23.1 percentage points in models 4 and 6, respectively. They suggest a upward bias in OLS estimates of about 6.2 (models 5 and 2) to 6.8 (models 6 and 3) percentage points; OLS predicts a smaller, weaker impact of crimes on performance than IV estimates. Taken together with the evidence of Hausman tests in section 5, I am confident that IV estimates are more consistent as asymptotically describe the true causal effect on performance. For any given candidate, a conviction at trial alone would reduce their probability of election by 23.1 percentage points, according to my preferred model (column 6).

This result supports the theoretical claim in section 4 and is aligned with similar results pre-

---

[11]City council elections are not necessarily decided in such manner; TSE tallies up all votes in a single election and divides them up by the number of seats available. All candidates who have more votes than this mark are automatically elected to office; remaining seats go to the coalitions who have rounded up more votes. Only rarely, however, all city councilors are elected this way. In most cases, votes are usually spread out across many candidates and coalitions, so being voted in as the last candidate within the number of available seats does guarantee their election and supports their coalitions to get further seats. In addition, this is a less strict way to define who is elected to city council such that, even if there are errors in the list of elected candidates, the correct measurement would decrease the number of elected candidates and reinforce the conviction effect.

sented in the literature. Ferraz and Finan (2008) report a smaller effect of seven percentage points for mayors when audit reports reveal corruption findings before elections in 2004. Though the effect here is larger for a less severe crime, the candidates in Ferraz and Finan (2008)'s sample are generally much more experienced than in this paper. The share of reelected mayors in Ferraz and Finan (2008) is 58.5, compared to 19.1 percent of experience politicians in this sample (the most closely related variable in this analysis), suggesting that ability would indeed offset some of the negative effect of crime (Winters and Weitz-Shapiro, 2013; Pereira and Melo, 2015).

In table 6, I report the results of the same regressions on a candidate's vote share. The OLS estimates are reported in columns 1-3, and show a negative and significant effect of crimes on candidate's vote share, ranging from 12.9 to 9.9 percentage points. The IV effect is about 3.9-3.3 smaller than OLS's. In the preferred model 6, the conviction effect significantly reduces vote share by 13.3 percentage points (p-value $< .01$). Though the difference between OLS and IV parameters here is twice as small as in the probability of election specifications, there are two reasons suporting the results in the IV model: (i) the Hausman test for the vote share specification in section 5.1 rejects the null of OLS consistency; and (ii) the 99 percent CIs around OLS and IV coefficients never overlap.[12] Together, this evidence points to a consistent IV estimation.

Compared to evidence in the literature, the effect size here is larger. Ferraz and Finan (2008) report a 10.4 percentage point decrease in vote share when mayors are running for reelection and have had corruption evidence released to the public prior to municipal elections in 2004. Chong et al. (2015) run an experiment before the municipal elections in three Mexican states in 2009 and find a 1.1 decrease in incumbent mayors' vote share when corruption information is revealed to the public. The differences in research design, however, explain why the effect is smaller in other studies. First, both studies are looking at the effect for incumbent politicians when there is evidence of corruption. These politicians are likely more skilled than the average and thus offset the negative impact of corruption with their ability. Second, they also only look at mayors, rather than city councilors, and the former have more visibility in local politics than the latter. When I reestimate the model in column 6 for the mayor-only sample (unreported here), the conviction effect remains significant and negative but falls to 0.9 percentage points – marginally smaller than Chong et al. (2015).

I lastly investigate the conviction effect for outcome three, vote distance to election cutoff. This effect represents how much candidates' choice of electoral tactic helped getting away (or closer) to the number of votes needed for election. In this analysis, I split the sample into city council and mayor candidates because of the meaningful differences in each office race. Mayor elections follow majority rule; city council elections follow proportional rule. As such, the number of candidates is much smaller, and the votes needed for election much larger, in mayor elections. Therefore, the distance to election is not uniform across office type; in other words, a one percentage point distance is much harder to come by in city council rather than mayor races.

---

[12]OLS and IV pairwise 99 percent CIs are $(-11.9, -14.0)$ and $(-15.6, -18.0)$; $(-7.5, -9.2)$ and $(-10.7, -12.8)$; $(-8.6, -11.3)$ and $(-11.6, -14.9)$.

Table 7 presents the results. OLS specifications are included in columns 1-2, and IV models in columns 3-4. I only report regressions with individual controls and fixed-effects. I find that being convicted at trial has again a negative and significant effect (at one percent) on the vote distance to the election cutoff across all models. For the city council sample, the IV coefficient points to 0.849 percentage point less in the distance to election than in the absence of crime; for the mayor sample, this effect is 7.4 percentage points. Thus, candidates accused, and found guilty, of violating electoral law generally place further away from the necessary votes to guarantee election – in line with the impact of convictions on outcomes one and two. Though the significance and direction of the effect are aligned with previous results, I am skeptical about the effect size reported in table 7. The difference between OLS and IV parameters is much smaller, and their 95 percent CIs overlaps. With a sample size of 9,442 candidates, the OLS and IV distributions are equally consistent.

Despite my skepticism regarding the effect size for outcome three, there is a robust, negative effect of electoral crime on electoral performance. When research different designs are accounted for, these results align well with previous evidence documented in the literature for the impact of other crimes on performance. In the following sections, I conduct multiple robustness checks to support the negative, unbiased, and significant effect of electoral crimes.

7   Exclusion Restriction Checks

In section 5.1, I carried out four tests validating the inclusion restriction of conviction on appeal as an instrument for conviction at trial for estimating the causal effect of electoral crimes on performance. In section 2, I also discussed how the TSE's judicial review process qualifies for exclusion restriction conditions. Though there are no empirical tests for the exclusion restriction, I conduct two indirect checks which support my instrument choice: (i) coefficient stability tests following Altonji et al. (2005) and Oster (2019) and (ii) correlation tests including the instrument in the second stage.

7.1   Coefficient Stability Tests

The most common way to address omitted variable bias is to include controls in the regression of interest. In this paper, I repeatedly report parameter estimates progressively including candidate controls, party, municipal, and election fixed-effects. In many cases, however, the set of controls does not fully identify confounding effects. In fact, scholars rarely use the full set of confounding factors; rather, they use the *observed*, available confounders. Unless the confounding set is fully captured by the available variables, selection on unobservables could still explain a significant portion of the parameters we are estimating in linear models.

Oster (2019) formalizes this point. She suggests that coefficient stability across regression models is only a reliable indication of unbiasedness if scaled by changes in the amount of regression variation explained by independent variables. In other words, the coefficient of interest should move relatively less than R^2, indicating that effect size is stable at the same time as one shifts explained

variation from the error term to the matrix of independent variables. The following equation in Oster (2019) translates this idea:

$$\beta^* = \tilde{\beta} - \delta \cdot [\beta^0 - \tilde{\beta}] \cdot \frac{R_{\max} - \tilde{R}}{\tilde{R} - R^0} \tag{7}$$

Where β^* is the bias-adjusted coefficient; $\tilde{\beta}$ is the coefficient in the unrestricted regression; β^0 is the coefficient in the restricted regression; and $\tilde{R}$ and R^* are their respective R^2. In this setting, we are interested in adjusting δ and $R_{\max}$ such that we can test how $\tilde{\beta}$ fares against β^*: δ is the ratio of the regression variance explained by unobserved and observed controls; $R_{\max}$ is the theoretical population explanation of the outcome. Altonji et al. (2005) and Oster (2019) suggest that $\delta = 1$ is a good threshold for coefficient stability, which means that unobservable variables would have to explain at least as much of the variation of the outcome as do observable variables to explain away $\tilde{\beta}$.

To support the exclusion restriction, I modify equation (7) such that we can compare across OLS and IV coefficients. Finding evidence that unobserved variation is unable to explain the shift between OLS and IV coefficients means that the instrument does not add bias to regression estimates, and thus provides support for the exclusion restriction. In other words, δ is the degree of selection necessary to attribute the difference between OLS and IV to omitted variable bias. Any $\delta > 1$ would mean excessive selection on unobservables is required for β_{iv} to be the biased effect, and not the true effect, of crimes on performance. Rearranging equation (7) and assigning subscripts to coefficients, I estimate the following equation in table 8:

$$\delta = \frac{\tilde{\beta}_{iv} - \beta^*_{ols}}{\beta^0_{iv} - \tilde{\beta}_{iv}} \cdot \frac{\tilde{R} - R^0}{R_{\max} - \tilde{R}} \tag{8}$$

In table 8, I compute δ for two different $R_{\max}$ (reported in columns 1 and 2 in each panel). When such empirical calculations yield a nonsensical $R_{\max} > 1$, I truncate $R_{\max} = 1$. In column 3, I calculate the necessary R^2 for $\beta_{iv} = \beta_{ols}$ under the threshold $\delta = 1$. I do not truncate $R_{\max}$ in column 3 to provide an indication of degree of variation required to make estimates converge. Both of these validation exercises follow Oster (2019).

Except for outcome two, all δ's are greater than one for $R_{\max} = R^2_{ur} + (R^2_{ur} - R^2_r)$. This means that selection on unobservables cannot fully explain the difference between OLS and IV estimates. The same is true when $R_{\max} = 2 \cdot R^2_{ur}$. Though the δ for outcome two in panel A suggests selection is a problem, my preferred models come from panel B, where all δ's are greater than one; the inclusion of the full set of fixed-effects represents a substantial improvement in observed variation and fixes potential selection in outcome two models. Therefore, I find no support for selection effects driving IV estimates, and this result serve as an indication in favor of the exclusion restriction for convictions on appeals.

7.2 Instrument Included in Second-Stage

Having confirmed the consistency of IV estimates, the second test in support of the exclusion restriction is relatively straightforward. I replace the endogenous variable (convictions at trial) for the instrument (convictions on appeals) in the second-stage regression, which is thus transformed in a OLS regression where the variable of interest is the ruling issued by the appellate panel. Such reduced-form approaches are only recommended if we are able to reject OLS consistency in favor of IV – which is the case here. In figure 2, there are four panels, one for each of the outcomes discussed in section 6. In light gray, I plot the coefficient on convictions at trial and its 99 percent CI for all outcomes (4) and regression models (3), yielding a total of 12 estimates and CIs. I do the same for convictions on appeal, the reduced-form, and produce the same estimates and CIs (in black).

It is clear that the instrument will be significantly correlated with the outcome of interest. No 99 percent CI includes zero. This is expected when the instrument passes inclusion restriction tests. The more important result, however, is the similarity of point estimates using either trial and appeals outcomes. Except for vote distance in the mayor candidates sample, all point estimates of appeals coefficients fall inside trial CIs. The correlation between either variable and covariates and fixed-effects is the same, meaning that the instrument is not adding anymore variance to the trial regression. In other words, there is additional support for $cov(z, X) = 0$. Almost all of the effect of instrument z on outcome y occurs via its correlation with x. Along with the evidence in the previous section and the institutional design of judicial review of candidacy registration cases, I can confidently say that there is no independent effect of the instrument on the outcome (or $cov(z, y) = 0$).

8 Alternative Explanations

There are additional threats to validity beyond the inclusion and exclusion restriction. In this section, I first explore how strategic changes of judge, voter, or candidate behavior during the judicial review process could bias the electoral performance result beyond the conviction issued by the TSE. Secondly, I provide another robustness check in which I show the validity of the IV effect even in the presence of weaker correlations between the endogenous variable and the instrument.

8.1 Heterogeneous Sentencing Across Review Stages

8.2 Voter Behavior Changes

8.3 Candidate Behavior Changes

8.4 IV Estimate is Contingent on Small Portion of Data

Though the inclusion restriction tests support convictions on appeal as instruments for convictions at trial, a potential source of concern is that the instrument is *too strongly* correlated with the

endogenous variable. A simple, unconditional Pearson correlation check yields a .796 correlation across conviction types. The appellate decision estimate in the first-stage regression from table 3, column 3 returns a .738 correlation in the presence of all covariates and fixed-effects. For three quarters of the sample, the distribution of trial outcomes is fully identified by the distribution of appeals outcomes. A valid concern, therefore, is whether the difference in estimates across OLS and IV is only coming from a small portion of the data for which trial outcomes are reversed – circa 17 percent of the whole sample.

To address this concern, I investigate how IV estimates change under different correlation co-efficients assumptions for the first stage. The reasoning is straightforward. Suppose I have a distribution of correlations between convictions at trial and on appeal, ranging from 0 to 1. For a weak enough correlation p_{low}, the instrument is weak and does not meet the inclusion restriction. I cannot identify causal effects. For an identity correlation $p = 1$, the instrument becomes meaningless as the endogenous variable is fully captured by the instrument. Similarly, I cannot identify causal effects. However, there is a range of correlations $\{p_{low} > 0, p_{high} < 1\}$ for which the instrument both meets the inclusion restriction and is a consistent estimator for the causal effect of interest. Since the empirical draw of β_{iv} has a high correlation with the endogenous variable and is asymptotically consistent, I can simulate effect sizes for the entire interval of strong instruments $\{p_{low}, p_{instrument} = .796\}$ and compare them to the observed OLS coefficient. If there is no overlap between CIs of OLS and IV parameters, then I more confident about the IV effect size being the true effect.

In practical terms, the simulation starts with the random draw of a (unconditional) correlation coefficient between 0 and .796 (empirical value). I then use the correlation draw to construct a simulated distribution of convictions on appeal with the same average and standard deviation as the empirical variable. Next, I run first and second-stage regressions using the simulated variable as instrument and store coefficients, standard errors, and first-stage F-statistics.[13] I repeat this process 10,000 times to recover the entire distribution of parameters under different correlation assumptions. Once this process is finished, I discard all weak instruments and their estimates using the Bound et al. (1995) cutoff of F-stat $= 10$. With the remaining coefficients from second-stage regressions, I construct figure 3 with various instrument correlation levels.

The first important result is no overlap between either 95 percent CI. The strong instrument IV parameters have a mean of $-.231$ with a standard error of .018. It is the same effect but larger standard error than in the empirical distribution (s.e. $= .016$). Thus, I are confident that the IV estimate pulls out only the exogenous part of convictions on trial. Second, and most importantly, the IV estimate remains statistically significant even when the correlation coefficient, measured by the first-stage coefficient between conviction at trial and on appeals, drops to .525. That is a 28 percent drop in correlation. It means that a third of the affirmed trial decisions would have to be reverted to threaten the IV effect size – an unrealistic high number of reversals. To put this in perspective, the overall reversal rate in this sample is 17.2 percent; the highest reversal rate in the

---

[13]I only use the most important outcome, the probability of election, in this test.

U.S. is 14.1 percent (U.S. Courts, 2018).[14]

9 Heterogeneous Electoral Punishment

10 Conclusion

This project aims at uncovering the effect of electoral crimes on performance. Supplemental to existing literature looking at severe electoral fraud and electoral malfeasance in non or partially democratic regimes, I provide evidence on less known, less understood electoral practices that can also shape the results of elections. I measure the effect of crimes in Brazil, one of the largest and highest quality electoral democracies in the world at the time of the study. Further contributions of this paper are the use of court documents as data and causal identification using the institutional design of electoral judicial review in Brazil.

I find substantial and significant negative effects of electoral crimes on ballot performance. Being convicted of an electoral crime reduces the probability of election for mayor and city councilor candidates in Brazil by 23.1 percentage points. It also reduces vote share by 13.3 percentage points. These results are robust to a series of checks on the inclusion and exclusion restriction, and to strategic changes of judge, voter, and candidate behavior. Contrary to expected, I find that voters do not impose differential punishments conditional on the severity of the electoral crime. There is a positive, independent effect of engaging in substantial rule-breaking on performance, about 4.8 percentage points both in the probability of election and vote share, but when interacted with the conviction effect, the type of conviction does not impose additional punishment beyond what the conviction had caused.

This research question is relevant for multiple policy discussions. I offer additional evidence claiming the existence of a negative relationship between crime and electoral performance beyond just corruption. Knowing that voters punish bad behavior, skilled politicians and policymakers can increase monitoring, detection, and prosecution of electoral crimes as a means of weeding out low-quality office-seeking candidates. A final implication of this project is, indeed, an honest discussion on the effectiveness of electoral oversight authorities in the first place, a feature shared by other developing countries such as Mexico, India, and South Africa. While on the one hand they might prevent low-quality candidates from running, and eventually, being elected, they might simultaneously create barriers to entry that are detrimental to political competition and to the democratic process in developing countries.

---

[14]Surprisingly, the Brazilian judiciary system has little information on reversal rates. The last official report came out in 2009 and did not have any information on reversal rates at the Electoral Court system at the time of publication.

References

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1):151–184.

Asunka, J., Brierley, S., Golden, M., Kramon, E., and Ofosu, G. (2017). Electoral Fraud or Violence: The Effect of Observers on Party Manipulation Strategies. *British Journal of Political Science*, pages 1–23.

Avis, E., Finan, F., and Ferraz, C. (2018). Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians. *Journal of Political Economy*.

Banerjee, A., Kumar, S., Pande, R., and Su, F. (2010). Do informed voters make better choices? Experimental evidence from urban India. *Unpublished manuscript*.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak. *Journal of the American Statistical Association*, 90:443–450.

Chong, A., De La O, A., Karlan, D., and Wantchekon, L. (2015). Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification. *The Journal of Politics*, 77:55–71.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., Bernhard, M., Fish, M. S., Cornell, A., Dahlum, S., Gjerløw, H., Glynn, A., Hicken, A., Krusell, J., Lührmann, A., Marquardt, K. L., McMann, K., Mechkova, V., Medzihorsky, J., Olin, M., Paxton, P., Pemstein, D., Pernes, J., Von Römer, J., Seim, B., Sigman, R., Staton, J., Stepanova, N., Sundström, A., Tzelgov, E., Wang, Y.-T., Wig, T., Wilson, S., and Ziblatt, D. (2018). V-Dem Country-Year Dataset 2018.

de Figueiredo, M. F., Hidalgo, F. D., and Kasahara, Y. (2011). When do voters punish corrupt politicians? Experimental evidence from Brazil. *Unpublished Manuscript, University of California Berkeley*.

Ferraz, C. and Finan, F. (2008). Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. *The Quarterly Journal of Economics*, 123:703–745.

Ferraz, C. and Finan, F. (2011). Electoral Accountability and Corruption: Evidence from the Audits of Local Governments. *American Economic Review*, 101:1274–1311.

Fujiwara, T. (2015). Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil. *Econometrica*, 83(2):423–464.

Gandhi, J. and Lust-Okar, E. (2009). Elections Under Authoritarianism. *Annual Review of Political Science*, 12(1):403–422.

Gans-Morse, J., Mazzuca, S., and Nichter, S. (2013). Varieties of Clientelism: Machine Politics during Elections. *American Journal of Political Science*, 58(2):415–432.

Ichino, N. and Schündeln, M. (2012). Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana. *The Journal of Politics*, 74(1):292–307.

Lambais, G. and Sigstad, H. (2018). *Judicial Subversion: Evidence from Brazil.* Mimeo, Harvard University.

Lehoucq, F. (2003). Electoral Fraud: Causes, Types, and Consequences. *Annual Review of Political Science*, 6(1):233–256.

Levitsky, S. and Way, L. (2002). The Rise of Competitive Authoritarianism. *Journal of Democracy*, 13(2):51–65.

Mares, I. and Young, L. (2016). Buying, Expropriating, and Stealing Votes. *Annual Review of Political Science*, 19(1):267–288.

Nunn, N. and Wantchekon, L. (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review*, 101(7):3221–52.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Pei, Z., Pischke, J.-S., and Schwandt, H. (2019). Poorly Measured Confounders are More Useful on the Left than on the Right. *Journal of Business & Economic Statistics*, 37(2):205–216.

Pereira, C. and Melo, M. A. (2015). Reelecting Corrupt Incumbents in Exchange for Public Goods: Rouba mas faz in Brazil. *Latin American Research Review*, 50(4):88–115.

Sanchez-Martinez, C. (2018). *Dismantling Institutions: Court Politicization and Discrimination in Public Employment Lawsuits.* PhD Dissertation, Stanford University.

Schedler, A. (2015). Electoral Authoritarianism. In Scott, R. A. and Kosslyn, S. M., editors, *Emerging Trends in the Social and Behavioral Sciences*, pages 1–16.

U.S. Courts (2018). Statistical Tables for The Federal Judiciary. Technical report.

Weitz-Shapiro, R. and Winters, M. S. (2017). Can Citizens Discern? Information Credibility, Political Sophistication, and the Punishment of Corruption in Brazil. *The Journal of Politics*, 79:60–74.

Winters, M. S. and Weitz-Shapiro, R. (2013). Lacking Information or Condoning Corruption: When Do Voters Support Corrupt Politicians? *Comparative Politics*, 45:418–436.

Tables and Figures

Table 1: Descriptive Statistics

	N	Mean	St. Dev.	Min	Max
Age	9,470	46.34	11.02	17	86
Male	9,470	.793	.405	0	1
Political Experience	9,470	.091	.287	0	1
Campaign Expenditures (in R$)	9,470	52,555	210,742	0	4,949,250
Convicted at Trial	9,470	.641	.480	0	1
Convicted on Appeal	9,470	.537	.499	0	1
Probability of Election	9,442	.191	.393	0	1
Total Vote Share (in p.p.)	9,442	10.13	17.98	0	100.00
Vote Distance to Election Cutoff (in p.p.)	9,442	−4.09	9.55	−92.82	12.83

Table 2: Electoral Crime Rulings

	Appeals		Percent
Trial	Affirmed	Reversed	Reversed
Not Convicted	3380	22	0.6
Convicted	5059	1009	16.6

Table 3: First-Stage Regressions

	Outcome: Convicted at Trial		
	(1)	(2)	(3)
Convicted on Appeal	.766***	.753***	.738***
	(.006)	(.007)	(.009)
Individual Controls	-	Yes	Yes
Fixed-Effects	-	-	Yes
Observations	9,470	9,470	9,470
Adjusted-R^2	.633	.649	.861
F-stat	16,364.9***	1,094.0***	21.7***

Note: First-Stage regressions here report the correlation between being convicted at trial and being convicted on appeal for all candidates who have had their candidacy challenged under charges of electoral irregularities. I present results including and excluding individual politician characteristics; municipal, electoral, and party fixed-effects; and use robust standard errors. *p<0.1; **p<0.05; ***p<0.01
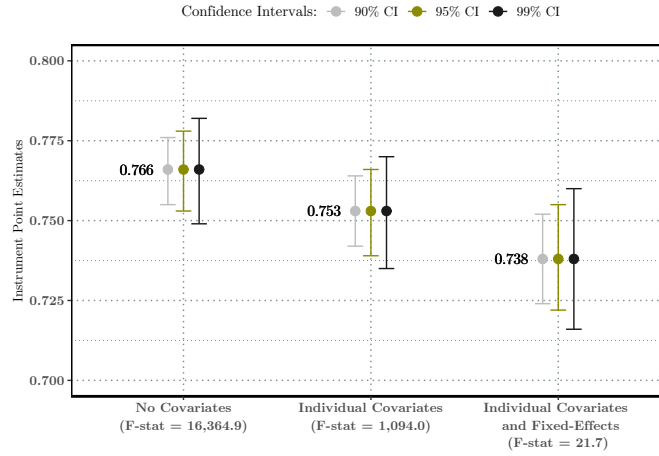
Figure 1: Instrument Point Estimates and CIs

Table 4: Hausman Test of Instrument Strength

Outcome	Hausman Statistic	p-value	Sample
1. Probability of Election	109.28	.000	Full
2. Total Vote Share	205.57	.000	Full
3. Vote Distance to Election Cutoff:	1.88	.170	Full
3.1. City Councilor	65.44	.000	Split
3.2. Mayor	93.43	.000	Split

Table 5: The Effect of Electoral Crime on the Probability of Election

	Outcome: Probability of Election					
	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Convicted at Trial	$-.208^{***}$	$-.151^{***}$	$-.163^{***}$	$-.272^{***}$	$-.213^{***}$	$-.231^{***}$
	(.009)	(.009)	(.014)	(.011)	(.010)	(.016)
Individual Controls	-	Yes	Yes	-	Yes	Yes
Fixed-Effects	-	-	Yes	-	-	Yes
Observations	9,442	9,442	9,442	9,442	9,442	9,442
Adjusted-R^2	.065	.149	.303	.059	.143	.300
F-stat	653.58^{***}	104.02^{***}	2.46^{***}	707.35^{***}	108.9^{***}	2.47^{***}

Note: The regressions here estimate the effect of being convicted at trial on the probability of election for all candidates who have had their candidacy challenged under charges of electoral irregularities. Columns 1 and 4 display models not including individual candidate characteristics; columns 2 and 5 include age, gender, marital status, education level, political experience, and the amount spent in their campaign; columns 3 and 6 also include municipal, electoral, and party fixed-effects. I report robust standard errors for all specifications in this table. $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 6: The Effect of Electoral Crime on the Total Vote Share

	Outcome: Total Vote Share (in p.p.)					
	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Convicted on Appeal	-12.945^{***}	-8.316^{***}	-9.943^{***}	-16.804^{***}	-11.765^{***}	-13.254^{***}
	(.418)	(.337)	(.529)	(.478)	(.399)	(.624)
Individual Controls	-	Yes	Yes	-	Yes	Yes
Fixed-Effects	-	-	Yes	-	-	Yes
Observations	9,442	9,442	9,442	9,442	9,442	9,442
Adjusted-R^2	.119	.379	.606	.109	.371	.602
F-stat	$1,278.91^{***}$	361.57^{***}	6.15^{***}	$1,360.8^{***}$	368.19^{***}	6.14^{***}

Note: The regressions here estimate the effect of being convicted at trial on the total vote share for all candidates who have had their candidacy challenged under charges of electoral irregularities. Columns 1 and 4 display models not including individual candidate characteristics; columns 2 and 5 include age, gender, marital status, education level, political experience, and the amount spent in their campaign; columns 3 and 6 also include municipal, electoral, and party fixed-effects. I report robust standard errors for all specifications in this table. *p<0.1; **p<0.05; ***p<0.01

Table 7: The Effect of Electoral Crimes on the Vote Distance to Election Cutoff

	Outcome: Vote Distance to Election Cutoff (in p.p.)			
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Convicted at Trial	$-.575^{***}$	$-.849^{***}$	-5.172^{***}	-7.381^{***}
	(.064)	(.075)	(1.905)	(2.184)
Individual Controls	Yes	Yes	Yes	Yes
Fixed-Effects	Yes	Yes	Yes	Yes
Sample	City Council	City Council	Mayor	Mayor
Observations	7,100	7,100	2,342	2,342
Adjusted-R^2	.431	.428	.384	.382
F-stat	3.54^{***}	1.86^{***}	3.55^{***}	1.85^{***}

Note: The regressions here estimate the effect of being convicted at trial on the distance to the election cutoff for candidates who have had their candidacy challenged under charges of electoral irregularities. All models include individual candidate characteristics and municipal, electoral, and party fixed-effects. Since election rules differ by office type, I split the sample into city council candidates (columns 1 and 2) and mayor candidates (columns 3 and 4). I report robust standard errors for all specifications in this table. *p<0.1; **p<0.05; ***p<0.01

Table 8: Coefficient Stability Tests: δ's and R^2 for $\beta_{ols} = \beta_{iv}$

	Panel A: Individual Covariate Models			Panel B: Individual Covariate and Fixed-Effects Models		
	(1)	(2)	(3)	(1)	(2)	(3)
	$R^2_{ur}+$ $(R^2_{ur} - R^2_r)$	$2 \cdot R^2_{ur}$	R^2 for $\beta_{ols} = \beta_{iv}$	$R^2_{ur}+$ $(R^2_{ur} - R^2_r)$	$2 \cdot R^2_{ur}$	R^2 for $\beta_{ols} = \beta_{iv}$
Probability of Election	1.05	0.63	-	1.69	1.49	-
	(.23)	(.29)	(.46)	(.96)	(1.00)	(3.07)
Vote Share	0.68	0.48	-	2.05	2.05	-
	(.64)	(.74)	(.99)	(1.00)	(1.00)	(3.01)
Vote Distance to Cutoff (City Councilor)	7.74	6.05	-	20.51	20.51	-
	(.21)	(.23)	(2.11)	(1.00)	(1.00)	(24.86)
Vote Distance to Cutoff (Mayor)	2.64	1.56	-	1.21	1.21	-
	(.23)	(.29)	(.64)	(1.00)	(1.00)	(1.51)

Note: In each panel, I compare the unrestricted coefficient for the model in the panel title ($\tilde{\beta}$) against the restricted coefficient for the bivariate model (β^0). The different outcomes are summarized across rows. Columns 1 and 2 in each panel display conditions for R_{max} calculations in the row just above table content. The first value in each cell is the δ for each model. R^2 values are reported inside parentheses. While I cap R^2 at one for δ calculations, I do not cap it for calculations of the necessary R^2 to yield $\beta_{ols} = \beta_{iv}$; therefore, some nonsensical $R^2 > 1$ appear in column 3 in each panel.

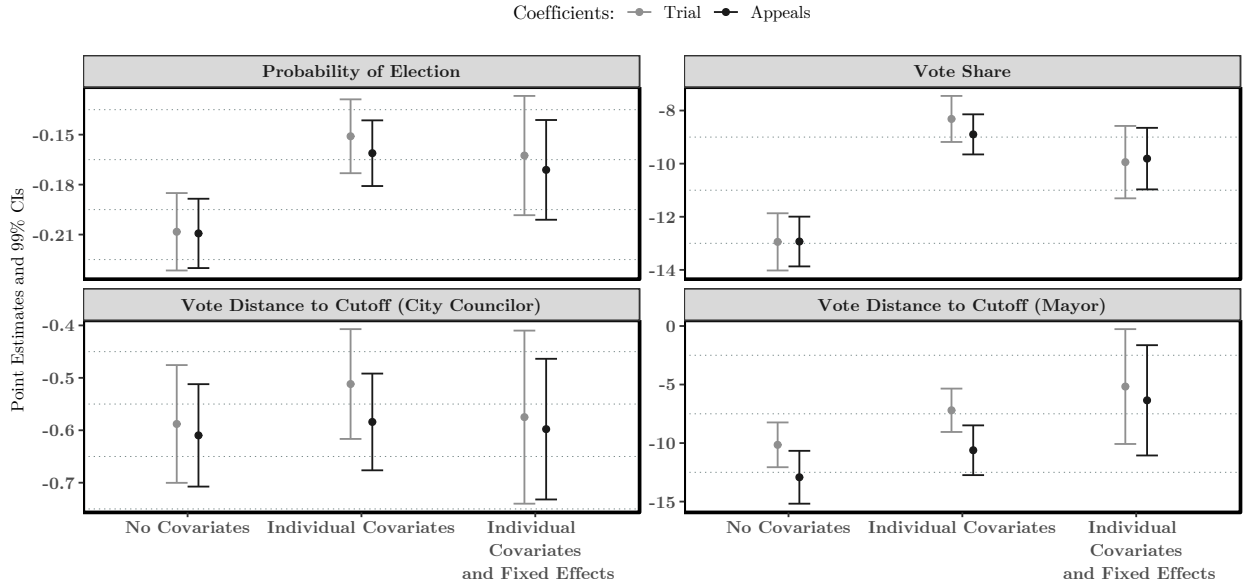Figure 2: Instrument Correlation with Covariates

Table 9: Heterogeneous Sentencing across Trial and Appeals

	(1)	(2)	(3)	(4)	(5)	(6)
	β_{trial}	β_{appeals}	$\beta_{\text{difference}}$	s.e.	t-stat	p-value
Elected to Office	−.223	−.267	.044	.085	.510	.610
Age	−.001	.000	−.001	.003	−.424	.671
Male	.029	.022	.007	.039	.176	.861
Political Experience	−.089	−.013	−.076	.079	−.964	.335
Campaign Expenditures (ln)	−.029	−.028	−.001	.029	−.034	.973
Marital Status:						
Divorced	−.006	.026	−.032	.038	−.839	.402
Legally Divorced	.066	.028	.039	.048	.795	.427
Single	−.008	.043	−.051	.040	−1.276	.202
Widowed	.029	−.011	.040	.064	.626	.532
Educational Levels:						
Completed ES/MS	−.160	−.234	.074	.090	.819	.413
Incomplete ES/MS	−.116	−.259	.143	.134	1.066	.286
Can Read and Write	−.066	−.286	.220	.174	1.268	.205
Completed HS	−.191	−.259	.068	.085	.799	.424
Incomplete HS	−.108	−.264	.156	.132	1.180	.238
Completed College	−.218	−.300	.083	.099	.833	.405
Incomplete College	−.177	−.270	.093	.125	.742	.458

Note: In this table, I report the coefficients of two regressions using the same covariates on the probability of receiving an unfavorable ruling at trial (column 1) and on appeals (column 2). I then recover the distributions of the differences in betas and test H0: $\beta_{\text{difference}} = 0$ for all covariates in the regressions (columns 3-6). Robust standard errors are clustered at the municipal-election pair level (equivalent to the judge-level error shared by all candidates in one municipality during one election period); party-fixed effects are included in both regressions but are not reported here.

Table 10: The Effect of Electoral Crimes on Voter Engagement

	Party-Level		Election-Level	
	Outcome: Voter Turnout (percent)	Outcome: Invalid Votes (percent)	Outcome: Voter Turnout (percent)	Outcome: Invalid Votes (percent)
	(1)	(2)	(3)	(4)
Share of Candidacies Invalid at Trial	.003 (.007)	.222*** (.076)	−.001 (.009)	.134* (.070)
Individual Controls	-	-	-	-
Fixed-Effects	Yes	Yes	Yes	Yes
Observations	5,322	5,322	3,757	3,757
Adjusted-R^2	.997	.973	.995	.946
F-stat	214.3***	354.1***	81.8***	124.5***

Note: The regressions here estimate the effect of the share of candidates convicted at trial overall the total office vacancies on voter turnout and the number of invalid votes (both logged). I aggregate observations up to the party and election level and control for municipality and election year fixed-effects. I report robust standard errors, clustered by elections and municipalities, for all specifications in this table. *p<0.1; **p<0.05; ***p<0.01

Table 11: Campaign Expenditure Across Ruling Group

Stage	Mean Campaign Spending in Ruling Group (in R$)		t-stat	p-value
	Favorable	Unfavorable		
Trial	84,766	34,497	9.45	.000
	[3,402]	[6,068]		
Appeals	73,275	34,658	8.62	.000
	[4,389]	[5,081]		
Unfavorable Ruling	Affirmed	Reversed	t-stat	p-value
Trial	34,346	34,527	−0.05	.961
	[5,059]	[1,009]		

Note: This table reports t-tests across different subsamples of candidates. The number of observations in each group is reported inside the squared brackets.

Table 12: Voter Sophistication and Benefit of Rule-Breaking

|		β_1: Substantial Violation	
		$\beta_1 = 0$	$\beta_1 > 0$
β_2: Convicted at Trial × Substantial Violation	$\beta_2 = 0$	**1.** Violation carries no electoral benefit. **2.** Voters impose same penalty for different electoral violations.	**1.** Violation helps candidate get elected. **2.** Voters impose same penalty for different electoral violations.
	$\beta_2 < 0$	**1.** Violation carries no electoral benefit. **2.** Voters impose harsher electoral penalties for substantial violations.	**1.** Violation helps candidate get elected. **2.** Voters impose harsher electoral penalties for substantial violations.

Figure 3: Simulation of IV Point Estimates

IV Coefficient Point Estimate Simulations

26
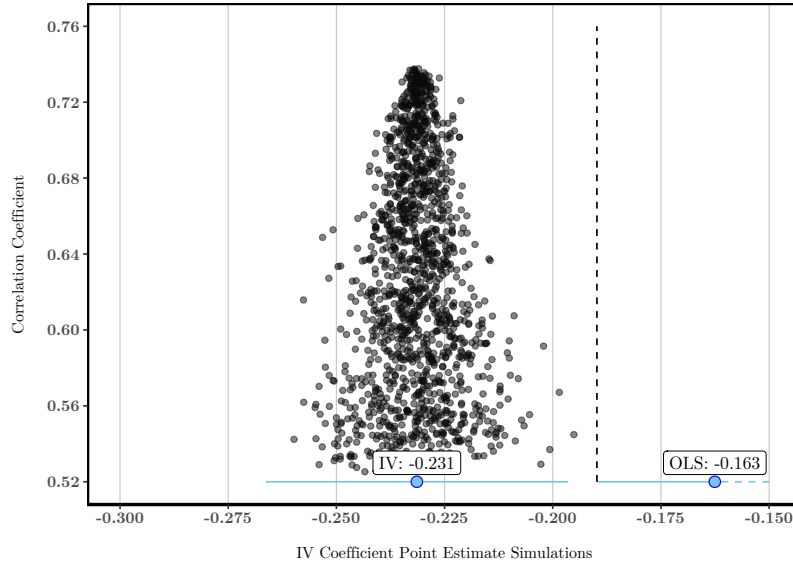
Table 13: Heterogeneous Effect of Electoral Ruling

	Full Sample		City Councilor	Mayor
	Outcome: Probability of Election	Outcome: Vote Share (in p.p.)	Outcome: Vote Distance to Cutoff (in p.p.)	Outcome: Vote Distance to Cutoff (in p.p.)
	(1)	(2)	(3)	(4)
Convicted at Trial	−.176***	−7.369***	−.713***	−6.653***
	(.020)	(.719)	(.084)	(2.101)
Substantial Violation	.047**	4.939***	.089	.169
	(.024)	(.723)	(.103)	(1.524)
Convicted at Trial × Substantial Violation	−.014	−4.952***	.015	1.644
	(.028)	(.915)	(.111)	(2.562)
Individual Controls	Yes	Yes	Yes	Yes
Fixed-Effects	Yes	Yes	Yes	Yes
Observations	4,717	4,717	3,465	1,252
Adjusted-R^2	.375	.697	.499	.380
F-stat	2.54***	6.84***	3.70***	1.73***

Note: The regressions here include the severity of the accusation brought against candidates running for municipal office. I recover the accusations from court documents and identify ruling type using linear support-vector machine classification (details in appendix A). In columns 1-4, I report the coefficients on ruling outcome (row 1), type (row 2), and their interaction (row 3). All regressions include municipal, electoral, and party fixed-effects. Robust standard errors are displayed inside parentheses. *p<0.1; **p<0.05; ***p<0.01

A   Appendix: Electoral Ruling Classification