Mar 2017 No.316

The Political Economy of Program Enforcement: Evidence from Brazil

Fernanda Brollo, Katja Maria Kaufmann and Eliana La Ferrara

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

The Political Economy of Program Enforcement: Evidence from Brazil*

Fernanda Brollo Katja Kaufmann Eliana La Ferrara

This version: March 2017

Abstract

Do politicians manipulate the enforcement of conditional welfare programs to influence electoral outcomes? We study the Bolsa Familia Program (BFP) in Brazil, which provides a monthly stipend to poor families conditional on school attendance. Repeated failure to comply with this requirement results in increasing penalties. First, we exploit random variation in the timing when beneficiaries learn about penalties for noncompliance around the 2008 municipal elections. We find that the vote share of candidates aligned with the President is lower in zip codes where more beneficiaries received penalties shortly before (as opposed to shortly after) the elections. Second, we show that politicians strategically manipulate enforcement. Using a regression discontinuity design, we find weaker enforcement before elections in municipalities where mayors from the presidential coalition can run for reelection. Finally, we provide evidence that manipulation occurs through misreporting school attendance, particularly in municipalities with a higher fraction of students in schools with politically connected principals.

*We thank Oriana Bandiera, Dan Bernhardt, Leonardo Bursztyn, Ana De La O, Horacio Larreguy, Salvo Nunnari, Jim Snyder, Debraj Ray, Alberto Solé-Ollé, Piero Stanig, and seminar participants at Harvard, LSE, University of Warwick, IDB, Stanford GSB, Barcelona GSE Summer Forum and Geneva EEA Congress for helpful comments. Giulia Zane, Simone Lenzu, Emanuele Colonnelli and Matteo Gamalerio provided excellent research assistance. Fernanda Brollo gratefully acknowledges financial support from the Centre for Competitive Advantage in the Global Economy (CAGE). Fernanda Brollo, University of Warwick, CAGE and CEPR, f.brollo@warwick.ac.uk; Katja Kaufmann, Mannheim University, CESifo and IZA, kaufmann@vwl.uni-mannheim.de; Eliana La Ferrara, Bocconi University, IGIER and LEAP, eliana.laferrara@unibocconi.it

1 Introduction

An increasing number of social welfare programs rely on some form of conditionality, such as proof that the beneficiary is actively seeking a job in order to receive unemployment benefits, regularly visiting a doctor or midwife in order to get maternity grants, or immunizing children and sending them to school to receive anti-poverty cash transfers. Conditional welfare programs are common both in industrialized and in developing countries. Regardless of the sector or region where they are implemented, a fundamental tenet on which these programs ultimately rest is that conditions are enforced according to the rules. Our paper focuses precisely on the enforcement of program conditions, and in particular on how political incentives may shape this enforcement.

When a government strictly enforces program conditions and penalizes beneficiaries who do not comply, voters might react in two opposite ways. On the one hand, voters who lose their benefits may be disgruntled and may not vote for the politician they believe is responsible. On the other hand, compliant beneficiaries and taxpayers may appreciate the government's action and reward it at the polls. Depending on which effect dominates, politicians may have incentives to implement the program according to the rules, or to manipulate enforcement by becoming less strict in election years. In the latter case, even programs that are nominally conditional may be *de facto* unconditional (or have less stringent conditions) due to electoral incentives.[1] If conditionality is important to achieve the objectives of the program, this will ultimately reduce program effectiveness.

The tradeoff between strict and lax implementation depends on the visibility of program enforcement. Reducing the enforcement of program conditions can be a direct way to reach a subset of the population without being too visible to other groups. For example, noncompliance by individuals may not be observable to others, hence the fact that the government pays benefits to them despite noncompliance may be more difficult to detect compared to say, mistargeting, where non-eligible individuals (e.g., too rich) receive the benefit. In other words, as social welfare programs become increasingly sophisticated, so does the ability of politicians to manipulate them for electoral purposes without being detected. This has important implications in terms of political accountability that have so far been neglected in

[1] For a review on the effects of conditional versus unconditional cash transfer programs, see Baird, Ferreira, Özler and Woolcock (2013).

the literature.

This paper addresses two questions. First, what are the electoral consequences of enforcing program conditions? Second, do politicians strategically manipulate enforcement in response to electoral incentives? We address these questions studying the implementation of the Brazilian program Bolsa Familia (BFP), the largest conditional cash transfer program in the world. BFP provides a monthly stipend to poor families that is conditional, among other things, on every school-aged child in the family attending at least 85 percent of school days every month. The program is enforced through a system of "warnings" which gradually increase in severity with subsequent instances of noncompliance, leading to suspension of the benefit and ultimately exclusion from the program.

In the first part of the paper we test whether stricter enforcement of BFP school attendance requirements affects electoral outcomes. To causally identify the impact of enforcement, we exploit a feature of the schedule of warnings combined with the timing of the 2008 mayoral elections in Brazil. The date of the month when BFP beneficiaries receive benefits and are notified of penalties depends on the last digit of their 11-digit Social Identification Number (NIS), which is random. The second round of the 2008 municipal elections was held on October 26^{th} and noncompliant beneficiaries with last digit of their NIS from 1 to 5 received notifications in the week before the election, while those with higher last digits received them in the week after the election. We exploit this random assignment by comparing zip codes within a given municipality where a higher fraction of noncompliant beneficiaries learnt about their penalties just before the elections, with zip codes where a higher fraction learnt about penalties just after the elections.

We find that voters respond negatively to the enforcement of BFP rules and that they associate this enforcement with the national government. Within municipalities, the vote share of candidates from the President's party (*Partido dos Trabalhadores*, PT) and its coalition is lower in zip codes where a higher fraction of noncompliant beneficiaries received warnings before as opposed to after the elections. A one standard deviation increase in the fraction of beneficiaries receiving warnings before the elections reduces the vote share of the mayoral candidate from the President's coalition by one percentage point. We do not find effects on local incumbents from non-coalition parties, consistent with the fact that Bolsa Familia was strongly associated with President Lula's social policy.

In the second part of the paper we investigate whether politicians strategically become more lenient in enforcement around the time of elections. In particular, we test if mayors who face stronger electoral incentives are more likely to manipulate enforcement, where electoral incentives are captured by the possibility of running for reelection. Mayors in Brazil are only allowed to run for a consecutive term once, so we analyze whether the enforcement of BFP requirements before elections differs between first and second term mayors. We use a regression discontinuity design (Lee, 2008), analyzing municipalities where the incumbent ran for re-election in the previous election and comparing municipalities where the incumbent won by a narrow margin with municipalities where he or she lost by a narrow margin.

We find that the fraction of beneficiaries who receive warnings is lower in municipalities where mayors from the President's coalition can run for reelection, compared to municipalities where mayors from the coalition cannot run for re-election. We find no difference between first and second term mayors who are not affiliated with the presidential coalition, consistent with the finding reported above that voters tend to associate BFP with the President's party. Manipulation of enforcement is concentrated in the election year, with no evidence of manipulation in the previous year. Moreover, we find no evidence that the number of beneficiary families is different in municipalities with first term mayors, suggesting little or no manipulation in the *allocation* of program benefits. Our results do not seem to reflect differences between first and second term mayors, such as differences in political ability, experience in office, or education-related policies.

Finally, we provide evidence on possible mechanisms for manipulation. We find that manipulation mostly occurs through misreporting of attendance: the fraction of students who did not meet the attendance requirement is lower in municipalities where first term mayors from the coalition won by a narrow margin. Furthermore, we exploit heterogeneity in the extent to which schools are linked to the local government based on whether school principals were politically appointed, as these principals may be more susceptible to political pressures. We compare municipalities where a higher or lower fraction of BFP students are enrolled in schools with politically appointed principals and we find that misreporting of attendance is driven by those municipalities with a higher fraction of BFP students in politically connected schools.

Our work is related to several strands of literature. Numerous contributions have shown that government spending affects election outcomes (e.g., Levitt and Snyder, 1997) and in particular that cash transfer programs bring electoral rewards to incumbents (e.g., Zucco, 2008; Manacorda, Miguel, and Vigorito, 2011; De La O, 2013; Labonne, 2013; Rodríguez-Chamussy, 2015).[2] While most of this literature analyzes how the introduction of the program or the enrollment of beneficiaries affect votes, we focus on implementation issues, and in particular on the enforcement of program conditions.

A second strand of literature has shown that politicians react to electoral incentives, generating political business cycles (e.g., Alesina, 1988; Drazen, 2001; Labonne, 2016) or strategically allocating spending to certain groups or regions (e.g., Levitt and Snyder, 1995; Solé-Ollé and Sorribas-Navarro, 2008; Brollo and Nannicini, 2012). With specific reference to conditional cash transfers, there is little evidence that politicians manipulate the allocation of benefits, which seems instead to be driven by programmatic criteria such as poverty rates (e.g., Green, 2006 ; Fried, 2012). We contribute to this literature by showing that politicians may not distort the allocation of the transfers, but might rather manipulate the enforcement of program conditions.

Another body of work studies the effects of term limits, which we use as a measure of electoral incentives. The evidence on the effects of term limits on the performance of elected officials is mixed. On the one hand, several papers find evidence that incumbents who are eligible for reelection perform better and that their policies are better aligned with the interests of voters (e.g., Besley and Case, 1995; Ferraz and Finan, 2011; de Janvry, Finan, and Sadoulet, 2012). On the other hand, other papers find that reelection incentives lead to stronger political budget cycles and more pork barrel spending (Aidt and Shvets, 2012; Labonne, 2016). Particularly related to our work is the paper by de Janvry et al. (2012), who study Bolsa Escola, the program that preceded Bolsa Familia. The authors find that the program was more effective in reducing school dropout in municipalities where mayors could be reelected, as they adopted more transparent practices (e.g., in registering beneficiaries or forming city councils). Aside from using a different identification strategy, our work differs from theirs because we focus on the enforcement of conditionality (conditions were not strictly enforced during the period studied by de Janvry et al. 2012) and on the electoral costs and

[2] Finan and Schechter (2012) study the mechanisms of reciprocity underlying the exchange of votes for material benefits.

incentives for manipulating enforcement.

The remainder of the paper is organized as follows. Section 2 describes the institutional setting and data. Section 3 analyzes whether voters respond to the enforcement of BFP conditions. Section 4 examines whether politicians manipulate program enforcement before elections, and Section 5 investigates the role of school principals in this manipulation. Section 6 concludes.

2 Institutional Setting and Data

2.1 Background on Bolsa Familia

Coverage: The Bolsa Familia Program, launched in 2003, is currently the largest conditional cash transfer program in the world, reaching around 14 million Brazilian families, that is, 60 million poor people (equivalent to about 30 percent of the Brazilian population). Funds invested in the program represent 0.5 percent of the Brazilian GDP and 2.5 percent of government expenditure.

Benefits: BFP provides a monthly stipend that depends on family income and the number of children. During the period we study, from January 2008 to July 2009, all families considered poor (monthly income per capita below 58 reais, approximately USD 30) were eligible to receive a monthly stipend of 62 reais. In addition, families with monthly per capita income below 120 Brazilian reais and with children under 16 years old attending school were eligible to receive 18 reais per child (20 reais after June 2008), for up to three children. The magnitude of the benefits is large for poor families in Brazil. For instance, a poor family with three children attending school would receive a monthly stipend representing 40 percent of its total family income.

BFP stipends are distributed directly to each family head through a "citizen card" which is mailed to the family. This card operates like a debit card and is issued by the Caixa Econômica Federal, a government-owned savings bank (and one of the largest banks in Brazil). The money can be withdrawn every month in over 14,000 Caixa locations. Beneficiaries receive their stipend on different days of the month depending on the last digit of their 11-digit Social Identification Number (NIS).

The targeting of the program is conducted in two steps. In the first step, the federal government allocates BFP quotas to municipalities according to poverty estimates based on poverty maps. In the second step, eligibility at the household level is determined. Families need to register with the municipal administration and declare their income. This information is transmitted to the federal government and collected in a central database known as the Cadastro Unico. Family eligibility is then determined by the Ministry of Social Development (MDS).

Monitoring conditionality: Benefits are conditional on school attendance by all school-aged children in the family, as well as vaccinations and medical checkups. Each school-aged child has to attend at least 85 percent of school days each month; absences due to health reasons can be excused and do not count towards the number of absent days. In addition, families are required to keep an up-to-date record of vaccinations and health checks for children younger than seven, and pregnant and lactating women must attend regular medical checkups. We focus our analysis on school attendance requirements as these are strictly monitored on a monthly basis. In contrast, monitoring of health-related requirements is soft and conducted only every six months.

Since 2006 the Brazilian central government significantly increased efforts to effectively monitor school attendance and enforce program rules. The Ministry of Education (MEC) is responsible for monitoring school attendance through the following procedure. The MDS feeds the system with information on all beneficiaries with school-aged children that should have their attendance monitored in a given month. MEC accesses this system and makes this information available to all municipalities. Every municipal administration has a conditionality manager who is in charge of accessing and distributing this information to all schools in the municipality. Each school receives a list of the current BFP beneficiaries in the school from the municipal administration. Data on daily school attendance for all children are collected by teachers, and consolidated by school directors, who can "justify" nonattendance, so that it does not count towards noncompliance. Every two or three months ("monitoring period"), monthly school attendance data for BFP beneficiaries are loaded into the system and sent to MEC, which consolidates all the information before reporting it to MDS. The conditionality manager in each municipality is responsible for collecting school attendance information, consolidating the information, and checking its quality. In schools that have computers and internet access, school principals may load daily attendance data into the system and send them directly to MEC.[3]

Sanctions: The program is enforced through a gradual system of *"warnings"*. The first time a family does not comply with the program requirements, the family receives a notification, without any financial repercussions. If noncompliance continues, a series of penalties is activated. In the second warning stage, benefits are blocked for 30 days; after this period the family receives the accumulated benefit for the previous and the current month. The third and fourth warning stages lead to a loss of benefits for 60 days each time. After the fifth warning stage, the benefit is cancelled and the family loses eligibility (the family can return to the program after 18 months, but the municipal administration can decide to allow a family back sooner).

Beneficiaries are well informed about program requirements and punishments for noncompliance: these rules are advertised on newspapers, radio and television, and written in a booklet that each beneficiary family receives. In case of noncompliance, a family receives a warning message at the time of withdrawing their monthly benefit at the bank. This message lists the instance(s) of noncompliance and reminds the family of the warning stage they are in and the punishment they may receive in the future in case of continued noncompliance.

2.2 Background on Brazilian Political Institutions

The layers of political and administrative organization in Brazil are the federal government, the states, and the municipalities. Municipalities are minor federative units with an autonomous local government, ruled by a mayor directly elected by citizens to a four-year mandate, and by a legislative body, also directly elected by voters. Mayors of municipalities above 200,000 voters (around 80 municipalities) are directly elected by a majority runoff rule, while mayors of municipalities below 200,000 voters (around $5,490$ municipalities) are directly elected with plurality rule. The elections of the President, governors, and members of Congress all take place at the same time every four years, while municipal elections are staggered by two years and also take place every four years. Before 1998 Brazilian mayors could not run for re-election, but since 1998 mayors were allowed to run for a second term.

[3] 60 percent of the schools in our sample have access to internet.

2.3 Data

To analyze the enforcement of the Bolsa Familia program we make use of a unique dataset assembled combining many different sources. First, we use administrative data from the Ministry of Social Development (MDS) to construct a dataset on all beneficiary families with children aged 6 to 15 containing information on monthly school attendance of each child for the period 2008-2012.[4] Second, we combine this dataset with monthly information from MDS on warnings and benefits that the family received and whether the benefit was blocked or suspended in a given month. These two sources allow us to know who should have received warnings or penalties based on insufficient attendance, and to verify whether these people were effectively sanctioned.

Electoral data come from the *Tribunal Superior Eleitoral* (Superior Electoral Court), which is the highest judicial body of the Brazilian Electoral Justice. We obtained electoral outcomes for the 2004 and 2008 mayoral elections at the polling station level, as well as information on candidate characteristics such as gender, education, political affiliation and political experience.

For our robustness checks and the analysis of the underlying mechanisms we rely on the following additional data sources.We use data from the 2008 and 2012 School Census conducted by the Ministry of Education and we construct the share of public schools in each municipality with different types of infrastructure (link to the water and sewage system, access to electricity, access to internet), total enrolment in public schools, and school characteristics such as teacher/pupil ratio, average class size, dropout and completion rates. Data on municipal characteristics such as income per capita, population, urbanization, infrastructure, and the presence of a local radio station, are obtained from the 2000 Brazilian demographic census. Finally, to identify politically connected schools, we use data from *Prova Brasil* 2007 and 2011, which tests school children's proficiency in language and math nationwide and which includes information on how school principals were appointed.

When matching the above data sources we aggregate the data at two levels. For the first part of the analysis, which estimates the effect of enforcement on election outcomes, we

[4] In particular, for children below the 85 percent attendance threshold we know the exact fraction of hours attended; for those above the threshold we only know that this condition was met. Attendance data are not available before 2008.

aggregate the data at the zip code level, which is the smallest geographic unit to which we can assign BFP beneficiaries.[5] For the second part of the analysis, which estimates the effects of electoral incentives on enforcement, we aggregate the data at the municipality level, since this is the level at which incentives for the mayor are defined (e.g., the possibility to run for re-election).

Table 1 reports the summary statistics for the main variables of interest. The detailed description of these variables, as well as of other controls used, is presented in Appendix Table A1.

[Insert Table 1]

Panel A shows zip code level statistics for the sample of municipalities where the President's coalition runs. On average, 49 percent of the votes in these areas go to the coalition, and 11 percent of beneficiaries in the sample receive a warning for noncompliance. Appendix Table A2 reports analogous statistics for the samples of municipalities where the president's party (*Partido dos Trabalhadores*, PT) runs and where an incumbent mayor from PT runs.

Panel B shows statistics for the variables used in the second part of the analysis, i.e., municipal level averages for the RDD sample. Analogous statistics are reported in Appendix Table A3 for the two samples of municipalities where the mayor is or is not affiliated with the presidential coalition, since we use those subsamples in our analysis.

3 Do Voters Respond to the Enforcement of Program Requirements?

This section analyzes whether the enforcement of the rules of Bolsa Familia affects local electoral outcomes, focusing on mayoral elections in 2008. According to retrospective voting theory (e.g., Fiorina, 1978), voters use politicians' performance as a signal about their ability and effort, which are not observable. As a result, politicians get credit, or are blamed, for

[5] We do not have data on where BFP beneficiaries vote, therefore we assume that they vote in the zip code where they live. For this analysis, our sample excludes individuals who live in zip code areas without polling stations, as for those zip codes there is no electoral data. For a small number of states we have information on where beneficiaries live and exactly where they vote. In those states, we find that 75 percent of the beneficiaries who live in a zip code area with a polling station, vote in that same area.

policy outcomes that voters believe they are responsible for.[6]

In our case, from a theoretical perspective it is not clear whether and how voters would respond to a strict enforcement of the program rules, for at least two reasons. First, consider the reaction of BFP beneficiaries who did not comply with program conditions and got penalized. If beneficiaries are perfectly rational and expect punishment with certainty, they should anticipate that they will receive warnings already after noncompliance. In this case, we should not find any effects on electoral outcomes coming from the actual receipt of a warning close to elections. On the other hand, if enforcement is not perfectly implemented or anticipated by noncompliant beneficiaries, they may respond to the actual receipt of warnings or punishments. Consistent with this argument, Brollo, Kaufmann, and La Ferrara (2015) show that households react to the actual receipt of a warning by increasing school attendance of their children.

Second, even if voters respond to the enforcement of the program, it is not clear how this response should affect electoral outcomes. Beneficiaries who lose their benefits due to noncompliance (or receive a notification that they might lose their benefits in the future) may be disgruntled and punish politicians at the polls. On the other hand, beneficiaries who comply with program rules may appreciate the fact that the program is implemented in a rigorous way, and reward politicians at the polls.[7] Moreover, taxpayers in general may also value strict enforcement of program rules. So whether the enforcement of the program has a positive or negative effect on electoral outcomes is an empirical question.

3.1 Methodology

Identifying the effects of the enforcement of program rules on voting behavior is challenging, as enforcement may be correlated with other factors that affect electoral outcomes. For instance, municipalities with better program enforcement may differ in terms of income, institutional quality and/or voter preferences, which are likely to affect electoral outcomes. To address this problem, we exploit random variation in the timing when different beneficiaries

[6] A number of papers also suggest that voters do not filter out outcomes that politicians are not responsible for, such as weather or commodity prices (see, among others, Healy and Malhotra, 2010; Wolfers, 2007).

[7] In Brollo, Kaufmann, and La Ferrara (2015) we show that BFP recipient families learn from the warnings received by others, i.e. there is communication about the extent of enforcement of the program, so that the information reaches beneficiary families that have not received warnings themselves.

learn about penalties for noncompliance. In particular, the exact date of the month when beneficiaries receive any notifications of penalties depends on the last digit of their Social Identification Number (NIS), which is random. The second round of the 2008 municipal elections was held on October 26^{th} and noncompliant beneficiaries whose last digit of their NIS was between 1 and 5 received notifications of penalties in the week before the elections, while those with higher last-digits received them in the following week.[8] We exploit this random assignment by comparing zip codes within a given municipality where a higher fraction of noncompliant beneficiaries received notifications *before* the elections to zip codes where a higher fraction were penalized *after* the elections.[9]

The identifying assumption is that, within municipalities, variation across zip codes in the fraction of beneficiaries with last digit of their NIS below and above 5 is orthogonal to the error term in our regression of interest. This is indeed the case due to the random nature of the last digit of the NIS.[10] Our identification strategy implies that, among other things, our results are not driven by differences across zip codes in the level of compliance with program requirements.

We conduct the analysis at the zip code level and construct different enforcement variables considering, alternatively, all program warnings or only those warnings that imply a financial cost for beneficiaries. We only include zip codes where at least one beneficiary family does not meet the school attendance requirement, given that only those families can potentially receive a warning or penalty.[11] We scale the number of beneficiary families that received a warning in October by the number of beneficiary families in the zip code during the relevant monitoring period, that is April-May 2008 (according to the monitoring calendar, low atten-

[8] The fortuitous occurrence of the election in the middle of the warning period for the month of October is the reason why we focus on 2008 for this part of the analysis.

[9] We work at the zip code level for two reasons. First, the larger the unit of analysis, the less variation there is across units in terms of fraction with NIS above/below 5. If we aggregated up to the municipality level, there would be no variation as every municipality has one half of BFP recipients with NIS 1 to 5 and one half with the remaining digits. The variation is driven by random allocation of NIS across zip codes within each municipality. Second, working at the zip-code level we are able to control for time-invariant municipal level characteristics. We tested if our main results (those reported in Table 2 below) are driven by zip codes that are below the median in terms of number of beneficiaries, and found that they are not.

[10] In Appendix Table AA1 we show that, after controlling for the fraction of BFP beneficiaries that receive a warning, the fraction receiving it before the election (i.e., the regressor $FractionTreated_z$ in equation (1) below) is uncorrelated with other zip code level controls. This is consistent with the random nature of the last digit of the NIS.

[11] We also estimated the regressions including all zip codes, even those where no beneficiary family fails the attendance requirements, and obtained similar results.

dance reported during April-May 2008 resulted in warnings in October 2008).[12] In terms of electoral outcomes, we consider the vote share of mayoral candidates from the President's party (PT) or coalition, as well as the vote share of incumbents from these and from different parties. Focusing on PT and its coalition is particularly useful given that Bolsa Familia is a federal program that was the cornerstone of President Lula's social policy and voters tend to associate the program with PT.

We estimate the following regression:

$$Y_{zi} = \alpha + \beta \mathit{FractionTreated}_{zi} + \delta \mathit{FractionWarned}_{zi} + \varphi X_{zi} + \theta_i + \epsilon_{zi}, \tag{1}$$

where Y_{zi} is, alternatively, the vote share of the mayoral candidate of the presidential party or coalition, in zip code area z in municipality i; $\mathit{FractionWarned}_{zi}$ is the number of families that get a warning in October 2008 divided by the number of beneficiary families; $\mathit{FractionTreated}_{zi}$ is the number of families that received a warning in the week before the elections (last-digit NIS from 1 to 5) divided by the number of beneficiary families; X_{zi} are additional control variables, including the ratio of the number of BFP beneficiary families to the electorate (*BFB/electorate*), the ratio of the number of BFP beneficiary families that did not comply to the number of beneficiary families (*Fail/beneficiaries*), and the ratio of the number of BFP beneficiary families that should receive the benefit before the elections (last-digit NIS from 1 to 5) to the number of beneficiary families (*NIS 1 to 5/beneficiaries*); θ_i denotes municipality fixed effects. We report heteroskedasticity robust standard errors.

Our coefficient of interest is β, which captures whether, within a given municipality, the vote share is different in zip codes where a higher fraction of noncompliant beneficiaries received warnings *before* the elections, compared to zip codes where a higher fraction received warnings *after* the elections. A positive value for β would indicate the presence of electoral gains from enforcement, while a negative value would indicate that enforcing BFP rules is politically costly.

[12] We also estimated our regressions scaling the number of families that receive a warning, alternatively, either by the electorate or by the number of families that do not meet the attendance requirement. For both measures we obtained similar results to those reported in Tables 2 and 3 below. Also, we winsorize all enforcement variables at the 99 percent level to handle outliers. Results are not sensitive to this procedure.

3.2 Results

[Insert Table 2]

Table 2 presents our results analyzing the effect of the enforcement of Bolsa Familia program requirements on voter behavior. The main results are presented in columns 1 to 4. Columns 5 to 8 present the results for a placebo test which considers warnings received after the elections. In columns 1, 2, 5 and 6 we study the electoral effect of all types of warnings, regardless of whether they carry an immediate financial penalty. In columns 3, 4, 7 and 8 we instead focus on warnings that imply financial losses. For these regressions, the numerators of the variables *Fraction warned* and *Fraction treated* include only families that received warnings in stages 2 to 5 (i.e., warnings that entail losing the transfer, at least temporarily). In columns 1, 3, 5 and 7 the dependent variable is the vote share of the mayoral candidate affiliated with the President's party (*share PT*); in columns 2, 4, 6 and 8 it is the vote share of the mayoral candidate affiliated with the coalition of the president (*share coal*).[13]

The results in Table 2 suggest that there are political costs of enforcing program requirements for candidates belonging to the President's party or coalition: the vote share of these candidates is significantly lower in zip code areas where a higher fraction of beneficiaries received a warning or penalty in the days before the elections (columns 1-4). These results hold both when analyzing all warnings and only warnings that imply financial losses, the effects being larger in the latter case. These findings suggest that the negative reactions of beneficiaries who get punished more than compensate for any potential political benefits of enforcement. In terms of economic magnitude, the results in column 2, for instance, imply that a one standard deviation increase in the fraction of beneficiaries who receive a warning before the elections reduces the vote share of the mayoral candidate from the President's coalition by one percentage point.

To further strengthen the credibility of our identification strategy, we conduct a falsification test considering as explanatory variables warnings received in November 2008, i.e., after

[13] Note that the sample of municipalities included in the regressions varies across columns, depending on the sample of parties that ran in the second round elections. For instance, in columns 1, 3, 5 and 7 when we analyze the vote share of the mayoral candidate affiliated with PT, we can only consider municipalities where this candidate ran in the second round. There are 15 municipalities where a PT candidate ran. Similarly, in columns 2, 4, 6 and 8 we can only include municipalities where the candidate from the presidential coalition made it to the second round. There are 18 municipalities where the candidate from the presidential coalition ran.

the elections. In this case, warnings to beneficiaries with last-digit NIS from 1 to 5 should have no effect on vote shares, because the elections were held before these warnings were issued. Columns 5 to 8 report the results of this exercise. As expected, we find no significant difference in electoral outcomes between zip codes with a higher-fraction of noncompliant beneficiaries with last-digit NIS from 1 to 5 and those with last-digit NIS above 5. This increases our confidence that our results are not due to some spurious association between the last digit of the NIS and vote shares.

We also analyze whether enforcing BFP requirements affects electoral turnout, which is defined as the ratio of the total number of votes to the electorate. The results are reported in Appendix Table AA2 and show that in zip codes where a higher fraction of noncompliant beneficiaries received warnings before the elections turnout tends to be lower, although the effects is not statistically significant in most cases.[14]

[Insert Table 3]

We next analyze whether BFP enforcement affects the vote share of incumbent mayors who are up for reelection. This is an interesting exercise because *a priori* one would not know whether voters attribute the enforcement of BFP to local authorities or to the federal government. Table 3 shows that voters do not punish local incumbents *per se* when they receive warnings: we only find a significant negative effect for incumbent mayors affiliated with the President's coalition (column 4). These results suggest that voters do not associate BFP enforcement with the local administration, but rather with the national government.[15]

Finally, it is worth mentioning that in order to conduct the analysis in this section, we must restrict the sample to those municipalities that held *second round* municipal elections in 2008, because it is only for the date of the second round elections that we observe a split in the beneficiaries warned before and after. Since this amounts to considering 30 municipalities, one may be concerned about the external validity of our results. To address this concern, we

[14] In Brazil, voting is compulsory for all literate citizens between 18 and 70 years old and non-compulsory for Brazilians aged 16?17 or over 70 or illiterate citizens of any age. Non-voting is only allowed after formally requesting an exemption due to travel or illness. Voters who do not receive an exemption are required to pay a fee equivalent to about 1 U.S. dollar. Voters who do not pay this fee cannot participate in civil service exams or public bidding processes, work in the government, obtain a passport, enrol in a public university, or obtain loans from state banks.

[15] The strong association between BFP and president Lula's party is corroborated by descriptive evidence (e.g., Zucco, 2008; Yoong, 2011).

estimate the correlation between electoral outcomes in the *first round* of the 2008 municipal elections (hence considering the universe of Brazilian municipalities) and BFP enforcement, as measured by the fraction of beneficiaries that received a warning or penalty in July 2008, that is the last warning period before the elections. The results are reported in Appendix Table A4 and confirm that there is a negative correlation between enforcement and vote share of the President's party or coalition, consistent with the results obtained for the second round of elections using our preferred methodology.[16]

4 Do Politicians Manipulate the Enforcement of Program Requirements Close to Elections?

The results discussed in the previous section show that program enforcement affects voter behavior. More specifically, program enforcement before the elections generates electoral losses for mayoral candidates associated with the ruling party. This could create incentives for mayors to reduce enforcement in the run-up to the election. However, manipulating enforcement could be costly in a number of ways. First, it would require the mayor to exert effort. Second, and more important, manipulation could have financial costs because the national government provides additional funding to schools and municipalities based on the quality of program enforcement at the local level.[17] Given these potential costs, we would expect that those mayors that have relatively strong electoral incentives should be more likely to engage in manipulation. We focus on the possibility of running for reelection as a measure of the intensity of local electoral incentives. Mayors in Brazil are only allowed to run for a consecutive term once, so we analyze whether the enforcement of BFP conditionality before elections differs between first and second term mayors.

[16]Regressions in Appendix Table A4 are at the zip code level considering, alternatively, the sample of municipalities where a candidate affiliated with PT or the presidential coalition runs. Controls (not shown) include state dummies, the ratio of beneficiaries to electorate and the ratio of noncompliant beneficiaries to electorate. The results are similar if we include micro-region fixed effects instead of state fixed effects. We do not include municipality fixed effects because 35 percent of municipalities in this sample have only one zip code. However, if we restrict the sample to municipalities with more than one zip code and include municipality fixed effects, the coefficient on *Fraction Warned* remains negative and significant.

[17]To improve the quality of the enforcement of BFP, the central government developed an index (*Indice de Gestão Descentralizada*) measuring the qualiy of implementation, which is used to allocate funds to municipal administrations (*Fundo Municipal de Assistência Social*).

4.1 Methodology

A simple comparison between municipalities with first and second term mayors will probably generate biased estimates due to endogeneity issues. For instance, municipality-specific characteristics such as voter preferences or demographics could affect both whether the mayor is on her first or second mandate and the quality of enforcement. To deal with this challenge and address the presence of both time-invariant and time-varying confounding factors, we implement a regression discontinuity design (RDD) in the spirit of Lee (2008). Exploiting the fact that in Brazil mayors can run only for one consecutive term, we analyze those municipalities where the incumbent ran for re-election in the previous election and compare municipalities where incumbents won by a narrow margin with municipalities where they lost by a narrow margin. This approach provides quasi-random assignment of first term mayors (municipalities where incumbents barely lost re-election in the previous election and a new mayor was elected) and second term mayors (municipalities where incumbents barely won re-election in the previous election). Close races are a relevant context for our study because incentives for manipulation may be higher, the higher the uncertainty of the electoral outcome.

As in the previous section, to measure BFP enforcement we consider, alternatively, the fraction of beneficiary families that receive any warnings or penalties (*fraction warned*) and the fraction of beneficiary families that receive a warning that entails losing the transfer (*fraction warned loss*) during the election year.[18] For this analysis, we restrict our sample to municipalities where the incumbent mayor during the mandate 2000-2004 (2005-2008) runs for re-election in 2004 (2008) and is one of the first two candidates in that election.

We consider the mayoral elections held in 2004 and 2008 to calculate the margin of victory of the non-incumbent mayoral candidate in each municipality i in state s and mandate t (MV_{ist}). At the threshold $MV_{ist} = 0$ there is a sharp change in terms of whether the mayor in power for the mandates 2005-2008 and 2009-2012 is first or second term: for $MV_{ist} < 0$ the incumbent is reelected in 2004 (2008), so she is a second term mayor in the mandate

[18] Specifically, we take averages of the fraction of BFP families that fail to comply with attendance during the electoral year before the elections and get warned in that year. Experimental work on retrospective evaluation suggests that recent events carry more weight (e.g., Langer, Sarin, and Weber, 2005). This, coupled with the high media coverage of BFP for campaign purposes, motivates our focus on enforcement during the election year.

2005-2008 (2009-2012) and cannot run for re-election again. For $MV_{ist} > 0$ the incumbent is not reelected in 2004 (2008) and a new mayor is in power in 2005-2008 (2009-2012), having the possibility to run for re-election in 2008 (2012). Those mayors elected in 2004 (2008) for the first time are eligible to run for reelection in 2008 (2012). In contrast, second term mayors are not allowed to run in 2008 (2012).

MV_{ist} can be considered as a random variable that depends on both observable and unobservable factors, as well as on random events on election day. The average treatment effect (ATE) of having a first term mayor (who can run for reelection) in close elections is given by:

$$E[\tau_{ist}(1) - \tau_{ist}(0)|MV_{ist} = 0] = \lim_{\epsilon \downarrow 0} E[\tau_{ist}|MV_{ist} = \epsilon] - \lim_{\epsilon \uparrow 0} E[\tau_{ist}|MV_{ist} = \epsilon]. \tag{2}$$

where τ is a local effect that captures the impact of being a first term mayor on the outcome only for municipalities around the threshold $MV = 0$ (i.e. for elections that were decided for a margin that is tiny enough).

We estimate the ATE in equation (2) fitting a second order polynomial in MV_{ist} on either side of the threshold $MV_{ist} = 0$:

$$\tau_{ist} = \sum_{k=0}^{2}(\rho_k MV_{ist}^k) + F_{ist}\sum_{k=0}^{2}(\pi_k MV_{ist}^k) + \delta_t + \sigma_s + \varepsilon_{ist}, \tag{3}$$

where MV_{ist} is the margin of victory of the non-incumbent candidate in the mayoral elections in municipality i in state s before term t; and F_{ist} is a dummy variable that equals one if the mayor is a first term mayor (eligible to run for re-election); δ_t are mandate fixed effects; σ_s are state fixed effects. We report heteroskedasticity robust standard errors.[19] The estimated coefficient $\hat{\pi}_0$ identifies the ATE at the threshold $MV_{ist} = 0$. We follow the standard procedure in the literature and show results for a second order polynomial and for local linear regressions with optimal bandwidth computed using the algorithm by Calonico, Cattaneo, and Titiunik (2014).

[19] We also tried clustering standard errors at the state level (there are 26 states in Brazil), adjusting for the small number of clusters using wild bootstrap (Cameron et al., 2008). The results were very similar.

4.2 Results

[Insert Table 4]

Table 4 reports our RDD estimates of the effects of electoral incentives on program enforcement. We show results for three different specifications: (i) OLS regressions that do not control for margin of victory on either side of the cutoff (columns 1 and 4); (ii) RDD regressions described in equation (3) using a second order spline polynomial specification (columns 2 and 5); and (iii) local linear regressions with optimal bandwidth calculated according to Calonico, Cattaneo and Titiunik (2014) (columns 3 and 6). The dependent variable in columns 1 to 3 is the share of BFP beneficiaries who received warnings (*fraction warned)* and in columns 4 to 6 it is the share who received warnings with financial penalties (*fraction warned loss)*. All variables are measured around the 2008 and 2012 municipal elections. Panel A presents results for all the municipalities where the incumbent ran for re-election in the previous election and was one of the first two candidates. Panel B restricts the sample to municipalities with mayors affiliated with the presidential coalition and Panel C presents results for mayors *not* affiliated with the presidential coalition.[20]

We find significant effects of electoral incentives on enforcement, but only in municipalities with mayors affiliated with the presidential coalition (Panel B). This is consistent with the idea that (in close races) mayors from the President's coalition have more incentives to manipulate program enforcement to try to reduce potential electoral costs of program enforcement. Quantitatively, the estimates in column 3 imply that first term mayors from the presidential coalition reduce the fraction of beneficiaries that receive a warning by 0.6 percentage points, which is about 20 percent of the mean of this variable. We find no effects for mayors who are not affiliated with the coalition, consistent with the finding in Section 3 that voters only punish candidates from the presidential coalition.[21]

[20]We split municipalities based on whether mayors belong to the presidential coalition or not, and not based on the President's party, because the sample of municipalities with mayors from the PT that ran for re-election is too small.

[21]Our sample considers all races where the incumbent runs for reelection and ends up being one of the first two candidates in the race. About half of the races in our sample have more than two candidates. We also estimate our main results considering only races with two candidates in cases where the incumbent runs for reelection. Despite the smaller sample size, our main estimates are very similar to those presented in Tables 4 and 7 both in terms of magnitude and significance. This attenuates possible concerns that heterogeneity in races with three or more candidates and two races decided by the same margin might affect our results (Cattaneo et al., 2016).

An interesting question is whether manipulation is confined to periods around elections, or it occurs throughout a mayor's mandate. If the main motive underlying lax enforcement is the fear of losing the votes of noncompliant beneficiaries, one should expect manipulation to be concentrated in the months around the election and not in earlier periods, when penalties may be less salient and easier to be forgotten. We test this hypothesis in Table 5.

[Insert Table 5]

Columns 1 to 4 show regressions similar to those in Table 4, but considering the year before the elections.[22] We find no significant difference between first and second term mayors of the presidential coalition in the year before the election, suggesting that the differences in Table 4 reflect program manipulation strictly linked to electoral reasons.

Finally, we test whether manipulation takes forms other than the enforcement of program conditions, e.g., targeting a broader set of beneficiaries. In columns 5 and 6 of Table 5 we estimate our RDD specification for election years, but we use as dependent variable the ratio of BFP beneficiaries to the electorate in a given municipality (BFP/electorate). We find no effects on the share of beneficiaries, consistent with earlier arguments that there is little or no manipulation in the allocation of program benefits (Fried, 2012).

[Insert Figure 1]

Visual inspection of Figure 1 confirms the results described above. When we consider the sample of municipalities with mayors from the presidential coalition (central graphs) we see visible discontinuities around the cutoff for the fraction of beneficiaries receiving any warning (Panel A) and warnings that entail financial losses (Panel B). We do not find significant discontinuities when considering the full sample of municipalities (leftmost graphs) and when considering the sample of municipalities with mayors that do not belong to the presidential coalition (rightmost graphs).[23] The graphical and non-graphical evidence is robust to the presence of specific outliers.

[22]For conciseness, we only report results for the second order polynomial specification (columns 1 and 3) and local linear regressions with optimal bandwidth (columns 2 and 4).

[23]In Figure 1, outcomes are averaged into bins of intervals of the margin of victory. Bins closer to the cutoff contain more observations, given that the density of our running variable (margin of victory of the first term mayor) is concentrated around zero.

4.3 Validity Checks

The standard RDD assumption is that potential outcomes must be a continuous function of the running variable at the threshold.[24] To formally test this assumption, we test the continuity of the density of the margin of victory, following McCrary (2008). The results for each of our samples are reported in Appendix Figure AA1. We find no evidence of discontinuities in the density of the margin of victory.

Our estimation strategy controls for municipality-specific characteristics. Therefore we should not expect any difference in municipal characteristics between treatment and control groups around the cutoff $MV_{ist} = 0$. To test this, we analyze whether a vast array of observable municipal pre-treatment characteristics are balanced around the cutoff. These characteristics include: (log of) total population, fraction of the population living in an urban area, fraction of houses linked to the water system, fraction of houses linked to the sewerage system, fraction of houses with access to electricity, (log of) per-capita income, and whether there is a local radio station in the municipality.

[Insert Table 6]

Table 6 presents the results of these validity tests for the sample of municipalities with mayors affiliated with the presidential coalition, given that we only find evidence of program manipulation for this sample.[25] Columns 1 to 7 of Panel A report our balance checks for the different municipal characteristics, showing that these characteristics are balanced around the cutoff.

One potential concern regarding the interpretation of our main results is that the difference in program enforcement between first and second term mayors may capture not only

[24] Some recent papers argue that the assumptions of RDD may be violated in recent U.S. House of Representatives elections (e.g., Caughey and Sekhon, 2011). However, Eggers et al. (2015) suggest that in no other case around the world elections exhibit a pattern that violates the RDD assumptions. Recent work on incumbency advantage in Brazilian municipalities (De Magalhães, 2012) also shows that the problem of manipulative sorting identified for the U.S. does not apply to Brazil.

[25] We also conducted the balance tests for the whole sample of municipalities used in the RDD analysis in Table 4 and confirmed that municipal pre-treatment characteristics are balanced around the cutoff. Also note that for the analysis in Table 6 we restrict the sample to those observations used in the local linear regressions with optimal bandwidth presented in Table 4. This way the results in Table 6 are comparable to our main results on the manipulation of program enforcement. Similar results are obtained if, instead, we derive the optimal bandwidth for each specific municipal characteristic or if we consider a second order polynomial including all observations for the coalition sample.

the effects of electoral incentives, but some other differences between first and second term mayors. For example, second term mayors may have higher political ability than first term mayors. If political ability is positively correlated with program enforcement, then the difference between first and second term mayor that we find may reflect differences in ability, not in electoral incentives. Also, second term mayors by construction have more consecutive years of experience in office than first term mayors. If more experienced mayors are better able to enforce program conditions, then our results may reflect differences in experience. Finally, the RDD method does not control for policies implemented after the elections. For instance, first term mayors affiliated with the presidential coalition may implement other policies that increase compliance with the school attendance requirements.

We address these concerns presenting several pieces of evidence. First, if our results were due to differences in ability or experience between first and second term mayors, we should expect to find differences in enforcement levels between these types of mayors in all municipalities. However, our results show significant differences only for the sample of municipalities with mayors affiliated with the presidential coalition.

Second, we checked whether there are significant differences in observable mayoral characteristics around the cutoff.[26] Columns 8 to 12 of Panel A in Table 6 report the results of balance checks for several mayoral characteristics: education, gender, marital status, and whether the individual will be elected for higher office after being a mayor. The results show that for all variables there is no discontinuity around the cutoff.

Third, as mentioned above, second term mayors by construction have more consecutive years of experience in office than first term mayors. So there could still be some concerns that the differences we find may reflect differences in experience and not in electoral incentives. To address this concern we re-estimated all our regressions restricting the sample of first term mayors to those with prior political experience. The results are reported in the Appendix Table AA3. We found that all our results, including those reported in the next section (Table 7), are robust to considering this sample, suggesting that our findings do not reflect differences in political experience between first and second term mayors.

Finally, if our results capture differences in policies implemented after the elections that affect school attendance, we should expect to find differences in other education-related policy

[26]The RDD design controls for municipality-specific characteristics, not for mayoral characteristics.

outcomes around the cutoff. Panel B of Table 6 reports the results of balance checks for several municipality-level educational outcomes: share of public schools linked to the water system, share of public schools that have access to electricity, that have a computer, that have access to internet, (log of) total enrollment in public schools, average teacher/student ratio, average class size, dropout rates, and completion rates separately for primary and high school. The results show that there are no significant differences in education-related policy outcomes between first and second term mayors around the cutoff. This helps ameliorate concerns that the lower noncompliance with the school attendance requirements that we find for first term mayors reflect differences in educational policies, rather than enforcement manipulation.

5 How Do Politicians Manipulate the Enforcement of Program Requirements?

In this section we analyze how program enforcement may be manipulated. Given the monitoring process described in Section 2, in principle authorities could reduce the penalties that beneficiaries receive in at least three ways. First, they could misreport attendance to reduce the number of students who fall below the 85 percent attendance threshold. Second, authorities could increase the number of students whose absence is excused, as excused absences do not count towards noncompliance. Finally, the Ministry in charge of the program (MDS) could directly manipulate warnings and not penalize some beneficiaries who do not meet the attendance requirement.

5.1 Manipulation at the local vs. central level

To analyze which of these channels account for the manipulation we uncovered in Section 4, we use the same identification strategy. Given that in Section 4 we found that manipulation is driven by mayors from the presidential coalition, we focus our analysis on municipalities where the mayor is affiliated with the coalition, and compare municipalities where incumbents won by a narrow margin with municipalities where they lost by a narrow margin. We consider five different dependent variables which reflect the three different forms of program enforcement manipulation described above: (i) the ratio of the number of students below the 85 percent threshold (considering all absences, including those later excused) to the number of BFP students; (ii) the ratio of the number of students below the threshold *whose absence was not excused* to the number of BFP students; (iii) the ratio of the number of students below the threshold whose absence was excused to the number of BFP students; (iv) the ratio of the number of students below the threshold whose absence was excused to the total number of students below the threshold; and (v) the ratio of the number of beneficiary families who received a warning or penalty to the number of beneficiary families who did not meet the attendance requirement (i.e., who attended less than 85 percent of the days and were not excused). If our findings are explained by authorities misreporting attendance, then we would expect variables (i) and (ii) to be lower in municipalities where first term mayors won by a narrow margin. If, instead, manipulation occurs through an increase in the number of students whose absence is excused (without misreporting missed attendance) then we would expect variables (iii) and (iv) to be higher. Finally, if our results are due to federal authorities directly manipulating warnings or penalties, by not penalizing some beneficiaries who do not meet the attendance requirement, we would expect variable (v) to be lower.

[Insert Table 7]

Table 7 contains our results. For each of the five dependent variables described above we report OLS coefficients and two different RDD specifications, one using a second order spline polynomial specification (indicated as POLY2 in the table) and one using a local linear regression considering the optimal bandwidth calculated following Calonico et al. (2014) (indicated as LLR). Columns 1 to 6 of Panel A show that the fraction of students below the 85 percent threshold and the fraction of students who do not meet the attendance requirement (i.e., those below the threshold whose absence was not excused) are both lower in municipalities where first term mayors from the coalition won by a narrow margin. In terms of magnitude, for instance, the estimates in column 3 imply that in municipalities with first term mayors the fraction of BFP students with attendance below the threshold is about 1 percentage point lower (than in municipalities with second term mayors), which is about 27 percent of the sample mean of this variable.

In columns 1 to 3 of Panel B we find a marginally significant negative effect on the ratio of BFP students whose absence was excused, but this seems to be driven by the lower reported missed attendance. In fact, once we use the number of noncompliant students as denominator (columns 4 to 6) we find no effects. We also find no evidence that federal authorities are directly manipulating warnings or penalties (columns 7 to 9). Overall, the results in Table 7 suggest that manipulation of program enforcement occurs mostly through misreporting of attendance.

Visual inspection of the outcomes in Figure 2 confirms the above results. There are visible discontinuities around the cut-off for the variables *Noncompliant /BFP students* and *Noncompliant & not excused/BFP students* (top panels in the figure) and no significant discontinuity for *Excused/Noncompliant* and *Warned/Noncompliant not excused* (bottom panels).[27] The graphical and non-graphical evidence is robust to the presence of specific outliers.

As a final check, we considered that if mayors manipulate enforcement to try to avoid electoral costs, we should not expect to find effects when elections are not close in time. To test this, we re-estimated the regressions in Table 7 for the year 2011 (municipal elections were held in 2012). We found no significant difference between first and second term mayors for this election-free year, consistent with the interpretation that the differences we uncover reflect manipulation due to electoral reasons.[28]

5.2 Politically connected school principals

To provide further evidence on how enforcement may be manipulated, we explore an additional source of heterogeneity across municipalities. In particular, the above results suggest that the bulk of manipulation consists in misreporting attendance, a task for which schools play a crucial role. We can then test if enforcement is weaker before the elections in municipalities where a large fraction of beneficiaries attend schools that are closely linked to the local government.

To measure whether schools are linked to the local government, we focus on how school principals where appointed. Traditionally, principals of municipal schools in Brazil were politically appointed: in fact, the position of school principal has been a well known source of patronage for politicians (Plank, 1996). Currently, there are four ways of selecting principals:

[27] In this figure, outcomes are averaged into bins of intervals of the margin of victory. Bins closer to the cut-off contain many more observations, given that the density of our running variable (margin of victory of the first term mayor) is concentrated around zero.

[28] Note that since attendance data are not available before 2008, we cannot conduct this test for 2007.

political appointment, public competition, election, and selective election. Political appointment is at the discretion of local politicians and is typically excused with the argument that the role of school principal is a " position of trust" , so politicians may be entitled to choose someone they trust for this position. Public competitions are instead decided by written exams open to any eligible candidate. In the case of elected principals, all decisions are made by the school community. Selective elections are a special case where candidates are first chosen through a competitive written exam and then voted on by the school community. On average, about 12 percent of the public schools in a given municipality in our sample have a politically appointed principal.

For the purpose of our analysis, the key idea is that politically appointed principals may be indebted to the politicians who selected them, and as a result they may be more susceptible to political pressures. For example, Akhtari, Moreira and Trucco (2016) show that party turnover in municipal elections significantly increases the replacement of school principals.

To analyze whether political connections of school principals play a role in the manipulation of BFP enforcement, we split the sample of municipalities in two groups. In particular, we define municipalities with high (low) intensity of politically connected schools as those where the fraction of BFP students enrolled in schools with politically appointed principals is above (below) the median. We then conduct the same analysis of manipulation as in Table 7 on these two subsamples.

[Insert Table 8]

Table 8 focuses on misreporting of attendance, which –based on our previous results– is the main mechanism for manipulation. In columns 1 to 3 the dependent variable is the share of BFP beneficiary students who fail to meet the attendance threshold (including those later excused), while in columns 4 to 6 only noncompliant students who are not excused are considered. For both outcomes, the negative effect of first term mayors is concentrated in municipalities with a high fraction of BFP students in politically connected schools (Panel A), while no significant effect is found for the low-connections sample (Panel B). This is consistent with the idea that the former set of schools may be more susceptible to political pressures.

A potential concern with the above analysis is that the distinction between municipalities with high and low intensity of politically connected schools may reflect not only how closely linked to the local government schools are, but also differences in other characteristics that affect program enforcement (e.g., differences in institutional capacity). We address this concern in two ways. First, we test for differences in observable characteristics. We estimate a linear probability model for whether the municipality has high politically connected schools as a function of a vast array of observable municipal characteristics and education-related policy outcomes. The results, reported in Appendix Table AA4, show that most variables are not statistically significant, with the exception of municipality size and urbanization.

Second, if the difference in school attendance between municipalities with high and low intensity of politically connected schools reflects some other underlying unobservables, then we should expect to find these differences for all periods. On the other hand, if this difference reflects manipulation of enforcement for electoral reasons, then we should expect to find it mainly in the election year. To test this, we estimated the RDD regressions for municipalities with a high and a low fraction of students in politically connected schools for the year before the elections. Given that we do not have attendance data before 2008, we conducted this placebo test only for the 2012 elections. The results, reported in Appendix Table AA5 (columns 1 to 4), show that, consistent with the results in Table 8, in the electoral year (2012) first term mayors are associated with a lower share of BFP beneficiary students who fail to meet the attendance requirements in those municipalities with a high fraction of BFP students in politically connected schools (Panel A), while no significant effect is found for the low-connections sample (Panel B). However, the results in columns 5 to 8 show that there is no significant difference between first and second term mayors of the presidential coalition in the year before the election (2011) for none of the samples, suggesting that the results in Table 8 reflect program manipulation due to electoral reasons.

6 Conclusions

Our paper focuses on the implementation of conditional programs and analyzes whether politicians manipulate the enforcement of program conditions to influence electoral outcomes. We analyze the system of conditionality of the Bolsa Familia Program (BFP), the largest conditional cash transfer program in the world. Our working hypothesis is that, if targeted government programs affect voter choices, then politicians may have incentives to strategically manipulate their enforcement, particularly when they face stronger electoral incentives.

To test this, we first analyze whether voters respond to the enforcement of program rules. Exploiting random variation in the timing when different beneficiaries learn about penalties for noncompliance, we find that beneficiaries respond negatively to stricter enforcement of BFP rules and that they associate this enforcement with the national government. Second, we study whether politicians strategically become more lenient around the time of elections, to avoid electoral costs of enforcement. Using a regression discontinuity design, we find that the fraction of beneficiaries who receive warnings or penalties is lower in municipalities where mayors from the President's coalition face reelection incentives. We also provide evidence on possible mechanisms: manipulation of enforcement seems to occur mostly through misreporting of attendance, and this effect is driven by municipalities where a high fraction of BFP students attend schools with politically connected principals.

Our results have relevant policy implications, as the manipulation of enforcement has the potential to reduce the effectiveness of welfare programs which include conditionality. In the limit, if voters internalize the incentives for manipulation, programs that are nominally conditional may *de facto* perform as if they were unconditional. Another interesting implication regards the ability of the relevant stakeholders to detect misbehavior by the authorities. The classic form of manipulation studied in the case of targeted transfer programs is mistargeting, in particular giving benefits to ineligible people. This form of manipulation is relatively easier to detect than the form we study in this paper: targeting errors may be very visible to neighbors who know the family, and could be uncovered by an audit scheme (of the type used for Brazilian municipalities). On the other hand, detecting misreported school attendance or missed warnings would require a more sophisticated analysis, one that keeps track of the entire attendance and payment history of the family. Our study suggests that, as the structure of transfer schemes becomes more complex (e.g., with an articulated system of requirements to qualify for the benefits), so does the sophistication of political actors who wish to manipulate such schemes for electoral purposes. Concentrating auditing efforts around election times (for example auditing schools and their attendance reporting practices) may limit the potential for the manipulation mechanisms uncovered in our paper. Whether such policies would be effective is an open question for future research.

References

Aidt, T.S. and Shvets, J., 2012. "Distributive politics and electoral incentives: Evidence from seven US state legislatures". *American Economic Journal: Economic Policy*, 4(3), pp.1-29.

Akhtari, M., D. Moreira and L. Trucco (2016), "Political Turnover, Bureaucratic Turnover, and the Quality of Public Services", mimeo, Harvard University.

Alesina, A., 1988. "Macroeconomics and politics". In NBER Macroeconomics Annual 1988, Volume 3 (pp. 13-62). MIT Press.

Alt, J., De Mesquita, E.B. and Rose, S., 2011. "Disentangling accountability and competence in elections: evidence from US term limits". *The Journal of Politics*, 73(01), pp.171-186.

Baird, S., C. McIntosh, and B. Özler, 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment", *The Quarterly Journal of Economics*, 126 (4), p. 1709-1753.

Baird, S., F. Ferreira, B. zler; M.. 2013. "Relative Effectiveness of Conditional and Unconditional Cash Transfers for Schooling Outcomes in Developing Countries: A Systematic Review", Campbell Systematic Reviews 9 (8): 1?124.

Besley, T. and Case, A., 1995. "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits" *The Quarterly Journal of Economics*, 110(3), pp.769?798.

Brollo, F., K. Kaufmann, and E. La Ferrara, 2015. "Learning about the Enforcement of Conditional Welfare Programs: Evidence from Brazil", mimeo.

Brollo, F. and T. Nannicini, 2012. "Tying Your Enemy's Hands in Close Races: The Politics of Federal Transfers in Brazil" . *American Political Science Review*, 106, p. 742-761.

Calonico, Sebastian, Matias Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82 (6), p. 2295?2326.

Matias D. Cattaneo, Luke Keele, Roco Titiunik, and Gonzalo Vazquez-Bare, 2016. "Interpreting Regression Discontinuity Designs with Multiple Cutoffs," The Journal of Politics 78, no. 4: 1229-1248.

Cameron, A.C., Gelbach, J.B. and Miller, D.L., 2008. "Bootstrap-based improvements for inference with clustered errors". *The Review of Economics and Statistics*, 90(3), pp.414-427.

Caughey, Devin and Jasjeet S Sekhon. 2011. "Elections and the Regression Discontinuity Design: Lessons from Close US House Races, 1942-2008." *Political Analysis* 19(4):385-408.

De Janvry, Alain, Frederico Finan, and Elisabeth Sadoulet., 2012. "Local Electoral Accountability and Decentralized Program Performance." *Review of Economics and Statistics* 94 (3), pp. 672-685.

De Janvry, A., Finan, F., Sadoulet, E. and Vakis, R., 2006. “Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks?”. *Journal of Development Economics*, 79(2), pp.349-373.

De La O, A. L. (2013), Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico. American Journal of Political Science, 57: 1?14.

De Magalhaes, L., 2012. “Incumbency Effects in Brazilian Mayoral Elections: A Regression Discontinuity Design”, Working Paper.

Drazen, A., 2001. “The political business cycle after 25 years”. In *NBER Macroeconomics Annual 2000*, Volume 15 (pp. 75-138). MIT Press.

Eggers, A.C., Fowler, A., Hainmueller, J., Hall, A.B. and Snyder, J.M., 2015. “On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races”. *American Journal of Political Science*, 59(1), pp.259-274.

Ferraz, C. and Finan, F., 2011. “Electoral accountability and corruption: Evidence from the audits of local governments”. *The American Economic Review*, 101(4), pp.1274-1311.

Fiorina, M.P., 1978. “Economic retrospective voting in American national elections: A micro-analysis”. *American Journal of Political Science*, pp.426-443.

Finan, F. and L. Schechter (2012), “Vote-Buying and Reciprocity,” *Econometrica*, 80(2): 863-881.

Fried, B. J., 2012. “Distributive Politics and Conditional Cash Transfers: The Case of Brazil’s Bolsa Familia” *World Development* 40, (5), pp. 1,042-1,053

Green T. 2006. “The Political Economy of a Social Transfer Program: Evidence on the Distribution of PROGRESA in Mexico”, mimeo, U,C., Berkeley.

Hainmueller, J. and Kern H. L., 2008. “Incumbency as a source of spillover effects in mixed electoral systems: Evidence from a regression-discontinuity design”, *Electoral Studies*, 4.

Healy, A. and N. Malhotra (2010), “Random Events, Economic Losses, and Retrospective Voting: Implications for Democratic Competence,” *Quarterly Journal of Political Science,* 5(2): 193-208.

Labonne, J. 2016. “Local Political Business Cycles: Evidence from Philippine Municipalities”. *Journal of Development Economics*, Vol. 121: 56-62.

Labonne, J. 2013. “The Local Electoral Impacts of Conditional Cash Transfers: Evidence from a field experiment”. *Journal of Development Economics*, 2013, Vol. 104: 73-88.

Langer, T., R. Sarin, and M. Weber. 2005. “The Retrospective Evaluation of Payment Sequences: Duration Neglect and Peak-and-End Effects.” *Journal of Economic Behavior and Organization* 58(1): 157–75

Lee, David S. 2008. “Randomized Experiments from Non-random Selection in the U.S. House Elections.” *Journal of Econometrics* 142: 675-697.

Lee, Davir, Enrico Moretti, and Matthew Butler. 2004. "Do voters affect or elect policies? Evidence from the U.S. House." *Quarterly Journal of Economics* 119 (3), 807-859.

Levitt, S.D. and Snyder Jr, J.M., 1995. "Political parties and the distribution of federal outlays". *American Journal of Political Science*, pp.958-980.

Levitt S.D. and J.M. Snyder, 1997. "The Impact of Federal Spending on House Election Outcomes", *Journal of Political Economy*, 105, (1), p. 30-53.

Manacorda, M., E. Miguel, and A. Vigorito, 2011. "Government Transfers and Political Support." *American Economic Journal: Applied Economics*, 3,(3), p. 1-28.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142: 698-714.

Plank, D. 1996. "The means of our salvation: Public education in Brazil, 1930? 1995." Boulder, CO: Westview Press.

Rodríguez-Chamussy, L. (2015), "Local Electoral Rewards from Centralized Social Programs: Are Mayors Getting the Credit?", IDB Working Paper No. 550.

Wolfers, J. (2007), "Are Voters Rational? Evidence from Gubernatorial Elections", mimeo, University of Pennsylvania.

Yoong, P. S. (2011), "Buying Out the Poor? Bolsa Familia and the 2010 Elections in Brazil", mimeo.

Solé-Ollé, A. and Sorribas-Navarro, P., 2008. "The effects of partisan alignment on the allocation of intergovernmental transfers. Differences-in-differences estimates for Spain". *Journal of Public Economics*, 92(12), pp.2302-2319.

Zucco Jr., Cesar., 2008. "The President's 'New' Constituency: Lula and the Pragmatic Vote in Brazil's 2006 Presidential Election", *Journal of Latin American Studies* , 40, (1), p. 29-55.

Table 1: Summary Statistics

	obs	mean	sd.
Panel A: Zip code level, second round elections sample			
Vote share coalition	1,395	0.4897	0.1495
Fraction warned	1,395	0.1110	0.1946
Fraction warned treated	1,395	0.0488	0.1091
Fraction warned with financial loss	1,395	0.0623	0.1512
Fraction warned with financial loss treated	1,395	0.0277	0.0857
Panel B: Municipality level, RDD full sample			
Noncompliant/BFP students	4,351	0.0333	0.0357
Excused/BFP students	4,351	0.0122	0.0248
Excused/Noncompliant	4,073	0.2966	0.2782
Noncompliant & not excused/BFP students	4,351	0.0218	0.0236
Warn/BFP families	4,351	0.0285	0.0295
Warn/Noncompliant & not excused	3,949	0.9229	0.1168
Warn loss/BFP families	4,351	0.0144	0.0179
Warn loss/Noncompliant & not excused	3,949	0.3957	0.2144
BFP families/electorate	4,351	0.1382	0.0792

Notes. Panel A considers observations at the zip code level for the sample of municipalities where a candidate from the presidential coalition runs in the second round of October 2008 municipal elections. Panel B considers the full RDD sample, taking municipal level averages during the electoral year. See Table A1 for the definition of the variables.

Table 2: The Effects of Warnings on PT and Presidential Coalition's Vote Share

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Warnings in October - month of municipal elections				Placebo: warnings in November - month after elections			
	Effects of warnings		Effects of warnings with financial loss		Effects of warnings		Effects of warnings with financial loss	
	share PT	share coal	share PT	share coal	share PT	share coal	share PT	share coal
Fraction treated	-0.1234**	-0.0936**	-0.1995***	-0.1514***	-0.0006	-0.0229	-0.0455	-0.0772
	(0.0497)	(0.0468)	(0.0610)	(0.0570)	(0.0575)	(0.0600)	(0.0440)	(0.0518)
Fraction warned	-0.0276	-0.0224	-0.0465	-0.0473	-0.0320	0.0360	-0.0369	0.0399
	(0.0294)	(0.0279)	(0.0368)	(0.0338)	(0.0302)	(0.0331)	(0.0283)	(0.0308)
Observations	1,130	1,395	1,130	1,395	1,130	1,395	1,130	1,395
Municipality fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Sample - municipalities where:	PT runs	coalition runs	PT runs	coalition runs	PT runs	coalition runs	PT runs	coalition runs

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Columns 1, 2, 5 and 6 consider all types of warnings; columns 3, 4, 7 and 8 only consider warnings that imply financial losses (warning stages 2 to 5). The dependent variable in columns 1, 3, 5 and 7 (2, 4, 6 and 8) is the vote share of the candidate affliated with the PT (with the presidential coalition). All regressions include municipality fixed effects and the following controls: the ratio of beneficiaries to electorate, the ratio of noncompliant beneficiaries to electorate and the fraction of beneficiaries with last digit of their NIS from 1 to 5.

Table 3: The Effects of Warnings on Incumbents' Vote Share

	(1)	(2)	(3)	(4)	(5)	(6)
	Effects of warnings			Effects of warnings with financial loss		
	share inc. PT	share inc. no PT	share inc. no coal	share inc. PT	share inc. no PT	share inc. no coal
Fraction treated	-0.1617	-0.0181	-0.0282	-0.3375***	0.0505	0.0604
	(0.0996)	(0.0562)	(0.0544)	(0.1173)	(0.1098)	(0.1074)
Fraction warned	-0.0314	0.0267	0.0437	0.0005	-0.0310	-0.0182
	(0.0496)	(0.0303)	(0.0272)	(0.0546)	(0.0612)	(0.0598)
Observations	126	523	431	126	523	431
Municipality fixed effects	yes	yes	yes	yes	yes	yes
Sample - Municipalities where:	incumbent PT runs	incumbent non PT runs	incumbent non coalition runs	incumbent PT runs	incumbent non PT runs	incumbent non coalition runs

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Columns 1, 2 and 3 consider all types of warnings; columns 4, 5 and 6 only consider warnings that imply financial losses (warning stages 2 to 5). The dependent variable in columns 1 and 4 is the vote share of the incumbent mayor who is affiliated with the PT; in columns 2 and 5 it is the vote share of the incumbent mayor who is not affiliated with PT; in columns 3 and 6 it is the vote share of the incumbent mayor who is not affiliated with the presidential coalition. All regressions include municipality fixed effects and the following controls: the ratio of beneficiaries to electorate, the ratio of noncompliant beneficiaries to electorate and the fraction of beneficiaries with last digit of their NIS from 1 to 5.

Table 4: The Effects of Electoral Incentives
on Warnings, RDD estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Warned/BFP families			Warned loss/BFP families		
	OLS	POLY 2	LLR	OLS	POLY 2	LLR
Panel A: Full Sample						
First term mayor	-0.0015	-0.0018	-0.0017	-0.0006	-0.0008	-0.0011
	(0.0010)	(0.0014)	(0.0018)	(0.0006)	(0.0008)	(0.0011)
Observations	4,351	4,351	2,668	4,351	4,351	2,713
h			0.147			0.153
Panel B: Mayors affiliated with the presidential coalition						
First term mayor	-0.0026	-0.0052**	-0.0064**	-0.0012	-0.0028**	-0.0033*
	(0.0016)	(0.0022)	(0.0030)	(0.0009)	(0.0013)	(0.0017)
Observations	1,817	1,817	1,045	1,817	1,817	1,035
h			0.139			0.138
Panel C: Mayors not affiliated with the presidential coalition						
First term mayor	-0.0007	0.0008	0.0023	-0.0001	0.0007	0.0005
	(0.0014)	(0.0018)	(0.0024)	(0.0008)	(0.0011)	(0.0014)
Observations	2,534	2,534	1,501	2,534	2,534	1,548
h			0.137			0.143

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. The dependent variable in columns 1-3 is the share of BFP families who received a warning of any type; in columns 4-6 it is the share of BFP families who received a warning that implies financial losses (warning stages 2 to 5). Columns 1 and 4 display OLS estimates not controlling for margin of victory on either side of the cut-off. Columns 2 and 5 display RDD estimates using a second order spline polynomial specification. Columns 3 and 6 display RDD estimates for local linear regressions with optimal bandwidth calculated according to Calonico et al. (2014). h denotes the interval of our running variable (for example: h=0.10 represents races where margin of victory is between -10% and 10%). All regressions include mandate and state fixed effects.

Table 5: The effects of Electoral Incentives on Warnings –
Placebo and Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
	Placebo: year before elections				Robustness check	
	warned/ BFP beneficiaries		warned loss/ BFP beneficiaries		BFP beneficiaries/ electorate	
	POLY 2	LLR	POLY 2	LLR	POLY 2	LLR
Panel A: Full sample						
First term mayor	-0.0009	-0.0011	-0.0006	-0.0005	0.0017	-0.0011
	(0.0014)	(0.0018)	(0.0010)	(0.0013)	(0.0025)	(0.0031)
Observations	4,351	2,668	4,351	2,713	4,351	2,867
h						0.166
Panel B: Mayors affiliated with the presidential coalition						
First term mayor	-0.0024	-0.0028	-0.0014	-0.0017	0.0034	0.0035
	(0.0022)	(0.0030)	(0.0014)	(0.0020)	(0.0039)	(0.0049)
	[0.216]	[0.304]	[0.248]	[0.310]	[0.426]	[0.434]
Observations	1,817	1,045	1,817	1,035	1,817	1,190
h						0.173
Panel C: Mayors not affiliated with the presidential coalition						
First term mayor	0.0002	-0.0012	0.0001	-0.0003	0.0001	-0.0046
	(0.0019)	(0.0025)	(0.0013)	(0.0017)	(0.0033)	(0.0043)
Observations	2,534	1,501	2,534	1,548	2,534	1,485
h						0.135

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. The dependent variable in columns 1 to 4 is the share of BFP families who received warnings in the year before the election; in columns 5 and 6 it is the ratio of beneficiary families to electorate. Columns 1, 3 and 5 display results using a second order spline polynomial specification; columns 2, 4 and 6 display RDD estimates for local linear regressions with optimal bandwidth calculated according to Calonico et al. (2014). h denotes the interval of our running variable (for example: h=0.10 represents races where margin of victory is between -10% and 10%). Panel A considers the full RDD sample. Panel B (Panel C) considers the RDD sample only for municipalities with mayors affliated (not affiliated) with the presidential coalition. All regressions include mandate and state fixed effects.

Table 6: Balance Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Municipal and mayoral characteristics												
	Municipal characteristics							Mayoral characteristics				
	population	urban	water	sewerage	electricity	income	radio	primary	college	male	married	future political career
First term mayor	0.0357	-0.0212	0.0198	0.0105	-0.0041	0.0031	-0.0013	0.0044	0.0425	-0.0286	-0.0558	-0.0041
	(0.1048)	(0.0231)	(0.0232)	(0.0185)	(0.0134)	(0.0439)	(0.0527)	(0.0390)	(0.0578)	(0.0359)	(0.0494)	(0.0278)
Observations	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045
Panel B: Education-related policy outcomes												
	school with water	school with electricity	school with computer	school with internet	enrollment	teacher/ student	primary class size	primary dropout rates	primary completion rates	high class size	high dropout rates	high completion rates
First term mayor	-0.0012	0.0001	0.0186	-0.0268	-0.0021	-0.2088	-0.2607	0.0508	-0.0813	-0.1640	-0.0813	0.2538
	(0.0270)	(0.0130)	(0.0259)	(0.0254)	(0.0025)	(0.4072)	(0.3973)	(0.2533)	(0.6367)	(0.5672)	(0.6367)	(0.7052)
Observations	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045	1,045

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Panel A reports balance checks for municipality pre-determined characteristics and mayoral characteristics. Panel B reports balance checks for outcomes related to education policy. The sample of municipalities is that of Table 4, column 3 of panel B. All regressions include mandate and state fixed effects. See Appendix Table 1A for variables definition.

Table 7: Manipulation at the Local vs. Federal Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	POLY 2	LLR	OLS	POLY 2	LLR	OLS	POLY 2	LLR
Panel A: Attendance									
	noncompliant/ BFP beneficiaries			noncompliant not excused/ BFP beneficiaries					
First term mayor	-0.0028	-0.0076***	-0.0086**	-0.0022*	-0.0043***	-0.0051**			
	(0.0019)	(0.0026)	(0.0037)	(0.0012)	(0.0017)	(0.0024)			
Observations	1,817	1,817	1,088	1,817	1,817	1,048			
h			0.147			0.139			
Panel B: Excuses and warnings									
	excused/ BFP beneficiaries			excused/ noncompliant			warned/ noncompliant not excused		
First term mayor	-0.0006	-0.0035*	-0.0034	0.0277	0.0066	-0.0093	0.0032	0.0056	0.0036
	(0.0015)	(0.0018)	(0.0024)	(0.0170)	(0.0230)	(0.0298)	(0.0063)	(0.0081)	(0.0107)
Observations	1,817	1,817	1,208	1,727	1,727	1,103	1,688	1,688	1,061
h			0.178			0.166			0.162

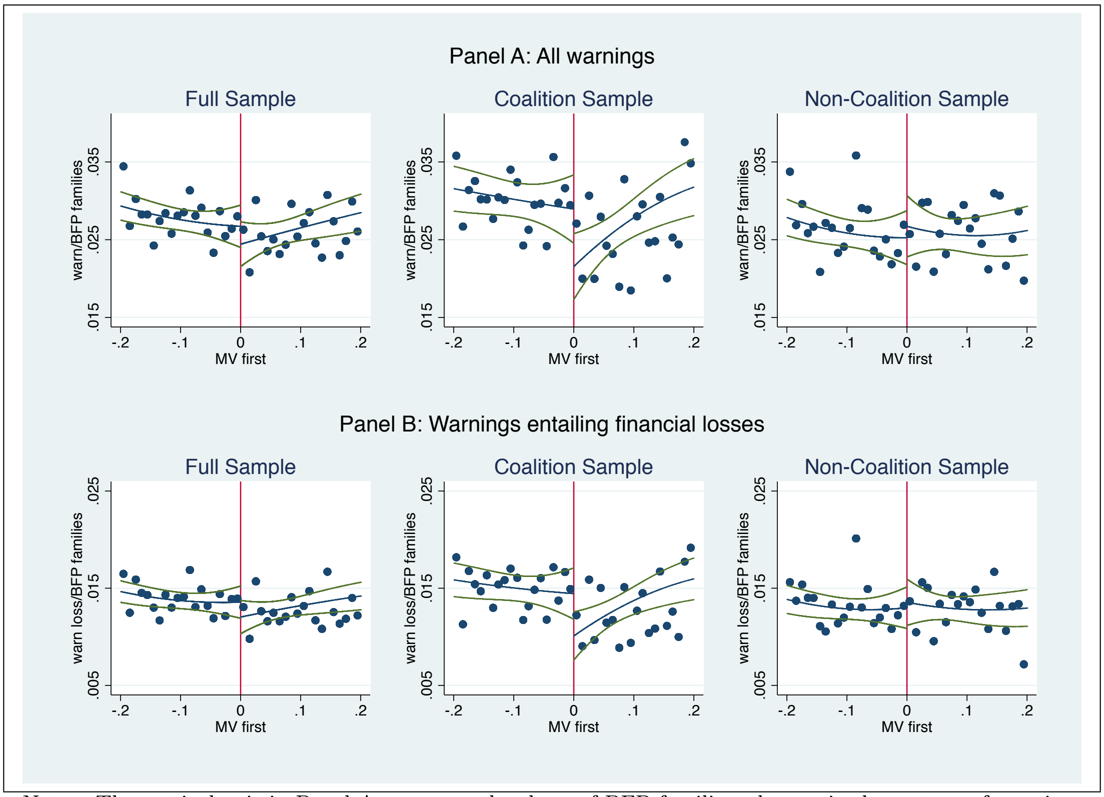
Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Dependent variables are: in columns 1-3 of Panel A the ratio of noncompliant to BFP beneficiary students; in columns 4-6 of Panel A the ratio of noncompliant and not excused to BFP beneficiary students; in columns 1-3 of Panel B the ratio of excused to BFP beneficiary students; in columns 4-6 of Panel B the ratio of excused to noncompliant students; in columns 7-9 the ratio of warned to noncompliant and not excused students. Columns 1, 4 and 7 display OLS estimates not controlling for margin of victory on either side of the cut-off. Columns 2, 5 and 8 display RDD estimates using a second order spline polynomial specification. Columns 3, 6 and 9 display RDD estimates for local linear regressions with optimal bandwidth calculated according to Calonico et al. (2014). h denotes the interval of our running variable (for example: h=0.10 represents races where margin of victory is between -10% and 10%). All regressions include mandate and state fixed effects.

Table 8: Attendance and Political Connections of School Principals

	(1)	(2)	(3)	(4)	(5)	(6)
	noncompliant/ BFP beneficiaries			noncompliantnot excused/ BFP beneficiaries		
	OLS	POLY 2	LLR	OLS	POLY 2	LLR
Panel A: Municipalities with high intensity of politically connected schools						
First term mayor	-0.0037	-0.0090***	-0.0099**	-0.0026*	-0.0058***	-0.0048
	(0.0024)	(0.0032)	(0.0043)	(0.0015)	(0.0021)	(0.0029)
Observations	939	939	607	939	939	554
h			0.156			0.132
Panel B: Municipalities with low intensity of politically connected schools						
First term mayor	-0.0016	-0.0061	-0.0056	-0.0019	-0.0029	-0.0033
	(0.0032)	(0.0043)	(0.0058)	(0.0020)	(0.0027)	(0.0035)
Observations	878	878	541	878	878	552
h			0.170			0.177

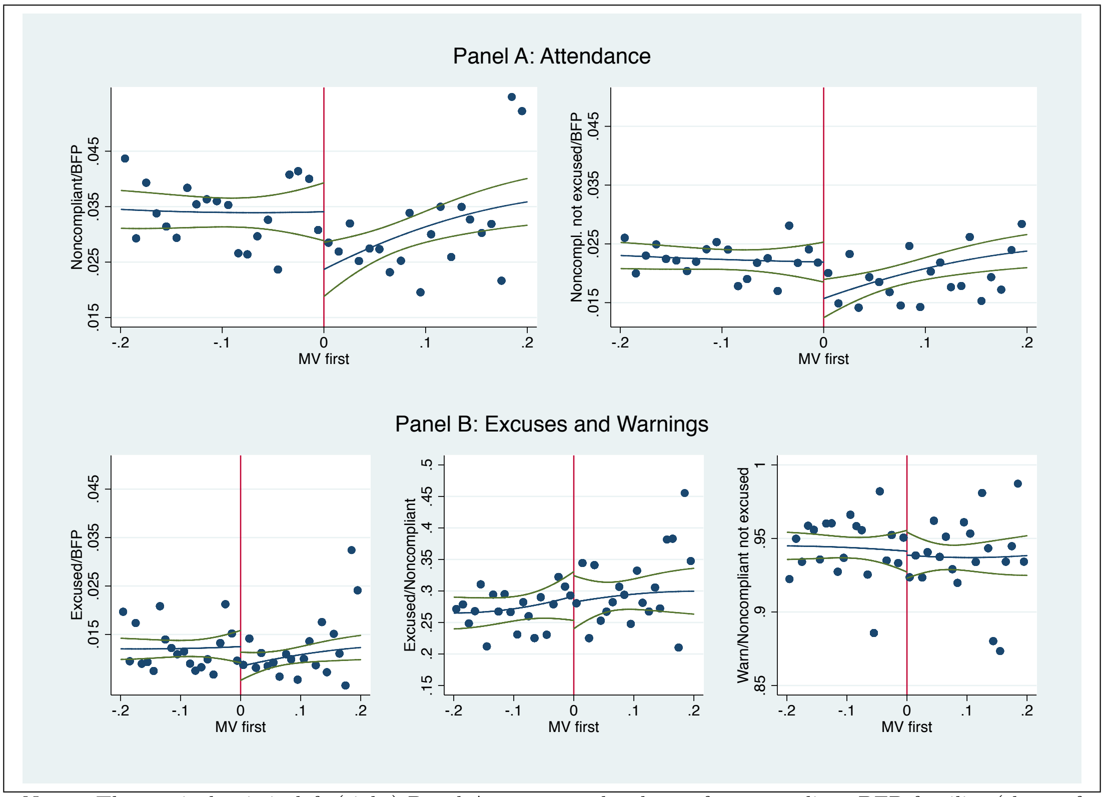
Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. The dependent variable in columns 1-3 is the ratio of noncompliant to BFP beneficiary students; in columns 4-6 it is the ratio of noncompliant and not excused to BFP beneficiary students. Columns 1 and 4 display OLS estimates not controlling for margin of victory on either side of the cut-off; columns 2 and 5 display RDD estimates using a second order spline polynomial specification; columns 3 and 6 display RDD estimates for local linear regressions with optimal bandwidth calculated according to Calonico et al. (2014). h denotes the interval of our running variable (for example: h=0.10 represents races where margin of victory is between -10% and 10%). Panel A (Panel B) reports results for the sample of municipalities where the fraction of students enrolled in schools with politically appointed principals is above (below) the median. The sample includes municipalities affiliated with the presidential coalition. All regressions include mandate and state fixed effects.

Figure 1: The Effects of Electoral Incentives on Warnings



Notes. The vertical axis in Panel A measures the share of BFP families who received any type of warning; in Panel B only warnings that imply financial losses. The horizontal axis displays the margin of victory (MV_{ist}) of the non-incumbent candidate in municipality i in state s and mandate t. The blue line is a split second-order polynomial in MV fitted separately on each side of the MV threshold at zero. $MV_{ist} > 0$ when the winner candidate in municipality i in state s and mandate t is the opponent candidate. $MV_{ist} < 0$ when the winner candidate in municipality i state s and mandate t is the incumbent candidate. Full sample considers races in 2004 and 2008 municipal elections when the incumbent runs for re-election and is one of the first two candidates. Coalition (Non-coalition) sample considers municipalities where the mayor is affiliated (not affiliated) with the presidential coalition.

Figure 2: The Effects of Electoral Incentives on Enforcement Outcomes



Notes. The vertical axis in left (right) Panel A measures the share of noncompliant BFP families (share of noncompliant and not excused BFP families). The vertical axis in left (middle) (right) Panel B measures the share of excused BFP families (ratio of excused to noncompliant BFP families) (ratio warned to noncompliant not excused BFP families). The horizontal axis displays the margin of victory (MV_{ist}) of the non-incumbent candidate in municipality i in state s and mandate t. The blue line is a split second-order polynomial in MV fitted separately on each side of the MV threshold at zero. $MV_{ist} > 0$ when the winner candidate in municipality i and mandate t is the opponent candidate. $MV_{ist} < 0$ when the winner candidate in municipality i in state s and mandate t is the incumbent candidate. Full sample considers races in 2004 and 2008 municipal elections when the incumbent runs for re-election and is one of the first two candidates. All regressions consider the sample of municipalities where the mayor is affiliated with the presidential coalition..

Table A1: Variable Definition

Panel A: Enforcement variables	
Fraction warned	Fraction of families in the program who received a warning or penalty.
Fraction treated	Fraction of families in the program with last digit NIS from 1 to 5 who received a warning or penalty in the week before the second round of the October 2008 municipal elections.
Noncompliant/BFP families	Ratio of the number of families in the program with at least one child below the 85 percent school attendance threshold to number of families in the program.
Noncompliant/BFP students	Ratio of the number of beneficiary students below the 85 percent school attendance threshold to number of students in the program.
Excused/BFP students	Ratio of the number of beneficiary students below the 85 percent school attendance threshold to number of students in the program.
Excused/fail	Ratio of the number of beneficiary students below the 85 percent school attendance threshold whose absence was excused to the total number of students in the program.
Noncompliant not excused/BFP students	Ratio of the number of beneficiary students below the 85 percent school attendance threshold whose absence was excused to the total number of beneficiary students below the 85 percent school attendance threshold.
Warned/BFP families	Ratio of the number of families who received a warning or penalty to the number of families in the program.
Warned/Noncompliant not excused	Ratio of the number of beneficiary families who received a warning or penalty to the number of beneficiary families who did not meet the attendance requirement.
Warned loss/BFP families	Ratio of the number of families who received a penalty to the number of families in the program.
Warned loss/Noncompliant not excused	Ratio of the number of beneficiary families who received a penalty to the number of beneficiary families who did not meet the attendance requirement.
Panel B: Electoral outcomes	
Vote share PT	Vote share of the candidate affiliated with the PT (presidential political party).
Vote share coalition	Vote share of the candidate affiliated with the party of the presidential coalition.
Vote share incumbent PT	Vote share of the incumbent affiliated with the PT (presidential political party).
Vote share incumbent no PT	Vote share of the incumbent that is not affiliated with the PT (presidential political party).
Vote share incumbent no coal	Vote share of the incumbent that is not affiliated with the party of the presidential coalition.
Turnout	Ratio of the total number of votes to the electorate.

Table A1 (continued)

Panel C: Municipal characteristics	
Population	Number of inhabitants.
% of people in urban areas	Fraction of people living in urban areas.
% of houses with access to water	Fraction of houses linked to the water system.
% of houses with access to sewerage	Fraction of houses linked to the sewerage system.
% of houses with electricity	Fraction of houses with access to electricity.
Income	per-capita income in 2000 in Brazilian *reais*.
Local radio station	Equals one if there is at least one local radio station.
% of schools with water	Fraction of public schools linked to the water system .
% of schools with electricity	Fraction of public schools with access to electricity.
% of schools with computer	Fraction of public schools with at least one computer.
% of schools with internet	Fraction of public schools with access to internet.
Enrollment	Ratio of the number of people enrolled in public schools to population.
Ratio teacher to students	Ratio of the number of teachers to the number of students.
Primary school class size	Average number of students per classroom in primary education.
Primary school dropout rates	Dropout rates of students enrolled in primary school.
Primary school completion rates	Completion rates of students enrolled in primary school.
High school class size	Average number of students per classroom in secondary education.
High school dropout rates	Dropout rates of students enrolled in high school.
High school completion rates	Completion rates of students enrolled in high school.
First	A dummy variable that equals one if the mayor is a first-term mayor (eligible to run for re-election).

Panel D: Mayoral characteristics	
Primary	Equals one if the mayor has at most primary education.
College	Equals one if the mayor has at college degree.
Male	Equals one if the mayor is male.
Married	Equals one if the mayor is married.
Future political career	Equals one if the mayor has run for higher offices after being a mayor.

Table A2: Summary Statistics - Zip Code Level

	obs	mean	sd.
Panel A: Sample of municipalities where the PT runs			
Vote share PT	1,130	0.4953	0.1560
Fraction warned	1,130	0.1191	0.2026
Fraction warned treated	1,130	0.0520	0.1133
Fraction warned with financial loss	1,130	0.0676	0.1587
Fraction warned with financial loss treated	1,130	0.0294	0.0891
Panel B: Sample of municipalities where the PT incumbent runs			
Vote share coalition	126	0.5622	0.1143
Fraction warned	126	0.1348	0.2101
Fraction warned treated	126	0.0460	0.0987
Fraction warned with financial loss	126	0.0741	0.1941
Fraction warned with financial loss treated	126	0.0171	0.0548

Notes. Panel A considers observations at the zip code level for the sample of municipalities where a candidate from the presidential party (PT) runs in the second round of October 2008 municipal elections. Panel B considers observations at the zip code level for the sample of municipalities where the PT incumbent runs in the second round of October 2008 municipal elections. See Table A1 for the definition of the variables.

Table A3: Summary Statistics: RDD Samples

	obs	mean	sd
Panel A: Mayors affiliated with the presidential coalition			
Noncompliant/BFP students	1,817	0.0332	0.0346
Excused/BFP students	1,817	0.0118	0.0221
Excused/Noncompliant	1,727	0.2839	0.2626
Noncompliant & not excused/BFP students	1,817	0.0218	0.0225
Warn/BFP families	1,817	0.0296	0.0293
Warn/Noncompliant & not excused	1,688	0.9427	0.0956
Warn loss/BFP families	1,817	0.0148	0.0176
Warn loss/Noncompliant & not excused	1,688	0.4099	0.2001
BFP families/electorate	1,817	0.1445	0.0796
Panel B: Mayors not affiliated with the presidential coalition			
Noncompliant/BFP students	2,534	0.0334	0.0366
Excused/BFP students	2,534	0.0124	0.0266
Excused/Noncompliant	2,346	0.3060	0.2889
Noncompliant & not excused/BFP students	2,534	0.0218	0.0244
Warn/BFP families	2,534	0.0277	0.0297
Warn/Noncompliant & not excused	2,261	0.9082	0.1284
Warn loss/BFP families	2,534	0.0140	0.0182
Warn loss/Noncompliant & not excused	2,261	0.3851	0.2240
BFP families/electorate	2,534	0.1337	0.0786

Notes. Panel A (Panel B) considers the RDD sample of municipalities where the mayor is (not) affiliated with the presidential coalition, taking municipal level averages during the electoral year. See Table A1 for the definition of the variables.

Table A4: Correlations – Warnings and Electoral Outcomes

	(1)	(2)	(3)	(4)
	Effects of warnings		Effects of warnings with financial loss	
	share PT	share coal	share PT	share coal
Fraction warned	-0.0663**	-0.0716**	-0.1120***	-0.0982**
	(0.0296)	(0.0317)	(0.0423)	(0.0489)
Observations	3,226	4,276	3,226	4,276

Notes. Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. The dependent variable in columns 1 and 3 (2 and 4) is the vote share of the candidate affiliated with the PT (with the presidential coalition). Columns 1 and 2 consider all types of warnings; columns 3 and 4 only consider warnings that imply financial losses (warning stages 2 to 5). All regressions are at the zip code level and include state fixed effects and the following controls: the ratio of beneficiaries to electorate and the ratio of noncompliant beneficiaries to electorate. See Table A1 in the appendix for the definition of the variables.

Appendix (For Online Publication)

This Appendix provides additional results and robustness checks, which are also discussed in the paper. In particular, we present the following:

- Balance checks for the last-digit NIS estimation strategy (Table AA1);
- The Effects of Warnings on Turnout (Table AA2);
- Validity checks for results in Table 4 and Table 7, excluding from the sample first term mayors with no previous political experience (Table AA3);
- Correlations between municipalities with high politically connected schools and other municipal characteristics (Table AA4);
- Placebo test for the results presented in Table 8 (Table AA5);
- McCrary density test of the running variable (margin of victory of the non-incumbent mayor) (Figure AA1)

Table AA1: Balance Checks – Last-digit NIS

	(1)	(2)	(3)	(4)
	fraction warned	fraction warned loss	fraction warned	fraction warned loss
BFB families/electorate	-0.0088	0.0008	-0.0112	-0.0018
	(0.0196)	(0.0131)	(0.0190)	(0.0128)
Fail/electorate	0.0160	-0.0477	0.0312	-0.0376
	(0.1302)	(0.0803)	(0.1313)	(0.0812)
BFB families	0.0000	0.0000	-0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Electorate	0.0084	0.0111**	0.0023	0.0039
	(0.0076)	(0.0052)	(0.0034)	(0.0028)
Observations	1,130	1,130	1,395	1,395
Sample - municipalities where:	PT runs	PT runs	coalition runs	coalition runs

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Columns 1 and 2 (3 and 4) consider all types of warnings (warning stages 2 to 5). All regressions include municipality fixed effects and the share of beneficiaries with warnings. See Table A1 in the appendix for the definition of the variables.

Table AA2: The Effects of Warnings on Turnout

	(1)	(2)	(3)	(4)
	Effects of warnings		Effects of warnings with financial loss	
	turnout	turnout	turnout	turnout
Fraction treated	-0.0169	-0.0105	-0.0247*	-0.0168
	(0.0108)	(0.0107)	(0.0126)	(0.0131)
Fraction warned	0.0005	-0.0005	-0.0060	-0.0051
	(0.0065)	(0.0060)	(0.0073)	(0.0072)
Observations	1,130	1,395	1,130	1,395
Municipality fixed effects	yes	yes	yes	yes
Sample - municipalities where:	PT runs	coalition runs	PT runs	coalition runs

Notes. Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Columns 1 and 2 (3 and 4) consider all types of warnings (warning stages 2 to 5). The dependent variable in all regressions is turnout. All regressions include municipality fixed effects and the following controls: the ratio of beneficiaries to electorate, the ratio of noncompliant beneficiaries to electorate and the fraction of beneficiaries with last digit of their NIS from 1 to 5. See Table A1 in the appendix for the definition of the variables.

Table AA3: Validity Checks - First Term Mayors
with Previous Political Experience

	(1)	(2)	(3)	(4)
Panel A: Warnings				
	fraction warned		fraction warned loss	
	POLY 2	LLR	POLY 2	LLR
First term mayor	-0.0083***	-0.0098***	-0.0044**	-0.0054**
	(0.0029)	(0.0037)	(0.0017)	(0.0022)
Observations	1,387	803	1,387	803
h		0.155		0.155
Panel B: Attendance				
	noncompliant/ BFP beneficiaries		noncompliant & not excused/ BFP beneficiaries	
	POLY 2	LLR	POLY 2	LLR
First term mayor	-0.0101***	-0.0113**	-0.0064***	-0.0077***
	(0.0036)	(0.0048)	(0.0022)	(0.0029)
Observations	1,387	803	1,387	803
h		0.155		0.155

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Dependent variables are: in columns 1-2 of Panel A the share of BFP beneficiary students with warnings; in columns 3-4 of Panel A the share of BFP beneficiary students with warnings that imply financial losses; in columns 1-2 of Panel B the ratio of noncompliant to BFP beneficiary students; in columns 3-5 of Panel B the ratio of noncompliant and not excused to BFP beneficiary students. Columns 1 and 3 display RDD estimates using a second order spline polynomial specification. Columns 2 and 4 display RDD estimates for local linear regressions with optimal bandwidth calculated according to Calonico et al. (2014). h denotes the interval of our running variable (for example: h=0.10 represents races where margin of victory is between -10% and 10%). All regressions include mandate and state fixed effects. All regressions consider the RDD sample of municipalities where the mayor is affiliated with the presidential coalition. First term mayors with no previous political experience are dropped from the sample. All regressions include mandate and state fixed effects. See Table A1 for the definition of the variables.

Table AA4: Correlations - School Political Connection and Municipal Characteristics

	politically connected
Population	-0.0637***
	(0.0239)
Urban	0.2767**
	(0.1108)
Water	-0.0467
	(0.0983)
Sewer	-0.1217
	(0.1010)
Electricity	-0.2497
	(0.1585)
Income	0.0214
	(0.0501)
Radio	0.0052
	(0.0342)
School with water	-0.0397
	(0.0805)
School with electricity	-0.0190
	(0.1364)
School with computer	0.0854
	(0.0760)
School with Internet	-0.0087
	(0.0694)
	[0.950]
Enrollment	0.7988
	(0.8199)
Teacher/students	0.0034
	(0.0052)
Primary class size	0.0009
	(0.0065)
Primary approval rates	0.0006
	(0.0035)
Primary dropout rates	-0.0036
	(0.0083)
High school class size	0.0055*
	(0.0031)
High school approval rate	0.0017
	(0.0025)
High school dropout rates	-0.0030
	(0.0035)
Observations	1,088

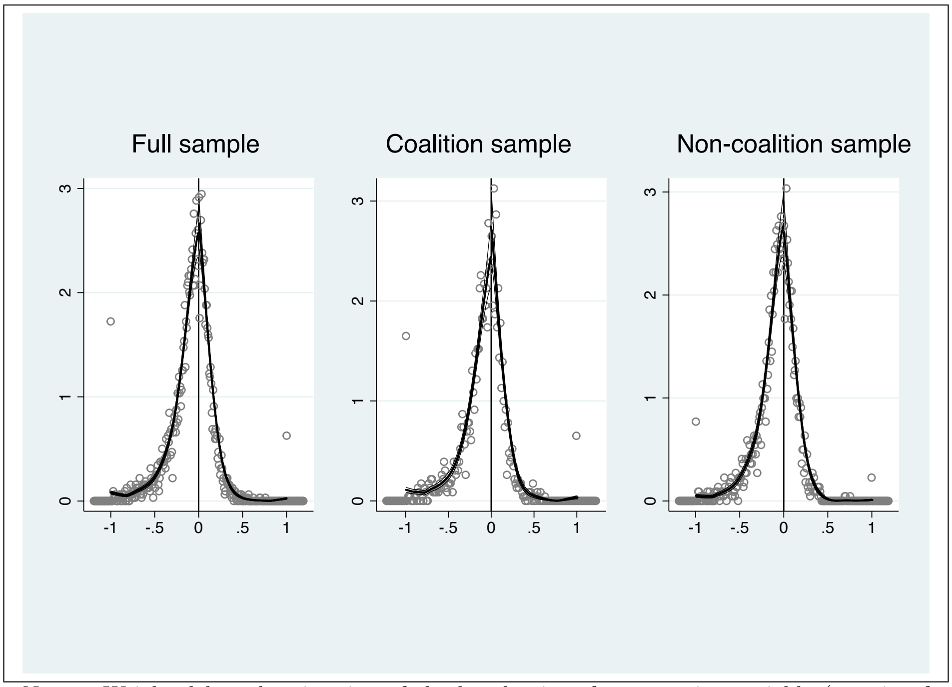
Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. The dependent variable is a dummy that equals one for municipalities where the fraction of students enrolled in schools with politically appointed principals is above the median (and zero otherwise). Robust standard errors in parenthesis. The sample of municipalities is that of Table 7, column 3 of panel A. All regressions include state and mandate dummies. See Table A1 for the definition of the variables.

Table AA5: Placebo Tests: Politically Connected, Year Before Elections

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Electoral year				Placebo: year before the election			
	noncompliant/ BFP beneficiaries		(noncompliant & not excused)/ BFP beneficiaries		noncompliant/ BFP beneficiaries		(noncompliant & not excused)/ BFP beneficiaries	
	POLY 2	LLR	POLY 2	LLR	POLY 2	LLR	POLY 2	LLR
Panel A: Municipalities with high intensity of politically connected schools								
First term mayor	-0.0107***	-0.0115**	-0.0073***	-0.0041	-0.0051	-0.0053	-0.0018	0.0004
	(0.0037)	(0.0055)	(0.0025)	(0.0035)	(0.0040)	(0.0052)	(0.0027)	(0.0035)
Observations	712	397	712	397	712	472	712	472
Panel B: Municipalities with low intensity of politically connected schools								
First term mayor	-0.0080	-0.0066	-0.0044	-0.0045	-0.0034	-0.0058	-0.0038	-0.0036
	(0.0051)	(0.0085)	(0.0032)	(0.0050)	(0.0046)	(0.0061)	(0.0029)	(0.0038)
Observations	648	302	648	302	648	373	648	373

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Dependent variable in columns 1, 2, 5 and 6 is the ratio of noncompliant to BFP beneficiary students; in columns 3, 4, 7 and 8 it is the ratio of noncompliant and not excused to BFP beneficiary students. Columns 1, 3, 5 and 7 display RDD estimates using a second order spline polynomial specification. Columns 2, 4, 6 and 8 display RDD estimates for local linear regressions with optimal bandwidth calculated according to Calonico, Cattaneo and Titiunik (2014). h denotes the interval of our running variable (for example: h=0.10 represents races where margin of victory is between -10% and 10%). Electoral year is 2012 and year before the election is 2011. Panel A reports the results for the sample of municipalities with high fraction of students enrolled in school with pollitically connected school principal (above the median). Panel B reports the results for the sample of municipalities with low fraction of students enrolled in school with pollitically connected school principal (below the median). All regressions consider only municipalities with affiliated with the presidential coalition. All regressions include mandate and state fixed effects. See Table A1 for the definition of the variables.

Figure AA1: McCrary Tests



Notes. Weighted kernel estimation of the log density of our running variable (margin of victory of the non-incumbent mayor) performed separately on either side of the zero Margin of Victory threshold. $MV_{it} > 0$ when the winner candidate in the municipality i and mandate t is the incumbent, $MV_{it} < 0$ when the winner candidate in the municipality i and mandate t is the opponent. Discontinuity estimate for the full sample is: point estimate -0.037 and standard error 0.067. Discontinuity estimate for the coalition sample is: point estimate -0.124 and standard error 0.095. Optimal bin-width and bin-size as in McCrary (2008). Discontinuity estimate for the non-coalition sample is: point estimate -0.021 and standard error 0.087. Optimal bin-width and bin-size as in McCrary (2008). Full sample considers races in 2004 and 2008 municipal elections when the incumbent runs for re-election and is one of the first two candidates. Coalition (Non-Coalition) sample consider municipalities the mayor is (not) affiliated with the presidential coalition.