The Political Economy of Program Enforcement: Evidence from Brazil*

Fernanda Brollo† Katja Kaufmann‡ Eliana La Ferrara§

This version: September 2018

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. 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 the editor, three anonymous referees, 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, University of Oslo, Barcelona GSE Summer Forum and Geneva EEA Congress for helpful comments. Giulia Zane, Simone Lenzu, Emanuele Colonnelli, Matteo Gamalerio and Rhayana Holz provided excellent research assistance. Fernanda Brollo gratefully acknowledges financial support from the Centre for Competitive Advantage in the Global Economy (CAGE) and Katja Kaufmann gratefully acknowledges financial support from the Elite program of the Baden-Wuerttemberg Foundation.

†University of Warwick, CAGE and CEPR; f.brollo@warwick.ac.uk

‡Mannheim University, BRIQ, CESifo and IZA; kaufmann@vwl.uni-mannheim.de

§Bocconi University, IGIER and LEAP; eliana.laferrara@unibocconi.it

1 Introduction

Are there electoral consequences from enforcing welfare program conditions and, if so, do politicians strategically manipulate enforcement around elections? We address these questions using original micro-level data on the universe of recipients of the Brazilian Bolsa Familia, the largest conditional cash transfer program in the world. We show that recipients who are penalized for not complying with program conditions punish mayoral candidates at the polls and that the latter react by becoming more lenient in enforcement before elections.

A large literature has studied political interference with economic policy and the emergence of political budget cycles. Most of this literature has relied on aggregate data and has found mixed results depending on whether the countries under study were high or low-middle income, on institutional details and on the empirical methodology used.[1] A more recent literature has focused on government welfare programs using micro level data and exploiting exogenous sources of variation to show that the introduction of cash transfers resulted in electoral gains for incumbents (e.g., Manacorda et al, 2010; Pop-Eleches and Pop-Eleches, 2012; De La O, 2013).

We take a different approach and focus not on the introduction or availability of social welfare programs but on their implementation, and, in particular, on the enforcement of the conditions that are often attached to these programs. This is motivated by a number of reasons. First, an increasing number of welfare programs around the world rely on some form of conditionality, such as proof that the beneficiary is actively seeking a job in order to receive unemployment benefits or sending children to school to receive anti-poverty transfers. A fundamental tenet, on which these programs rest, is that conditions are enforced according to the rules. If politicians fail to enforce conditionality due to electoral incentives, programs that are nominally conditional may de facto become unconditional. If conditionality is deemed to be important for achieving the objectives of the program, this will ultimately reduce program effectiveness. Furthermore, inconsistent enforcement of program rules could create uncertainty and perceptions of unfairness, which could undermine the functioning of the program.

A second reason for focusing on enforcement is that its manipulation may be a particu-

[1] See e.g., Brender and Drazen (2005) and Shi and Svensson (2006).

larly appealing tool from a politician's perspective. Consider the cost of introducing a new program or of expanding the coverage of an existing one as an electoral strategy: this will require increases in the budget and it will create entitlements that are difficult to revoke after the election. Lax enforcement, instead, will have much milder consequences on the budget (transfers to existing recipients are already budgeted as if compliance were complete) and it will always be possible to tighten enforcement again after elections. In this sense our focus on enforcement is close to Holland's (2015) analysis of "forbearance" , that is, the intentional and revokable non-enforcement of the law as an electoral strategy.

Finally, enforcement manipulation has clear implications in terms of visibility: it can be a way to reach a subset of the population without being too visible to other groups. For example, the fact that the government pays benefits to individuals despite their noncompliance with school attendance conditions 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.

At a general level, by focusing on the politics of enforcement, we aim to contribute to a better understanding of the political implications of the design of social protection programs, which are not yet well understood, especially for developing countries (Mares and Carnes, 2009). By providing micro evidence on how and why implementation may fail in a large nation-wide program, we also speak to the challenges of scaling up solutions that may be perfectly implemented at a smaller scale, for instance in controlled experiments.

We study the implementation of the Bolsa Familia program (BFP) in Brazil. BFP provides a monthly stipend to poor families that is conditional, among other things, on every school-aged child 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 that leads 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. For the universe of municipalities that held mayoral elections in 2008 we show a negative correlation between the fraction of beneficiary families that received warnings before the election and the vote share of the president's party

(*Partido dos Trabalhadores*, PT) and of its coalition.

To causally identify the impact of enforcement we exploit a feature of the schedule of warnings combined with the timing of the 2008 elections. In any given month, the exact date 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 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 learned about their penalties just before the elections, with zip codes where a higher fraction learned about penalties just after the elections (conditional on the overall fraction of beneficiaries punished in the zip-code in that month).

We confirm that voters respond negatively to enforcement and that they associate it with the national government. Within municipalities, the vote share of candidates from PT or its coalition is lower in zip codes where a higher fraction of noncompliant beneficiaries received warnings before, as opposed to after, the election. A one standard deviation increase in the fraction receiving warnings before the election 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 allowed to run for only one consecutive term, so we analyze whether the enforcement of BFP requirements before elections differs between first and second term mayors. We use a regression discontinuity design (RDD), analyzing municipalities where the incumbent ran for reelection 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 they cannot. 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. We also provide evidence that our results do not reflect differences between first and second term mayors in terms of political ability, experience in office, or education-related policies.

Finally, we investigate the 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 mayors from the coalition can run for reelection. 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 show that misreporting of attendance is higher in municipalities where a larger fraction of BFP students are enrolled in schools with politically appointed principals.

Our work is related to several strands of literature. Within the literature on political institutions in developing countries, recent contributions have started to analyze the determinants of weak enforcement. Levitsky and Murrillo (2009) explain how the non-enforcement of formal rules may be a result of politicians' deliberate interests or of weak state capacity, e.g., due to lack of resources or to agency problems between bureaucrats and politicians. Burgess et al. (2012) show that local officials' incentives affect the enforcement of forest policy in Indonesia, giving rise to violations of the law. Holland (2015, 2016) theorizes and provides evidence about the intentional and revocable leniency about law enforcement ("forbearance") in contexts where such enforcement would decrease support from poor voters. We contribute to this literature by uncovering strategic leniency in enforcement in the context of one type of welfare programs (conditional cash transfers) that is increasingly used in the developing world and that typically affects a significant fraction of the population. Furthermore, our RDD approach allows us to hold constant state capacity and isolate the role of electoral incentives.

A second strand of literature has shown 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). 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.

Numerous studies find that politicians strategically allocate spending to certain groups or regions (e.g., Levitt and Snyder, 1995; Solé-Ollé and Sorribas-Navarro, 2008; Brollo and Nannicini, 2012; Hodler and Raschky, 2014). With specific reference to conditional cash transfers, however, 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).[2] We contribute to this literature by showing that politicians may not distort the allocation of transfers, but might rather manipulate the enforcement of program conditions.

Finally, the evidence on the effects of term limits on the performance of elected officials is mixed. Several papers find that incumbents who are eligible for reelection perform better (e.g., Besley and Case, 1995; Ferraz and Finan, 2011), while others find that reelection incentives lead to stronger political budget cycles and more pork barrel spending (Aidt and Shvets, 2012; Labonne, 2016). In particular de Janvry, Finan and Sadoulet (2012) study Bolsa Escola, the program that preceded Bolsa Familia. They find that the program was more effective in reducing school dropout in municipalities where mayors could be reelected, as they adopted more transparent practices. Aside from using a different identification strategy, our work differs because we focus on the enforcement of conditionality (conditions were not strictly enforced during the period studied by de Janvry et al.) and on the electoral costs and incentives for manipulating enforcement.

The remainder of the paper is organized as follows. Section 2 offers a conceptual framework to understand the link between conditionality enforcement and electoral outcomes. Section 3 describes the institutional setting and data. Section 4 analyzes whether voters respond to the enforcement of BFP conditions. Section 5 examines whether politicians manipulate program enforcement before elections, and section 6 investigates the role of school principals in this manipulation. Section 7 concludes.

[2] An exception is Camacho and Conover (2011), who provide evidence on manipulation of eligibility for a social program in Colombia that is identified by a discontinuity in the density of the score used for targeting the poor.

2 Conceptual Framework

A number of theoretical arguments could be advanced to understand how voters may respond to a strict enforcement of program rules and, as a consequence, whether politicians might have incentives to manipulate enforcement.

Punishment or reward?

First of all, it is unclear whether a politician who enforces program conditions would be penalized or rewarded. Consider the reaction of beneficiaries who did not comply with BFP conditions and got penalized. From a behavioral perspective, these beneficiaries may punish politicians at the polls because they feel disgruntled, thus attributing the responsibility for negative outcomes to the politicians despite them not being responsible for (Wolfers, 2007; Healy and Malhotra, 2010).[3]

Alternatively, BFP beneficiaries (both those who got punished and those who complied with the rules) may be fully rational but take the "harsh" implementation of program conditions as a signal that the politicians do not have pro-poor preferences. This may lead these voters to update their priors on the policy positions of the candidates and move away from candidates that have revealed "less friendly" positions towards the poor (Drazen and Eslava, 2010).[4]

Also, the strict implementation of the norm may reveal something about the candidate's valence, e.g., in terms of competence or honesty (Fiorina, 1978; Rogoff, 1990; Persson and Tabellini, 1990). In this case, the capacity to implement a program according to the rules should play in favor, and not against, the politician. Related to this, 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, consistent with the idea of intrinsic reciprocity.[5] The view that voters are "fiscal conservatives" (Peltzman, 1992) who dislike, rather than appreciate, fiscal manipulation should also imply that stricter enforcement is rewarded with

[3]Since BFP was not enforced before 2006, voters might be used to *de facto* unconditional transfers and therefore believe that they are "entitled" to such transfers.

[4]The literature on distributive politics has long sustained that voters support politicians who target policies at their group (e.g., Cox and McCubbins, 1986; Dixit and Londregan, 1996). In our case, BFP beneficiaries may expect politicians who withold the benefit of noncompliance to disregard the needs of the poor in other policy choices.

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

electoral gains.

Timing of voters' reaction

For a given reaction of the voters, the timing with which one may uncover it in the data is not obvious. If beneficiaries were perfectly rational and expected punishment with certainty, they should anticipate that they will receive warnings already after noncompliance. In this case, we may 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 perfectly anticipated by noncompliant beneficiaries, the latter may respond to the actual receipt of warnings or punishments.[6] Behavioral arguments related to salience would also lead to the prediction that warnings received closer to the election should have a stronger impact.

Who is held accountable?

A nontrivial question in the context under study is who is held accountable by the voters for any penalties they may receive as BFP beneficiaries. Most of the literature on pocketbook voting assumes that incumbent politicians get the credit or blame for the targeted benefits that they provide. As we explain in detail in the next section, however, Bolsa Familia is a federal program that involves a division of responsibilities between the central government and the municipality, so that voters can, in principle, associate enforcement either with the party of the president (PT) and its coalition, or with the incumbent mayor. Existing studies document the strong association of BFP with former President Lula and its party, which led many recipients to believe that "the benefit is federal" and "Brasilia" decides (Sugiyama and Hunter, 2013, p. 54).

Politicians' incentives to manipulate

If BFP beneficiaries, who have been penalized for noncompliance, react by punishing or rewarding politicians, incentives could be created for mayors or central government officials to manipulate enforcement in the run-up to the election. However, such manipulation could be costly in a number of ways. First, it would require the politician to exert effort. Second, in case it is discovered, manipulation would have significant reputational costs. Third, manip-

[6] Consistent with this argument, Brollo, Kaufmann, and La Ferrara (2017) show that households react to the actual receipt of a warning by increasing school attendance of their children. This suggests that enforcement was not perfectly anticipated by the families, otherwise attendance would have already increased in the months after noncompliance and before the warning.

ulation could have financial costs, especially when done by municipal governments: in fact, the national government provides additional funding to schools and municipalities based on the quality of program enforcement at the local level.[7] Given these potential costs, we would expect that mayors that have relatively strong electoral incentives should be more likely to engage in manipulation. Following the literature on term limits (e.g., Besley and Case, 1995; Ferraz and Finan, 2011) 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.

How can manipulation occur?

Politicians who want to manipulate enforcement by reducing the number of warnings that are issued around the time of elections could do so in different ways. First, they could go to the source and manipulate attendance records so that fewer students fall below the threshold that is used to determine compliance. This has the advantage of making every administrative record internally consistent but requires the collaboration of schools, as they are the ones that provide attendance records to the municipal administration. For this reason, we may expect that this channel of manipulation should be easier to use when school directors and candidates are politically connected (we define political connectedness in section 6).

A second method of manipulation would be to use an 'ad hoc' tool and excuse students' absences so that they do not count towards noncompliance. While this could be done without the collaboration of the schools, an abnormally high number of students excused before elections may raise a red flag, so politicians may be wary of using this method.

Finally, manipulation may occur at the federal level if the Ministry of Social Development (MDS) decides not to issue warnings to noncompliant beneficiaries. This would require an alignment between the incentives of local politicians and those of the federal government, and also introduce an inconsistency between administrative records and program implementation that may expose the government to criticisms, if not to a scandal.

In our empirical analysis we try to shed light on which method of manipulation is used in the context under study.

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

3 Institutional Setting and Data

3.1 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 most of the period studied, from January 2008 to July 2009, all families considered poor (monthly income per capita below 58 reais *(R$)*, approximately USD 30) were eligible to receive a monthly stipend of 62 *R$*. In addition, families with monthly per capita income below 120 *R$* and with children under 16 years old attending school were eligible to receive 18 *R$* per child (20 *R$* after June 2008), for up to three children. The magnitude of the benefits is substantial: for instance, a poor family with three children attending school would receive a monthly stipend representing 40 percent of its family income.

BFP stipends are distributed directly to each family head through a "citizen card". 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. In the second step, 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 every 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 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 the schools, so 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.[8]

Sanctions: The program is enforced through a gradual system of *"warnings"*. The first time a family does not comply with the requirements it 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

[8]In schools that have computers and internet access, school principals may load daily attendance data into the system and send them directly to MEC. 60 percent of the schools in our sample have access to internet.

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 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.

3.2 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 by the 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 reelection, but since 1998 mayors are allowed to run for a second term.

3.3 Data

To analyze the enforcement of BFP we make use of a unique dataset assembled combining 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.[9] 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 those people were effectively sanctioned.

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

Electoral data come from the Tribunal Superior Eleitoral, which is the highest judicial body of the Brazilian Electoral Justice. We obtained electoral outcomes for the 2004, 2008 and 2012 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 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 that have different types of infrastructure (link to the water and sewage system, access to electricity, access to internet); we also use 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 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 aggregate the data at the zip code level, which is the smallest geographic unit to which we can assign BFP beneficiaries.[10] 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 reelection).

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 party (*Partido dos Trabalhadores*, PT) runs in the municipal election of October 2008,

[10]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 where they vote exactly. 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.

first round. On average, 32 percent of the votes in these areas go to the PT, and 7 percent of beneficiaries in the sample receive a warning for noncompliance. Panel B considers the sample of municipalities where the PT runs in the second round of the October 2008 elections. In this sample, 49.5 percent of the votes go to PT, and 12 percent of beneficiaries receive a warning for noncompliance.[11] Panel C shows statistics for the variables used in the second part of the analysis, i.e., municipal level averages for the RDD sample.[12]

4 Do Voters Respond to the Enforcement of Program Requirements?

As discussed in section 2, a priori it is not obvious whether stricter enforcement of conditionality would result in fewer or more votes for the ruling party. In this section we address this question using two strategies. The first strategy has the advantage of being representative of all municipalities, but does not allow for a rigorous causal interpretation. The second strategy applies to a subsample of municipalities that participate in second round elections, and employs a tight identification strategy. As we will show, both strategies yield very consistent results.

As a first step, we establish the correlation between enforcement and the vote share of PT in the first round of the 2008 municipal elections. We focus on PT because BFP is a federal program that was the cornerstone of President Lula's social policy and voters tend to associate the program with PT (see e.g., Sugiyama and Hunter, 2013). We estimate the following equation:

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

where Y_{zi} is the vote share of the mayoral candidate affiliated with the president's party (*share PT*) in zip code area z in municipality i; $\mathit{FractionWarned}_{zi}$ is the share of beneficiary

[11] Appendix Table A2 reports analogous statistics for the samples of municipalities where the president's coalition runs and where an incumbent mayor from PT (or coalition) runs. Note that the sample of municipalities where an incumbent mayor from PT or coalition runs for the second round in October 2008 is the same.

[12] 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.

families.[13] X_{zi} is the ratio of beneficiaries to electorate in the zip code, and θ_i denotes municipality fixed effects. Working at the zip-code level allows us to control for time-invariant municipal characteristics. We cluster standard errors at the municipality level.

[Insert Table 2]

Table 2 presents the results of this exercise. The main results are reported in columns 1 to 4. These regressions consider warnings in July 2008, which is the last period noncompliant families received warnings before the first round of elections. In columns 5 to 8 we instead consider warnings during the corresponding period one year before the elections, i.e., in 2007. We construct different enforcement variables using all program warnings (cols. 1, 2, 5, 6) or only those that imply a financial cost for beneficiaries (cols. 3, 4, 7, 8).[14] Finally, we report specifications with micro-region or municipality fixed effects, alternatively.

The results in columns 1 to 4 show a robust negative association between the fraction of beneficiaries warned in a zip code and the vote share of PT in that zip code, consistent with the idea that program beneficiaries respond negatively to the enforcement of the rules. However, we should be cautious when interpreting these estimates, as the fraction of beneficiaries that receive warnings in a given zip code might be correlated with other factors that affect electoral outcomes.

To the extent that the correlation between enforcement and PT vote share reflects time invariant differences across zip codes, we should expect to find a similar correlation at any point in time. To investigate this, in columns 5 to 8 of Table 2 we consider an analogous period in 2007, the year before the elections. We find no correlation between enforcement and PT vote share during this period. All in all, these results suggest that there might be political costs of enforcing program requirements for candidates belonging to the president's party. The results are also consistent with the idea that recent warnings are more likely to affect voters' behavior than warnings received a long time ago. However, we cannot interpret

[13]We divide the number of beneficiary families that received a warning by the number of beneficiary families in the zip code during the relevant monitoring period (e.g., for families that received a warning in July 2008, the relevant monitoring period was February-March 2008 – according to the official BFP monitoring calendar, low attendance reported during February-March 2008 resulted in warnings in July 2008).

[14]For the case of financial losses, the numerator of the variable *Fraction warned* includes only families that received warnings in stages 2 to 5.

these results in a causal way, as time varying unobservable factors across zip codes might correlate with both enforcement and electoral outcomes.

To address this problem, we employ an alternative strategy and exploit random variation in the timing when different beneficiaries learn about penalties for noncompliance. Noncompliant beneficiaries receive notifications of penalties in the second half of the month, but the exact date depends on the last digit of their Social Identification Number (NIS), which is random. In October 2008, when the second round of municipal elections were held, noncompliant beneficiaries whose NIS last digit 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.[15] We exploit this random exposure to enforcement and compare 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.[16]

We estimate the following regression:

$$Y_{zi} = \alpha + \beta \mathit{FractionTreated}_{zi} + \delta \mathit{FractionWarned}_{zi} + \varphi X_{zi} + \theta_i + \epsilon_{zi}, \tag{2}$$

where Y_{zi} is, alternatively, the vote share of the mayoral candidate of the presidential party or coalition in zip code area z within 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} include 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 families that should receive the benefit before the elections to the number of beneficiary families (*NIS 1*

[15] Given the timing of the elections and warnings, we can conduct this analysis only for the second round of the 2008 municipal elections that were held on October 26^{th}. 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.

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

to 5/beneficiaries); θ_i denotes municipality fixed effects. We report heteroskedasticity robust standard errors.[17]

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*, as opposed to *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.

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 (2). This implies that, among other things, our results are not driven by differences across zip codes in the level of compliance with program requirements (which is picked up by the variable $FractionWarned_{zi}$). Orthogonality holds thanks to the random nature of the last digit of the NIS, as we show with a number of tests in Appendix Tables AA1 and AA2.[18]

It is worth noting that the variation we exploit stems from the random allocation of NIS across smaller areas within municipalities. While the reason for staggering cash withdrawals based on NIS was to avoid bottlenecks at Caixa points, this was implemented at the municipality level (where exactly half the BFP recipients have a NIS below 5). However, different zip codes ended up having different fractions of beneficiaries with NIS ending in 1-5, which is the variation we are using. We test whether our results are driven by zip-codes with only few beneficiaries but find that this is not the case (see Appendix Table AA3 where we show that our main results –as reported in Table 3 below– are not significantly different for zip codes that are below or above the median in terms of number of beneficiaries).

[Insert Table 3]

[17] 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 attendance reported during April-May 2008 resulted in warnings in October 2008). We also estimated our regressions scaling the number of families that receive a warning 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 Table 3. Also, we winsorize all enforcement variables at the 99 percent level to handle outliers. Results are not sensitive to this procedure.

[18] In the Online Appendix Table AA1 we show that, after controlling for $FractionWarned_{zi}$, the share of beneficiaries receiving warnings before the election ($FractionTreated_{zi}$) is uncorrelated with other zip code level controls. In Table AA2 we show that the number of beneficiary families with last digit of their NIS below and above 5 is not associated with $FractionWarned_{zi}$.

Table 3 presents the results from estimating (2). The main estimates are presented in columns 1 to 4, while columns 5 to 8 present the results of a placebo test. In columns 1, 2, 5 and 6 we consider all types of warnings, and in columns 3, 4, 7 and 8 we focus on warnings that imply financial losses. 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*).[19,20]

Table 3 confirms our previous finding 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. Based on the estimates in column 2, a one standard deviation increase in the fraction of beneficiaries who receive a warning before the elections reduces the vote share of the candidate from the president's coalition by one percentage point.[21]

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

[19] 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 we can only consider municipalities where the PT candidate ran in the second round (15 municipalities). 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 (18 municipalities).

[20] We measure alignment with the federal government based on whether the mayor belongs to a political party that is part of the presidential coalition in Congress in a given year, following the literature (Brollo and Nannicini, 2012). Considering only parties whose members formally occupy federal ministries to identify the government coalition leads to a similar classification.

[21] The estimates in column 2 imply that a 10 percentage point increase in the share of BFP beneficiary families warned before the elections in a given zip code (corresponding to about 3 extra families being warned) would lead to a 0.9 percentage points reduction in the vote share of candidates affiliated with the presidential coalition in that zip code, which is equivalent to about 30 fewer votes. This implies that one extra family being warned before elections leads to about 10 fewer votes for the candidates associated with the presidential coalition, suggesting sizeable spillovers. This evidence of spillovers is consistent with research on peer effects in voter mobilization interventions in developing countries (Gine and Mansuri, 2011; Fafchamps and Vicente, 2013; Fafchamps, Vaz, and Vicente, 2018), which finds that individuals who are not directly targeted exhibit large responses to these interventions.

were issued. As expected, columns 5 to 8 show no significant effects. This increases our confidence that our results are not due to some spurious association between the last digit of the NIS and vote shares.

Finally, we 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 AA4 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.[22] The strong association between BFP and president Lula's party is consistent with the way in which the program was marketed and is corroborated by descriptive evidence (e.g., Zucco, 2008; Yoong, 2011). For example, Sugiyama and Hunter (2013) report that in their fieldwork the expression "Brasilia decides" was commonly used by beneficiaries to describe ownership of the program.[23]

In addition, we test whether enforcing BFP affects electoral turnout, which is defined as the ratio of the total number of votes to the electorate. The results, reported in Appendix Table AA5, show that in zip codes where a higher fraction of noncompliant beneficiaries received warnings before the elections turnout tends to be lower, although the effect is not statistically significant in most cases.[24]

[22]This result differs, for example, from Pop-Eleches and Pop-Eleches (2012), who find that the beneficiaries of a cash transfer program in Romania expressed greater trust in the local government (who administered the coupons) than in the central government who launched it and funded it.

[23]Ibidem, p. 54.

[24]A possible reason for the modest effects on turnout is that in Brazil, voting is compulsory for all literate citizens between 18 and 70 years old. 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, and if they do not pay they 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.

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

The results discussed so far show that program enforcement generates electoral losses for mayoral candidates associated with the ruling party. In this section we test whether this creates incentives for these mayors to reduce enforcement in the run-up to the election. In particular, we test whether mayors who have stronger electoral incentives (because they can run for a second term) are more likely to manipulate BFP enforcement.

5.1 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 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 reelection 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 reelection and a new mayor was elected) and second term mayors (municipalities where incumbents barely won reelection). Close races are a relevant context for our study because incentives for manipulation may be higher, the higher the uncertainty of the electoral outcome.

Enforcement is measured using the fraction of beneficiary families that receive any warnings or penalties (*Fraction warned*) and the fraction receiving a warning that entails losing the transfer (*Fraction warned loss*) during the election year.[25] Given the empirical design, we restrict our sample to municipalities where the incumbent mayor during the mandate 2000-2004 (2005-2008) runs for reelection in 2004 (2008) and is one of the first two candidates in

[25] 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.

that election.[26]

We consider the mayoral elections held in 2004 and 2008 to calculate the margin of victory of the non-incumbent 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 whether the mayor in power 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 2005-2008 (2009-2012) and cannot run for reelection. For $MV_{ist} > 0$ the incumbent is not reelected in 2004 (2008) and a new mayor is in power from 2005-2008 (2009-2012), and only has the possibility to run for reelection again in 2008 (2012).

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

$$E[\tau_{ist}(1) - \tau_{ist}(0)|MV_{ist} = 0] = \lim_{\epsilon \downarrow 0} E[\tau_{ist}|MV_{ist} = \epsilon] - \lim_{\epsilon \uparrow 0} E[\tau_{ist}|MV_{ist} = \epsilon]. \tag{3}$$

where τ captures the impact of being a first term mayor for municipalities around the threshold $MV_{ist} = 0$ (i.e. for elections that were decided for a margin that is tiny enough). We estimate the LATE in equation (3) by fitting a second order polynomial in MV_{ist} on either side of the threshold $MV_{ist} = 0$:

$$\tau_{ist} = \sum_{k=0}^{2}(\rho_k MV_{ist}^k) + F_{ist}\sum_{k=0}^{2}(\pi_k MV_{ist}^k) + \delta_t + \sigma_s + \varepsilon_{ist}, \tag{4}$$

where F_{ist} is a dummy variable that equals one if the mayor is a first term mayor; δ_t are mandate fixed effects and σ_s are state fixed effects. We report heteroskedasticity robust standard errors.[27] 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).

[26] About half of the races in our sample have more than two candidates. We also estimated our results considering only races with two candidates in cases where the incumbent runs for reelection, and we found very similar results. 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).

[27] Results are virtually unchanged if we include micro-region fixed effects and/or cluster standard errors at the micro-region level or if we cluster 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).

5.2 Results

[Insert Table 4]

Table 4 reports our RDD estimates using 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 (4) 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 any warning and in columns 4 to 6 it is the share who received warnings with financial penalties. Our sample covers the 2008 and 2012 municipal elections. Panel A includes all municipalities where the incumbent ran for reelection in the previous election and was one of the first two candidates; while in the remaining two panels we split the sample based on whether the mayor is affiliated (Panel B) or not affiliated (Panel C) with the presidential coalition.[28]

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 findings of the first part where we show that coalition mayors are punished by voters but not others. Quantitatively, the estimates in column 3 imply that first term mayors from the 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.[29] 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.[30]

[28]We split municipalities based on whether mayors belong to the presidential coalition, and not based on the president's party because the sample of municipalities with mayors from the PT that ran for reelection is too small.

[29]Based on the estimates in column 3, first term mayors affiliated with the presidential coalition warn about 16 fewer families before the elections, compared to second term mayors from the coalition. Our estimates in the first part of the paper (Table 3) suggest that one extra family being warned before the elections results in 10 fewer votes for mayoral candidates affiliated with the presidential coalition, suggesting the presence of spillovers. Putting these two estimates together would imply that manipulation by first term mayors from the presidential coalition would result in about 160 extra votes, which is equivalent to 0.8 percentage points of the electorate of the average municipality in our sample.

[30]We conducted additional RDD analyses comparing mayors that belong to the presidential coalition with mayors that do not belong to the coalition and restricting the sample to first-term mayors who are re-elected (and thus become second term-mayors). For this restricted sub-sample we found some evidence that second term mayors that belong to the coalition tend to enforce BFP conditions more than second term mayors that

An interesting question is whether manipulation is confined to periods around elections, or if 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 columns 2-3 and 5-6 of Table 4, but considering the year before the elections, instead of the election year.[31] 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 the ratio of BFP beneficiaries to the electorate in a given municipality, (*BFP families/electorate*), as the dependent variable. We find no effects on the share of beneficiaries, consistent with earlier research on BFP that reports little or no manipulation in the allocation of program benefits (Fried, 2012; Sugiyama and Hunter, 2013).

[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

are not affiliated with the coalition. This would be consistent with the idea that mayors that belong to the presidential coalition may be more concerned about the functioning of BFP, as it is closely identified with their party, and, in the absence of electorate incentives to manipulate enforcement, may put more effort into making the program work effectively than mayors from other parties. However, we think that it is difficult to draw conclusions from this analysis, as mayors belonging to different parties may differ along several dimensions, notably policy preferences, and it is not obvious how these may affect enforcement. This is why in the paper we do not focus on differences across parties, but rather compare mayors belonging to similar parties, who are more likely to have similar policy preferences, and focus instead on differences in electoral incentives.

[31]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).

(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).[32] The graphical and non-graphical evidence is robust to the presence of specific outliers.

It is worth discussing the relation between the evidence on manipulation we just presented and the results on the costs of enforcement in the first part of the paper. While prima facie it may seem that finding manipulation in this second part of the analysis might invalidate the identification strategy of the first part, this is not the case unless mayors are micro-managing their manipulation to only target families that receive warnings before the election, that is families with NIS last digit 1 to 5. If this were the case the timing of when beneficiaries learn about penalties for noncompliance within a given month, or more specifically whether they learn about the penalty before or after the election would not be random any more. However, apart from the fact that this type of manipulation would be informationally and logistically very cumbersome, we find no evidence of this type of micro-management of manipulation in our data. In particular, we find that the overall share of families warned before the second round elections (0.0488) is very similar to the share warned after the elections (0.0532). And we do not find a significant correlation between the fraction of beneficiaries warned in a given zip code and the fraction of beneficiaries with the last digit NIS between 1 and 5 in the same zip code, suggesting that mayors were not explicitly targeting beneficiaries who receive warnings before the second round elections (Appendix Table AA2).

5.3 Validity Checks

The standard RDD assumption is that potential outcomes must be a continuous function of the running variable at the threshold.[33] To formally test this assumption, we test the continuity of the density of the margin of victory, following McCrary (2008). The results are

[32]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.

[33]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.

reported in Appendix Figure AA1, which shows no evidence of discontinuities in the density of the margin of victory.

Do our results capture municipal characteristics?

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 test 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 tests for the sample of municipalities with mayors affiliated with the presidential coalition (which is the sample where we find evidence of manipulation).[34] 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.

Do our results capture mayor's characteristics?

One potential concern regarding the interpretation of our results is that the difference in program enforcement between first and second term mayors may capture not only the effects of electoral incentives, but other differences between first and second term mayors. For example, second term mayors may have higher political ability and, if political ability is positively correlated with program enforcement, then the differences that we find may reflect ability, not electoral incentives. Also, second term mayors by construction have more consecutive years of experience in office. If more experienced mayors are more able to enforce program conditions, then our results may reflect differences in experience.

[34]We also conducted the balance tests for the whole sample of municipalities used in the RDD analysis of 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.

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 these differences in all municipalities. However, our results show significant effects only for the sample of municipalities with mayors affiliated with the presidential coalition.

Second, we tested whether there are significant differences in observable mayoral characteristics around the cutoff. Panel A of Table 6 (columns 8 to 12) shows balance checks for mayoral characteristics such as education, gender, marital status, and whether the individual will be elected for higher office after being a mayor. For all these variables there is no discontinuity around the cutoff.

Third, to address the concern that our results may reflect differences in experience we re-estimated our regressions, restricting the sample to first term mayors with prior political experience. The results are reported in Appendix Tables A4 and A5. All our estimates are robust to considering this sample, suggesting that our findings do not reflect differences in political experience between first and second term mayors.

Do our results capture differences in policies?

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 policies that increase compliance with the school attendance requirements. In this case the lower levels of enforcement we uncover would not reflect manipulation but a changed environment. If this interpretation were correct, then we should expect to find differences not only in enforcement but also in other education-related policy outcomes around the cutoff. Panel B of Table 6 reports balance checks for numerous policy outcomes related to education: 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 no significant differences in any of the outcomes between first and second term mayors around the cutoff. This helps alleviate concerns that the lower levels of warnings that we find for first term mayors reflect differences in educational policies, rather than enforcement manipulation.

6 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 3.1, authorities could, in principle, 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.

6.1 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. We focus our analysis on municipalities where the mayor is affiliated with the coalition (i.e., the ones where manipulation is found) and compare municipalities where incumbents won by a narrow margin with those where they lost by a narrow margin. We consider five different dependent variables, which reflect the three different forms of enforcement manipulation: (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 students whose absence was excused to the number of BFP students; (iv) the ratio of students whose absence was excused to the total number of students below the threshold; and (v) the ratio of the number of families who received a warning or penalty to the number of beneficiary families who did not meet the attendance requirement and whose absence was 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 mayors can run for reelection. 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, we would expect variable (v) to be lower.

[Insert Table 7]

Table 7 contains our results. For each dependent variable we report OLS coefficients and two different RDD specifications, one using a second order spline polynomial specification (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 share of noncompliant students is 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 these municipalities the fraction of BFP students with attendance below the threshold is about 1 percentage point lower than in municipalities with second term mayors. This is about 27 percent of the sample mean.

In column 2 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) and no significant discontinuity for *Excused/BFP*, *Excused/Noncompliant* and *Warned/Noncompliant not excused* (bottom panels).

As a final check, we considered that if mayors manipulate enforcement to try to avoid electoral costs, we should expect less action 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.[35]

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

6.2 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 were 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: political appointment, public competition, election, and selective election. Political appointment is at the discretion of local politicians and is typically defended 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 have a politically appointed principal in a given municipality in our sample.

For the purpose of our analysis, the key idea is that politically appointed principals may be more likely to internalize the electoral incentives of local mayors for two reasons. One is reciprocity: having been appointed by politicians, they feel greater loyalty to them. The other is the threat that if a mayor from the same party is not elected, they may lose their job. 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: 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 perform our RDD estimates 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 caveat about this 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 schools in a municipality are highly political as a function of a vast array of observable municipal characteristics and education-related policy outcomes. The results, reported in Appendix Table AA6, show that none of the variables are statistically significant, with the exception of municipality size (with the prevalence of politically connected schools being lower in bigger municipalities). In Appendix Table AA7, we show that the results of Table 8 are not sensitive to the inclusion of municipality characteristics as controls (including municipality size).

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 and not just in the election year. To test this, we re-estimated Table 8 for a year before the elections. As we do not have attendance data before 2008, we did this for 2011, the year before the 2012 elections. The results are reported in Appendix Table AA8. In columns 1 to 4 we show that, consistent with the results in Table 8, in the electoral year 2012 first term mayors are associated with a lower share of noncompliant BFP students in those municipalities with a high fraction of 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 in the year before the election (2011), suggesting that the differential effects by school political connections reflect program manipulation due to electoral reasons.

Finally, we conducted balance tests for a vast array of observable municipal pre-treatment characteristics, mayoral characteristics and education-related policy outcomes for the sample of municipalities with mayors affiliated with the coalition of the president that have a high fraction of BFP students in politically connected schools. These results, reported in Table AA9, show that there are no discontinuities around the cutoff for this sample of municipalities.[36]

7 Conclusions

This paper studies the implementation of conditional programs and in particular whether politicians manipulate the enforcement of program conditions to influence electoral outcomes. We analyze the 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 enforcement when this imposes costs on voters, particularly when politicians face stronger electoral incentives.

To shed light on this, we first show that 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 and that they associate this enforcement with the national government. Second, we test 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 occurs mostly through misreporting of attendance, and this effect is driven by municipalities where a high fraction of BFP students attend schools

[36] We restrict the sample to those observations used in the local linear regressions with optimal bandwidth presented in Table 8, panel A, column 3.

with politically connected principals.

Our results have important policy implications. First, the manipulation of enforcement has the potential to reduce the effectiveness of welfare programs which include conditionality. To an extent, if voters internalize the incentives for manipulation, programs that are nominally conditional may de facto perform as if they were unconditional. A second 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 compared to 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 them 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. Also, limiting the discretion of politically motivated public officials (such as politically appointed school principals) or insulating the method of appointment from political considerations seems important. Whether, and to what, extent 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.

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.

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.

Brender, A. and A. Drazen (2005), "Political budget cycles in new versus established democracies," *Journal of Monetary Economics,* 52, 1271–1295.

Brollo, F., K. M. Kaufmann, and E. La Ferrara, 2017. "Learning about the Enforcement of Conditional Welfare Programs: Evidence from Brazil", IZA Working Paper no.10654.

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.

Burgess, R., M. Hansen, B. Olken, P. Potapov and S. Sieber (2012)"The Political Economy of Deforestation in the Tropics," *Quarterly Journal of Economics*, 127 (4).

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

Cattaneo M. D. , L. Keele, R. Titiunik, and G. Vazquez-Bare, 2016. "Interpreting Regression Discontinuity Designs with Multiple Cutoffs," The Journal of Politics 78, no. 4: 1229-1248.

Camacho, A. and E. Conover (2011), "Manipulation of Social Program Eligibility," *American Economic Journal: Economic Policy*, 3(2), 41-65.

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.

Cox, G. W. and M. D. McCubbins (1986), "Electoral Politics as a Redistributive Game," *Journal of Politics*, 48(2), 370–89.

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.

Dixit, A. and J. Londregan (1996), “The Determinants of Success of Special Interests in Redistributive Politics,” *Journal of Politic*s, 58(4), 1132–55.

Drazen, A. and M. Eslava (2010), “Electoral manipulation via voter-friendly spending: Theory and evidence,” *Journal of Development Economics*, 92(1), 39-52.

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.

Fafchamps, M. and Vaz, A. and Vicente, P. C., 2018. “Voting and Peer Effects: Experimental Evidence from Mozambique”. CEPR Discussion Paper No. DP12580.

Fafchamps, Marcel and Vicente, Pedro C., (2013). “Political violence and social networks: Experimental evidence from a Nigerian election”. *Journal of Development Economics*, 101, pp 27-48.

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

Gine, X.; Mansuri, G., 2011. “Together we will : experimental evidence on female voting behavior in Pakistan”. Policy Research working paper; no. WPS 5692. Washington, DC: World Bank.

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.

Holland, A. (2015), “The Distributive Politics of Enforcement,” *American Journal of Political Science*, 59(2), 357-371.

Holland, A. (2016), “Forbearance,” *American Political Science Review*, 110(2), 232-246.

Hodler, R. and P. A. Raschky (2014), “Regional Favoritism”, *The Quarterly Journal of Economics*, 129(2), 995–1033.

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.

Levitsky, S. and M. V. Murillo (2009), “Variation in Institutional Strength,” *Annual Review of Political Science*, 12 (1), 115–33.

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.

Mares, I. and M.E. Carnes (2009), “Social Policy in Developing Countries,” *Annual Review of Political Science*, 12, 93-113.

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.

Persson, T. and G. Tabellini (1990), *Macroeconomic Policy, Credibility and Politics*, Harwood Academic Publishers, New York NY.

Peltzman, S. (1992), “Voters as fiscal conservatives,” *Quarterly Journal of Economics*, 107, 327–361.

Pop-Eleches, C. and G. Pop-Eleches (2012), "Targeted Government Spending and Political Preferences," *Quarterly Journal of Political Science*, 7(30).

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

Rogoff, K. (1990), "Equilibrium political budget cycles," *American Economic Review*, 80, 21–36.

Shi, M. and J. Svensson (2006), "Political budget cycles: do they differ across countries and why?", *Journal of Public Economics,* 90, 1367–1389.

Sugiyama, N. and W. Hunter (2013), "Whither Clientelism? Good Governance and Brazil's Bolsa Familia Program", *Comparative Politics*, 46(1), 43-62.

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	Std. Dev.
Panel A: Sample of municipalities where PT runs (first round elections)			
Vote share PT	8,314	0.3227	0.2025
Fraction warned	8,314	0.0679	0.1511
Fraction warned with financial loss	8,314	0.0351	0.1094
Panel B: Sample of municipalities where PT runs (second round elections)			
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 C: RDD full sample - observations at the municipality level			
Warn/BFP families	4,351	0.0285	0.0295
Warn loss/BFP families	4,351	0.0144	0.0179
BFP families/electorate	4,351	0.1382	0.0792
Noncompliant/BFP beneficiaries	4,351	0.0333	0.0357
Excused/BFP beneficiaries	4,351	0.0122	0.0248
Excused/Noncompliant	4,073	0.2966	0.2782
Noncompliant & not excused/BFP beneficiaries	4,351	0.0218	0.0236
Warn/Noncompliant & not excused	3,949	0.9229	0.1168

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 first round of October 2008 municipal elections. Panel B 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 C considers the full RDD sample, taking municipal level averages during the electoral year. See Appendix Table 1A for variable definitions.

Table 2: Correlations – Warnings and Electoral Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Warnings in 2008				Warnings in 2007 (year prior to election)			
	Effects of warnings		Effects of warnings with financial loss		Effects of warnings		Effects of warnings with financial loss	
Dependent variable	share PT	share PT	share PT	share PT	share PT	share PT	share PT	share PT
Fraction warned	-0.0507**	-0.0289**	-0.0697**	-0.0415*	-0.0267	0.0102	0.0112	0.0161
	(0.0227)	(0.0133)	(0.0289)	(0.0241)	(0.0518)	(0.0189)	(0.0733)	(0.0317)
Observations	8,314	8,310	8,314	8,310	8,314	8,310	8,314	8,310
Micro-region fixed effects	yes	no	yes	no	yes	no	yes	no
Municipality fixed effects	no	yes	no	yes	no	yes	no	yes

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Standard errors clustered at the municipality level in parenthesis. The dependent variable is PT vote share. Columns 1 to 4 consider warnings in July 2008 (the last warning period before the elections). Columns 5 to 8 consider warnings during similar period but in the year before the elections. Columns 1, 2, 5 and 6 report results for all types of warnings. In columns 3, 4, 7 and 8 we instead focus only on warnings that imply financial losses. Columns 1, 3, 5 and 7 (2, 4, 6 and 8) include micro-region (municipality) fixed effects. All regressions consider the ratio of beneficiaries to electorate in the municipality i and zip code z as control. See Table A1 in the appendix for the definition of the variables.

Table 3: 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	
Dependent variable	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 *Partido dos Trabalhadores*, 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 4: The Effects of Electoral Incentives
on Warnings, RDD Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	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	-0.0015	-0.0006	-0.0008	-0.0008
	(0.0010)	(0.0014)	(0.0020)	(0.0006)	(0.0008)	(0.0012)
Observations	4,351	4,351	2,201	4,351	4,351	2,218
h			0.113			0.114
Panel B: Mayors affiliated with the presidential coalition						
First term mayor	-0.0026	-0.0052**	-0.0073**	-0.0012	-0.0028**	-0.0045**
	(0.0016)	(0.0022)	(0.0037)	(0.0009)	(0.0013)	(0.0021)
Observations	1,817	1,817	770	1,817	1,817	773
h			0.095			0.089
Panel C: Mayors not affiliated with the presidential coalition						
First term mayor	-0.0007	0.0008	0.0032	-0.0001	0.0007	0.0017
	(0.0014)	(0.0018)	(0.0028)	(0.0008)	(0.0011)	(0.0016)
Observations	2,534	2,534	1,205	2,534	2,534	1,268
h			0.102			0.109

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 –
Salience and Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
	Warnings in the year before elections				Robustness check	
Dependent variable	Warned/ BFP families		Warned loss/ BFP families		BFP families/ electorate	
	POLY 2	LLR	POLY 2	LLR	POLY 2	LLR
Panel A: Full sample						
First term mayor	-0.0008	-0.0016	-0.0004	-0.0013	0.0017	-0.0033
	(0.0014)	(0.0018)	(0.0010)	(0.0014)	(0.0025)	(0.0035)
Observations	4,351	2,758	4,351	2,218	4,351	2,385
h						0.126
Panel B: Mayors affiliated with the presidential coalition						
First term mayor	-0.0023	-0.0029	-0.0012	-0.0010	0.0034	0.0042
	(0.0022)	(0.0036)	(0.0014)	(0.0025)	(0.0039)	(0.0052)
Observations	1,817	770	1,817	773	1,817	1,086
h						0.147
Panel C: Mayors not affiliated with the presidential coalition						
First term mayor	0.0002	-0.0007	0.0003	-0.0002	0.0001	-0.0020
	(0.0018)	(0.0029)	(0.0013)	(0.0020)	(0.0033)	(0.0048)
Observations	2,534	1,205	2,534	1,268	2,534	1,246
h						0.107

Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. The dependent variable in columns 1 and 2 is the share of BFP families who received warnings two years before the election; in columns 5 to 8 is the share of BFP families who received warnings in the year before the election; in columns 9 and 10 it is the ratio of beneficiary families to electorate. Columns 1, 3, 5, 7 and 9 display results using a second order spline polynomial specification; columns 2, 4, 6, 8 and 10 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				
Dependent variable	population	urban	water	sewerage	electricity	income	radio	primary	college	male	married	future political career
First term mayor	-0.1379	-0.0489*	-0.0035	-0.0055	-0.0044	-0.0372	-0.0702	-0.0028	0.0230	-0.0344	0.0071	0.0023
	(0.1247)	(0.0278)	(0.0281)	(0.0217)	(0.0159)	(0.0524)	(0.0637)	(0.0482)	(0.0691)	(0.0414)	(0.0594)	(0.0321)
Observations	770	770	770	770	770	770	770	770	770	770	770	770
Panel B: Education-related policy outcomes												
Dependent variable	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	0.0008	-0.0011	0.0019	-0.0066	0.0016	-0.5225	-0.4763	-0.1654	0.3599	-0.3362	0.3599	0.9782
	(0.0326)	(0.0157)	(0.0307)	(0.0298)	(0.0028)	(0.4922)	(0.4818)	(0.3137)	(0.7697)	(0.6754)	(0.7697)	(0.8553)
Observations	770	770	770	770	770	770	770	770	770	770	770	770

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									
Dependent variable	noncompliant/ BFP beneficiaries			noncompliant & not excused/ BFP beneficiaries					
First term mayor	-0.0028	-0.0076***	-0.0111**	-0.0022*	-0.0043***	-0.0066**			
	(0.0019)	(0.0026)	(0.0045)	(0.0012)	(0.0017)	(0.0029)			
Observations	1,817	1,817	788	1,817	1,817	737			
h			0.097			0.089			
Panel B: Excuses and warnings									
Dependent variable	excused/ BFP beneficiaries			excused/ noncompliant			warned/ noncompliant & not excused		
First term mayor	-0.0006	-0.0035*	-0.0045	0.0277	0.0066	-0.0010	0.0032	0.0056	-0.0040
	(0.0015)	(0.0018)	(0.0029)	(0.0170)	(0.0230)	(0.0335)	(0.0063)	(0.0081)	(0.0117)
Observations	1,817	1,817	935	1,727	1,727	911	1,688	1,688	913
h			0.122			0.127			0.129

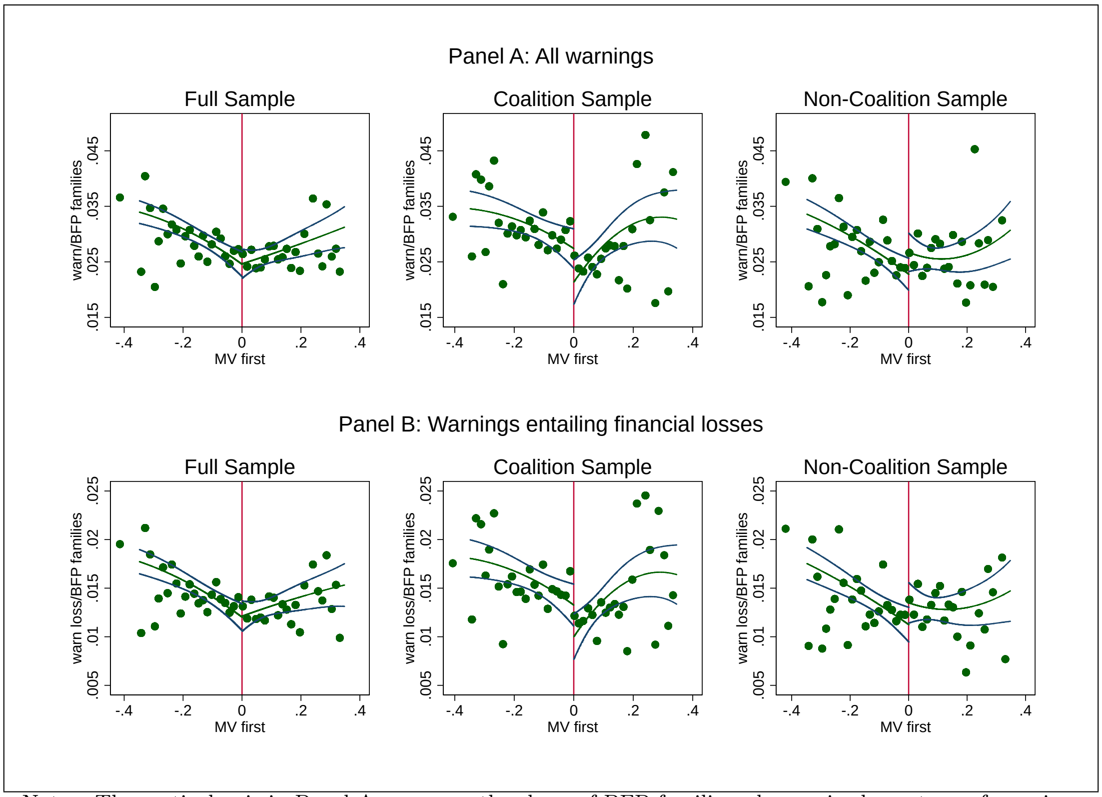
Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Dependent variables are: in columns 1-3 of Panel A the ratio of noncompliant to BFP beneficiary students; in columns 4-6 of Panel A the ratio of noncompliant and not excused to BFP beneficiary 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)
Dependent variable	Noncompliant/ BFP beneficiaries			Noncompliant not 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***	-0.0101**	-0.0026*	-0.0058***	-0.0052
	(0.0024)	(0.0032)	(0.0050)	(0.0015)	(0.0021)	(0.0037)
Observations	939	939	485	939	939	431
h			0.114			0.096
Panel B: Municipalities with low intensity of politically connected schools						
First term mayor	-0.0016	-0.0061	-0.0073	-0.0019	-0.0029	-0.0046
	(0.0032)	(0.0043)	(0.0073)	(0.0020)	(0.0027)	(0.0046)
Observations	878	878	398	878	878	375
h			0.111			0.105

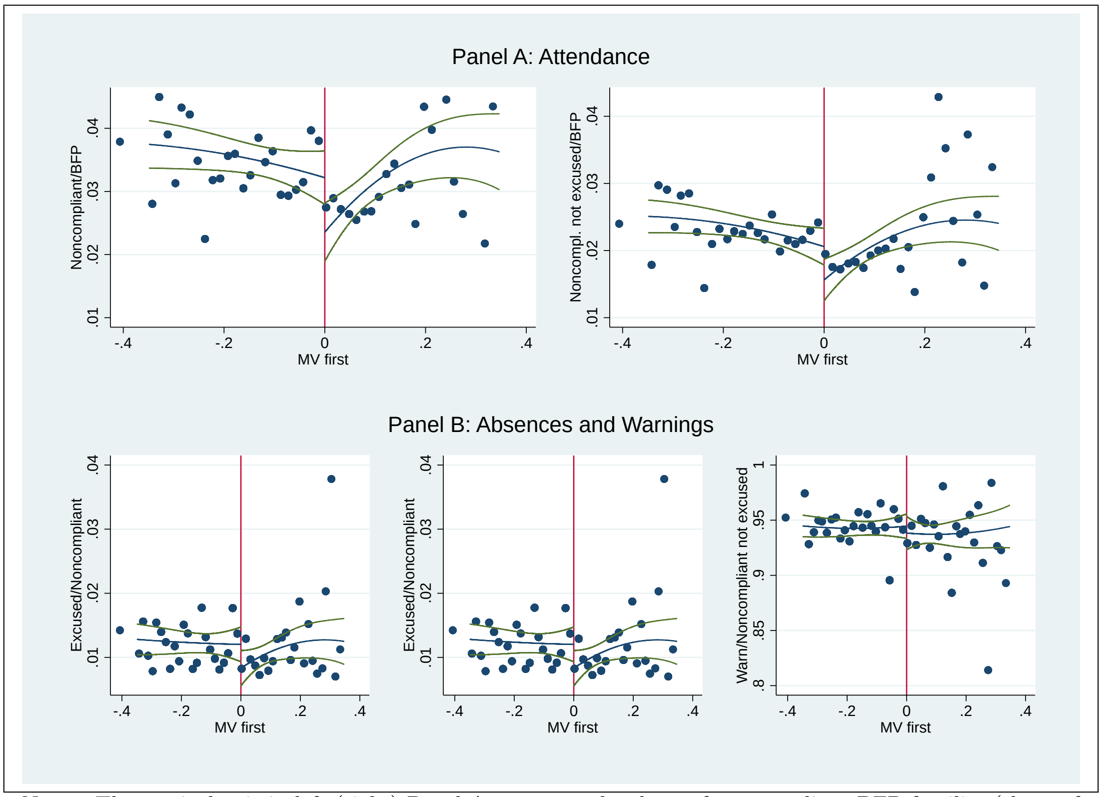
Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. The dependent variable in columns 1-3 is the ratio of noncompliant to BFP beneficiary students; in columns 4-6 it is the ratio of noncompliant and not excused to BFP beneficiary students. 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.
Warned/BFP families	Ratio of the number of families who received a warning or penalty to the number of families in the program.
Warned loss/BFP families	Ratio of the number of families who received a penalty (i.e. lost part of their transfers) to the number of families in the program.
BFP families/electorate	denotes the ratio of the number of BFP families to the electorate.
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 beneficiaries	Ratio of the number of beneficiary students below the 85 percent school attendance threshold to number of students in the program.
Excused/BFP beneficiaries	Ratio of the number of beneciary students below the 85 percent school attendance threshold, whose absence was excused to the total 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 number of students who fail to comply.
Noncompliant & not excused/BFP beneficiaries	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.
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 and was not excused.
NIS 1-5	Number of families with last digit NIS from 1 to 5
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 a 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.
Panel D: Mayoral characteristics	
First	A dummy variable that equals one if the mayor is a first-term mayor (eligible to run for re-election).
Primary	Equals one if the mayor has at most primary education.
College	Equals one if the mayor has a 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 coalition runs			
Vote share PT	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: Sample of municipalities where the PT/coalition 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			
Warn/BFP families	1,817	0.0296	0.0293
Warn loss/BFP families	1,817	0.0148	0.0176
BFP families/electorate	1,817	0.1445	0.0796
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/Noncompliant & not excused	1,688	0.9427	0.0956
Panel B: Mayors not affiliated with the presidential coalition			
Warn/BFP families	2,534	0.0277	0.0297
Warn loss/BFP families	2,534	0.0140	0.0182
BFP families/electorate	2,534	0.1337	0.078
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/Noncompliant & not excused	2,261	0.9082	0.1284

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: Validity Checks - First Term Mayors with Previous Political Experience

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Warned/BFP families			Warned loss/BFP families		
	OLS	POLY 2	LLR	OLS	POLY 2	LLR
Panel A: Full sample						
First term mayor	-0.0015	-0.0036*	-0.0028	-0.0008	-0.0023*	-0.0019
	(0.0014)	(0.0021)	(0.0026)	(0.0008)	(0.0013)	(0.0015)
Observations	3,264	3,264	1,909	3,264	3,264	1,913
h			0.151			0.152
Panel B: Mayors affiliated with the presidential coalition						
First term mayor	-0.0039*	-0.0083***	-0.0103**	-0.0019	-0.0044**	-0.0062**
	(0.0021)	(0.0029)	(0.0048)	(0.0012)	(0.0017)	(0.0027)
Observations	1,387	1,387	544	1,387	1,387	544
h			0.095			0.095
Panel C: Mayors not affiliated with the presidential coalition						
First term mayor	0.0006	0.0001	0.0024	0.0003	-0.0006	0.0007
	(0.0020)	(0.0031)	(0.0036)	(0.0012)	(0.0019)	(0.0022)
Observations	1,877	1,877	1,133	1,877	1,877	1,080
h			0.155			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%). First term mayors with no previous political experience are dropped from the sample. All regressions include mandate and state fixed effects.

Table A5: Manipulation at the Local vs. Federal Level- First Term Mayors with Previous Political Experience

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	POLY 2	LLR	OLS	POLY 2	LLR	OLS	POLY 2	LLR
Panel A: Attendance									
Dependent variable	noncompliant/ BFP beneficiaries			noncompliant & not excused/ BFP beneficiaries					
First term mayor	-0.0042*	-0.0101***	-0.0150**	-0.0032**	-0.0064***	-0.0084**			
	(0.0025)	(0.0036)	(0.0062)	(0.0016)	(0.0022)	(0.0037)			
Observations	1,387	1,387	525	1,387	1,387	544			
h			0.092			0.095			
Panel B: Excuses and warnings									
Dependent variable	excused/ BFP beneficiaries			excused/ noncompliant			warned/ noncompliant & not excused		
First term mayor	-0.0010	-0.0039*	-0.0052	0.0205	0.0047	-0.0286	0.0030	0.0081	-0.0100
	(0.0018)	(0.0023)	(0.0034)	(0.0208)	(0.0298)	(0.0491)	(0.0079)	(0.0100)	(0.0177)
Observations	1,387	1,387	621	1,326	1,326	495	1,297	1,297	467
h			0.113			0.091			0.086

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%). First term mayors with no previous political experience are dropped from the sample. All regressions include mandate and state fixed effects.

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);
- Correlations between the last-digit NIS and warnings (Table AA2);
- The effects of warnings in big vs small zip-codes (Table AA3);
- The effects of warnings on the incumbent party (Table AA4);
- The effects of warnings on turnout (Table AA5);
- Correlations between municipalities with high politically connected schools and other municipal characteristics (Table AA6);
- Attendance and political connection presented in Table 7, including municipal controls (Table AA7);
- Placebo test for the results presented in Table 8 (Table AA8);
- Balance checks for the results presented in Table 8 (Table AA9);
- 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)
	Dependent variable = fraction treated			
	All warnings	Warnings with financial loss	All warnings	Warnings with financial loss
BFP 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)
BFP 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. "The dependent variable in columns 1,3 is the fraction of individuals with NIS 1-5 among those who received any type of warning. In columns 2,4 it is the fraction of individuals with NIS 1-5 among those who received warnings that implied a financial loss. 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: Correlation Between Number of Families with Last-digit NIS 1-5 and Warnings

	(1)	(2)	(3)	(4)
Dependent variable	fraction warned		fraction warned with financial loss	
NIS 1-5 (before elections)	-0.0230	-0.0161	-0.0067	-0.0056
	(0.0310)	(0.0295)	(0.0244)	(0.0230)
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. *, ** 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 AA3: The Effects of Warnings on PT and Presidential Coalition's Vote Share - Heterogeneity, Big vs Small Zip-codes

	(1)	(2)	(3)	(4)
	Effects of warnings		Effects of warnings with financial loss	
Dependent variable	share PT	share coal	share PT	share coal
Fraction treated	-0.1313**	-0.0998**	-0.1895***	-0.1329**
	(0.0540)	(0.0490)	(0.0667)	(0.0596)
Fraction warned	0.0156	-0.0062	-0.0115	-0.0383
	(0.0320)	(0.0298)	(0.0395)	(0.0351)
Fraction treated*Big	0.0335	-0.0508	-0.0900	-0.2409
	(0.1552)	(0.1900)	(0.1800)	(0.2314)
Fraction warned*Big	-0.1407	0.0573	-0.1254	0.1035
	(0.1019)	(0.1268)	(0.1308)	(0.1608)
Big	0.1224***	0.0509**	0.1154***	0.0529**
	(0.0265)	(0.0251)	(0.0248)	(0.0235)
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 columns 1, and 3 (2 and 4) is the vote share of the candidate affliated with the PT (with the presidential coalition). BIG is a dummy variable that equals 1 for zip-codes with number of beneficiary families greater than the median. 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 AA4: The Effects of Warnings on Incumbents' Vote Share

	(1)	(2)	(3)	(4)	(5)	(6)
	Effects of warnings			Effects of warnings with financial loss		
Dependent variable	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 1-3 (4-5) consider all types of warnings (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 AA5: The Effects of Warnings on Turnout

	(1)	(2)	(3)	(4)
	Effects of warnings		Effects of warnings with financial loss	
Dependent variable	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 AA6: Correlations -
School Political Connection
and Municipal Characteristics

Variables	politically connected
Population	-0.0727**
	(0.0296)
Urban	0.2101
	(0.1350)
Water	0.0183
	(0.1155)
Sewer	-0.0799
	(0.1201)
Electricity	-0.2144
	(0.1943)
Income	0.0368
	(0.0603)
Radio	-0.0111
	(0.0408)
School with water	-0.0420
	(0.0960)
School with electricity	-0.1727
	(0.1719)
School with computer	0.0369
	(0.0890)
School with Internet	-0.0681
	(0.0799)
Enrollment	0.6203
	(1.1344)
Teacher/students	0.0055
	(0.0061)
Primary class size	0.0048
	(0.0075)
Primary approval rates	0.0030
	(0.0042)
Primary dropout rates	-0.0074
	(0.0097)
High school class size	0.0046
	(0.0037)
High school approval rate	-0.0002
	(0.0030)
High school dropout rates	-0.0037
	(0.0042)
Year dummy	0.0716
	(0.0557)
Observations	788

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 AA7: Attendance and Political Connections of School Principals - Including Municipality Characteristics as Controls

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	noncompliant/ BFP beneficiaries			noncompliant & not excused/ BFP beneficiaries		
	OLS	POLY 2	LLR	OLS	POLY 2	LLR
Panel A: Municipalities with high intensity of politically connected schools						
First term mayor	-0.0036	-0.0091***	-0.0094*	-0.0025*	-0.0059***	-0.0043
	(0.0024)	(0.0033)	(0.0051)	(0.0015)	(0.0021)	(0.0037)
Observations	939	939	485	939	939	431
h			0.114			0.096
Panel B: Municipalities with low intensity of politically connected schools						
First term mayor	-0.0016	-0.0063	-0.0079	-0.0019	-0.0031	-0.0050
	(0.0032)	(0.0042)	(0.0071)	(0.0020)	(0.0026)	(0.0046)
Observations	878	878	398	878	878	375
h			0.111			0.105

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. All regressions include as controls the fraction of people leaving in urban areas and population .

Table AA8: Politically Connected, Year Before Elections

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Electoral year				Year before the election			
Dependent variable	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	-0.0126***	-0.0131*	-0.0089***	-0.0053	-0.0059	-0.0060	-0.0024	-0.0006
	(0.0039)	(0.0070)	(0.0026)	(0.0047)	(0.0042)	(0.0064)	(0.0027)	(0.0043)
Observations	680	289	680	289	680	347	680	347
Panel B: Municipalities with low intensity of politically connected schools								
First term mayor	-0.0060	-0.0099	-0.0028	-0.0084	-0.0026	-0.0079	-0.0029	-0.0063
	(0.0049)	(0.0098)	(0.0030)	(0.0060)	(0.0044)	(0.0075)	(0.0027)	(0.0046)
Observations	680	248	680	248	680	291	680	291

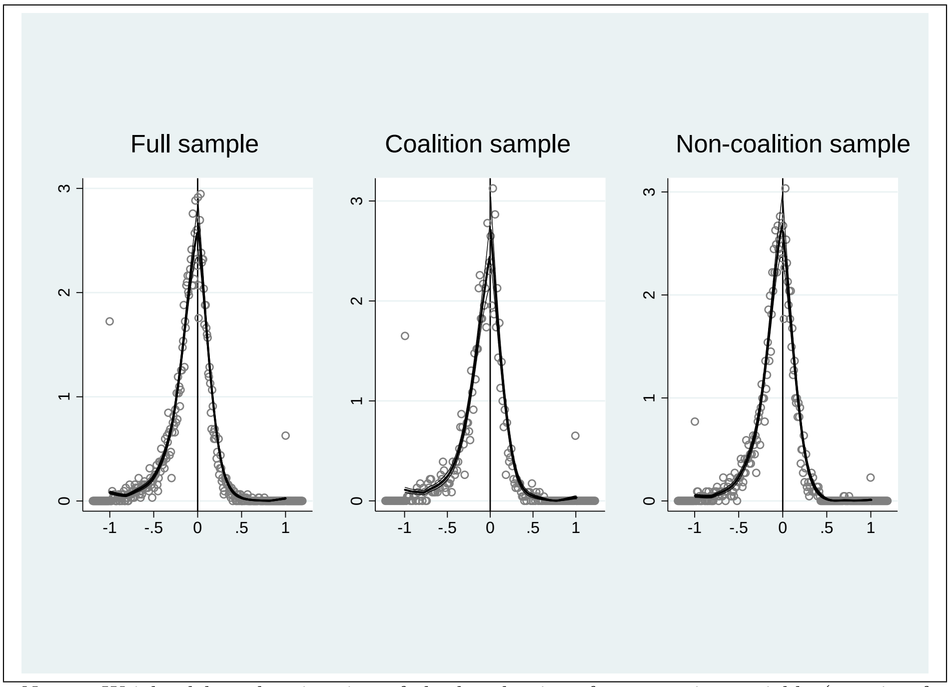
Notes. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust standard errors in parenthesis. Dependent variable in columns 1, 2, 5 and 6 is the ratio of noncompliant to BFP beneficiary students; in columns 3, 4, 7 and 8 it is the ratio of noncompliant and not excused to BFP 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.

Table AA9: Balance Checks, Political Connections of School Principals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Municipal and mayoral characteristics												
	Municipal characteristics							Mayoral characteristics				
Dependent variable	population	urban	water	sewerage	electricity	income	radio	primary	college	male	married	future political career
First term mayor	-0.0325	-0.0197	-0.0215	-0.0519	-0.0050	-0.5407	-0.5147	0.0226	0.0605	-0.2922	0.0605	-0.3241
	(0.0390)	(0.0199)	(0.0384)	(0.0395)	(0.0031)	(0.6099)	(0.6146)	(0.4307)	(1.0289)	(0.8508)	(1.0289)	(1.1075)
Observations	485	485	485	485	485	485	485	485	485	485	485	485
Panel B: Education-related policy outcomes												
Dependent variable	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	0.2350*	-0.0454	0.0101	0.0068	-0.0090	-0.0398	0.0395	-0.0790	-0.0454	-0.0769	-0.0315	0.0487
	(0.1216)	(0.0330)	(0.0322)	(0.0240)	(0.0222)	(0.0627)	(0.0792)	(0.0582)	(0.0873)	(0.0567)	(0.0773)	(0.0343)
Observations	485	485	485	485	485	485	485	485	485	485	485	485

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 8, column 3 of panel A. All regressions include mandate and state fixed effects. See Appendix Table 1A for variables definition.

Figure AA1: McCrary Tests



Notes. Weighted kernel estimation of the log density of our running variable (margin of victory of the winner candidate) performed separately on either side of the zero Margin of Victory threshold. Discontinuity estimate for the full sample is: point estimate -0.008 and standard error 0.036. Optimal bin-width and bin-size as in McCrary (2008).