June 6-7, 2019. Paper prepared for the Accelerator Workshop on Interest Groups and Education Policy at the Radcliffe Institute. Please find the latest version here.

Absence: Electoral Cycles and Teacher Absenteeism in India*

Emmerich Davies
Harvard Graduate School of Education
emmerich_davies_escobar@gse.harvard.edu

June 6, 2019

Front-line service worker absence has commonly been cited as a reason for the poor performance of developing country public services. Teachers and health care workers are often absent and, when present, not working. This absenteeism is expensive: a nationally representative sample of villages across India finds that teacher absence costs $1.5 billion a year. This paper argues that one explanation for variation in absenteeism is the differential attention politicians pay to public services over the cycle of their tenure. Using a panel of all government schools across India between 2006 and 2017, I find that teacher absenteeism decreases substantially in election years. Placebo tests on private school absenteeism finds smaller and inconsistent effects of election years on absenteeism in the private sector, lending support for a channel of political control of the bureaucracy during elections. I argue that political control of the bureaucracy has a strong effect on bureaucratic performance and that electoral accountability focuses political attention.

*I am grateful to Francisco Lagos, Sophie Litschwartz, and Fernanda Ramírez-Espinosa for fantastic research assistance on this project. Francesca Refsum-Jensenius was instrumental in acquiring electoral data and Sandip Sukhtankar provided the ultimate public good by making his electoral constituency maps publicly available. Comments from Rikhil Bhavnani, Jane Gingrich, Saad Gulzar, James Pickett, Rabia Malik, Gilles Vernier, Torsten Figueiredo Walter and conference participants at the 2017 Annual Conference on South Asia, the 2018 Political Economy of Education Workshop at Nuffield College, the 2018 North-East Universities Development Consortium annual conference, and the 2019 annual meeting of the Midwest Political Science Association have helped to greatly improve the paper. The responsibility for all errors rests solely with me.

Public sector worker absence has commonly been cited as a reason for the poor performance of low- and middle-income country public services (Chaudhury et al., 2006; Alcázar et al., 2006; Callen et al., 2017). Teachers and health care workers are often absent and, when present, not working. This absenteeism is expensive. In a nationally representative sample of villages across India, Muralidharan et al. (2017) find that teacher absenteeism costs $1.5 billion a year. In contexts of weak state capacity, low levels of accountability, and large informational asymmetries in the face of principal-agent problems, high absenteeism is not a mystery. There is, however, large variation in the levels of absenteeism between and within regions. What explains this variation?

This paper argues that one explanation for variation in absence is the differential attention that politicians give to public services over the course of their tenure. Using theory and evidence from education in India, I argue that differential attention paid to education over an electoral cycle determines a large amount of front-line service worker absence. Although accountability for public sector workers is often low, elected officials wield a number of tools to encourage them to show-up including the power of hiring, firing, and transferring under-performing public sector workers. If desired, politicians can encourage public sector workers to work.

Using a school-level dataset of all government schools in India from 2006 to 2017 matched to electoral data over the same period, I find that teachers are less likely to be absent in the year immediately preceding an election and in election years. Specifically, I find that in election years, within school absenteeism declines to zero. I argue that this is a result of political pressure applied by elected officials in election years to show up for work. While I am unable to disentangle whether this is a result of "cooking the books" – whereby elected officials pressure school administrators to falsify data on attendance – or actual decreased absence, the effect is large and consistent across a number of specifications. The results are robust to a number of different estimation strategies, including using the time to election, individual election year dummies, and modeling the full electoral cycle. I do not find a similar electoral cycle in the private sector. This suggests that increased political attention near election years is an important source of real or reported attendance. Unlike Gulzar and Pasquale (2017) I cannot conclude that this is a positive result for development and service provision: increased monitoring is likely a result of the election cycle, *not* greater accountability.

Across low- and middle-income countries, front-line service workers are frequently involved in electoral politics as poll monitors, census enumerators, and more partisan political activities as members of political parties or collective bargaining unions (Kingdon and Muzammil, 2009; Kingdon and Teal, 2010; Larreguy, Montiel Olea and Querubin, 2017). This paper adds to this literature on front-line service provider absenteeism by exploring a previously unexplored channel for front-line service provider absenteeism: the electoral cycle. Scholars have recently begun to turn to the bureaucracy as an area of study for its potentially large effects on economic and political outcomes. Pepinsky, Pierskalla and

Sacks (2017) argue that there is a distinction between literature that focuses on principal-agent problems and the developmental state literature. This papers sits in between these two literatures by trying to understand the impact of principal-agent problems on development outcomes. Additionally, the paper uses rich administrative data to answer an important political problem.

The paper closest to this one is Fagernäs and Pelkonen (2016) that uses the same data to look at teacher *transfers* over the electoral cycle. The key difference between this paper and theirs is that I am interested in top-down pressure from politicians on bureaucratic performance while theirs lies in bureaucratic sanctioning. This paper also is interested in the effects of competition on bureaucratic performance by looking at the marginal effects of political competition and electoral alignment on bureaucratic performance.

THEORETICAL EXPECTATIONS

There are two pressures teachers face to perform. First, embedded in a principal-agent relationship, there is the top-down pressure from principals (in this case higher-level bureaucrats and elected politicians) to perform. Here they are subject to sanctions, such as transfers to undesirable locations, for poor performance. Second, across the Global South, and in India specifically, teachers, either as members of political parties or other interests groups such as unions, are frequently engaged in partisan political activities during elections. It is likely that the first pressure — punishment for poor performance — increases during elections and result in reduced absenteeism, while the second pressure — partisan activities near elections — results in higher absenteeism.

Given these relationships, it is unclear whether elected politicians will exert more or less control in an election year. Theories of democratic accountability suggest that politicians would exert greater control over teachers, especially closer to an election year as they seek reelection (Lenz, 2012). On the other hand, teachers across the developing world often engage in non-teaching activities that are closely related to, if not direct, partisan political activities in support of elected politicians that would take them away from their schools in election years. In this case, politicians might reduce accountability pressures if they believe teachers can help them win elections.

Principal-Agent Problems in Education: Monitoring and Transferring Teachers

Elected politicians in India control teachers through two channels. First, politicians control postings either directly or indirectly through putting pressure on mid-level bureaucrats who can re-assign teachers to favorable or less desirable teaching positions (Béteille, 2015).[1]

[1]These are most often District Education Officers (DEOs) who sit at the district-level, the third level of Indian bureaucratic administration, and under responsibilities for Indian education policy implementation, have control of teacher postings and transfers within their districts.

Second, politicians can monitor teachers directly by visiting schools and observing whether teachers are present.[2]

In a series of unannounced audit studies, Chaudhury and Hammer (2004), Kremer et al. (2005), and Alcázar et al. (2006) explored service provider absenteeism across the world. While rates of absenteeism varied from a low of 11 percent for government school teachers in Peru (Alcázar et al., 2006), to a high of 74 percent of government doctors in clinics in Bangladesh (Chaudhury and Hammer, 2004), public service delivery across these varied contexts was characterized by high levels of absenteeism.

More recent work has attempted to move beyond this and understand why front-line service providers are absent and address high levels of absenteeism. In an intervention that provided financial incentives as well as monitoring through the use of cameras, Duflo, Hanna and Ryan (2012) found that providing financial incentives for attendance significantly reduced absenteeism while also raising learning outcomes. Callen et al. (2017) implement a smartphone monitoring system to facilitate inspections of health centers in Pakistan. They find that although the intervention increased inspections, it only decreased absenteeism in politically competitive constituencies. Taken together, these suggest significant principal-agent problems in motivating front-line workers. When provided with extrinsic incentives or increased monitoring, agents, in this case teachers and health-care workers, are more likely to show-up to work. Callen et al. (2017) also suggest a mechanism other than material incentives. Political pressures, in the form of increased electoral competition was the key driver behind decreasing absence.

Callen et al. (2017) find their results are conditional on competitive elections, suggesting that the control of the bureaucracy by local politicians is an important channel through which accountability is ensured. Gulzar and Pasquale (2017) find that in areas where politicians can fully internalize credit from successful development projects, development projects will be more successful. Better monitoring has also been found to reduce teacher absenteeism (Muralidharan et al., 2017).

Rogger and Rasul (2013) find that increased autonomy for Nigerian civil service workers led to improved project completion rates and quality, while performance incentives (or extrinsic motivations) reduced completion rates and quality. Taken together, these series of studies suggest a significant role for some combination of intrinsic and extrinsic motivations for bureaucrats to perform their job. Extrinsic incentives solve some basic performance problems by encouraging front-line workers to show-up to work, while they are less useful for encouraging better performance while at work.

Béteille (2009, 2015) finds that transfers are a powerful form of sanctioning teachers for poor performance in India, and the decision to transfer teachers lies with DEOs. Beteille's work builds on a larger series of anthropological studies by Robert Wade (1985) that looks at

[2] Interview with M. Somi Reddy, District Education Officer, Ranga Reddy District, Andhra Pradesh, September 2013.

the power of bureaucrat transfer and assignment as a powerful source of patronage. Wade (1985) outlined the specific mechanisms through which the Indian state held its employees accountable. Through the use of transfers and premiums on more desired positions, various bureaucracies in the Indian state create an "internal labor market," through which they can sanction non-performing workers as well reward supporters. Béteille (2015) tests these mechanisms and finds that a para-statal organization has emerged to facilitate the ability of the state to create an internal labor market and transfer teachers.

My own field work suggests that mid-level bureaucrats are subject to political pressure from elected officials to ensure front-line service provider attendance. The responsibility to sanction teachers, including the hiring and firing of teachers, rests at the district level, particularly with the DEO. DEOs across the state of Andhra Pradesh frequently cited political pressure, particularly from elected officials, as a key driver in their decision to monitor certain schools and sanction teachers. In a context of low capacity and high information asymmetries, DEOs relied on information and pressure from elected representatives to decide which schools to monitor.

Politicians insert themselves in these relationships of accountability by putting pressure on mid-level bureaucrats to sanction underperforming teachers or reward supporters. Theories of democratic accountability suggest that they are most likely to exert these pressures in election years to ensure voters reward them for good performance (Lenz, 2012). As a result, we should see increased monitoring and sanctioning of teachers in election years and decreased absenteeism as a result.

Vested Interests in Education: Teachers as Partisan Political Actors

While embedded in principal-agent relationships, teachers are also vested interests in the political system (Moe, 2015). Government teachers are engaged in a variety of political and administrative tasks unrelated to their official role as *teachers*. Government teachers are often the most educated members and the most constant representative of the state in rural communities (Béteille, 2009), and serve as poll booth monitors and census enumerators where they work (Béteille, 2009; Neggers, 2018). They also engage in partisan political activities such as mobilizing voters, and acting as teachers union representatives. Teachers unions are particularly powerful in India (Kingdon and Muzammil, 2009; Moe and Wiborg, 2016), and have been credited with bringing down the Chief Minister of Andhra Pradesh in the late 1990s (Rudolph and Rudolph, 2001).

In Mexico, Larreguy, Montiel Olea and Querubin (2017) find that the Mexican National Educational Workers Union (or by its Spanish acronym SNTE), the largest Mexican teachers union, serves as a partisan machine by delivering votes to the parties they support in elections. They deliver votes by monitoring voters as this effect is only present in polling stations located in schools, a monitoring function Mexican teachers share with teachers in India. During elections in India, teachers are called upon to man poll booths (Neggers, 2018), tally votes, and are often affiliated with political parties. Neggers (2018) finds poll-booth monitors privilege co-ethnics, suggesting a second mechanism of encouraging those similar to them to turn out.

This literature suggests two contradictory expectation. On the one hand, increased monitoring has reduced absenteeism across a number of contexts. Politicians will be likely to increase monitoring in election years. This suggests we should see reduced absenteeism in election years as politicians look to win subsequent elections by improving service provision near elections. On the other hand, the number of official and unofficial political activities that teachers are engaged in India and other developing countries should increase in election years. Working in their official capacity as poll booth monitors and in their unofficial capacity as part of teacher's unions, teachers are engaged in a number of partisan activities around elections that suggest they may be more absent in election years.

The two mechanisms discussed above — increased monitoring of teachers by principals in election years and teachers own partisan political activities — provide contrary expectations for teacher absenteeism in elections years. On one hand, election years increase incentives for principals to increase monitoring of the agent. Looking to win re-election, politicians will use all the tools at their disposal such as transferring non-performing teachers, to ensure lower teacher absenteeism. On the other hand, teachers are also embedded in partisan networks of their own, either as members of teachers unions, poll booth monitors, co-ethnics, or members of partisan political machines. This should increase election year absences as they are working to help politicians get re-elected. In the next section, I attempt to tease out which mechanism should prevail and provide credibly causal evidence that increased monitoring by the principal is the mechanism that dominates.

DATA & METHODS

This paper draws on two sources of data to create a school-level panel of schools across India. I combine data from the District Information System for Education (DISE) School Report Cards with assembly constituency election data.

District Information System for Education School Report Cards

The primary data source used in this paper is the DISE School Report Cards. The data consists of self-reported data on school-level infrastructure, enrollment, educational outcomes, resources, and labor for every year from 2006 to 2017.[3] School headmasters are responsible for reporting the data to the National University of Education Planning and Administra-

[3] I refer to years here as the second year in the school year. For example, the 2005-2006 academic year is referred to as 2006. This is to correspond with the electoral year each academic year would correspond to.

tion (NUEPA) at the beginning of the academic year for the previous academic year.[4] All registered schools are mandated to report this data, meaning that all government schools in the country, as well as private schools that meet government standards for registration are included in the dataset. NUEPA and DISE send the data reporting sheet to unrecognized schools they are aware of, so the data represents an undercount of unrecognized schools as the Government often has poor records of unrecognized schools (Rangaraju, Tooley and Dixon, 2012). Given that we are interested in absence in government run or aided schools, the missingness of private unrecognized schools is less of a concern.

Electoral Data

I then match the School Report Card Data with electoral data at the assembly constituency level from 2004 to 2018.[5] Assembly constituencies are India's state-level legislative assemblies and are equivalent to state houses in the United States and other two-tiered federal systems. Each assembly constituency elects one Member of the Legislative Assembly (MLA) in first-past-the-post single-member district.

I match the school report cards data to assembly constituencies using the postal pincodes provided for schools in the school report cards data. Each school observation in the school report cards data reports the postal pincode in which the school is located. I geo-locate these pincodes to spatial points, and then merge these points to assembly constituencies.[6] Postal pincodes are small units of and are nested within electoral constituencies, allowing for a clean identification of the assembly constituency in which a school is located.

Responsibilities for government schools lies at the *district* level, which is the third tier of decentralization below the Centre and States in India's administrative system. Assembly constituencies are nested within districts and there are between four and ten assembly constituencies in each district, so one DEO will respond to various MLAs. Ultimately, DEOs are appointed at the state-level by the Chief Minister (CM) of each state.

Summary statistics for all data sources are presented in Table 1. It is important to note that rates of absence are much lower than those found in independent audits. Only 13 percent of schools reported *any* absences over an academic year and about 5 days missed

[4]NUEPA is a federal public university tasked with training education administrators and researchers as well as collecting nationally representative data on education at the primary and secondary level. Headmasters are responsible for filling out forms, which are then checked by cluster and district education officials. District officials compile the DISE data for all schools in a given district and send it to the state office. Each state then collects the information and sends it NUEPA located in Delhi. There is a five percent back check to verify information.

[5]Data was downloaded from the Trivedi Centre for Political Data at Ashoka University and more details of the data collection process can be found in Jensenius (2016) and Jensenius and Verniers (2017).

[6]Sandip Sukhtankar has provided the ultimate public good by making his assembly constituency maps publicly available here. The schools in the state of Madhya Pradesh do not provide any geo-located information, so all analyses are conducted without data from the state of Madhya Pradesh and accounts for much of the unmatched data.

per school, or one per teacher. Most schools in the sample are also rural and government schools, consistent with the distribution of schools in India.

Table 1: Summary Statistics

	N	Mean	SD	Min	Max
Number of Absences per Teacher	15032804	1.14	7.68	0.00	851.00
Absent	15378545	0.12	0.32	0.00	1.00
No. of Teachers	15032804	5.49	5.61	1.00	100.00
Private School	14712991				
Government	12017497	0.82	0.39	0.00	1.00
Private	2695494	0.18	0.39	0.00	1.00
Rural School	15531546				
Urban	2148318	0.14	0.35	0.00	1.00
Rural	13383228	0.86	0.35	0.00	1.00

I run two analysis. First, I run an OLS of the form:

$$Y_{i,t} = \beta_1 \text{Election Year}_{c,t} + Z_{i,t} + \gamma_i + \zeta_t + \mu_{i,t,d}, \tag{1}$$

Where $Y_{i,t}$ is the either a binary indicator for whether any teachers were absent in school i in year t, or the logged total number of days teachers were absent in school i in year t. Election Year is the main variable of interest that takes the value of 1 if there is an election in constituency c in year t, and $Z_{i,t}$ is a vector of controls, including whether the school is urban or rural, private or government run, and the number of teachers in the school, γ_i are school-level fixed effects, and ζ_t are year fixed effects. $\mu_{i,t,d}$ is the error term clustered at the district level.

Next, I adopt the model used by Akhmedov and Zhuravskaya (2004) in their study of political business cycles in Russia, and Kapur and Vaishnav (2011) in their study of political business cycles and campaign finance in the cement industry in India. Akhmedov and Zhuravskaya (2004) construct a dataset of elections and monthly budgetary expenditures in Russia's states in order to identify the influence of political opportunism on government spending. Kapur and Vaishnav (2011) construct a dataset of elections and cement consumption in India's states in order to identify the influence of political business cycles on campaign donations from the construction industry, of which cement is a key input. Although the subjects in both papers are different the model suits this particular empirical puzzle. Specifically, I estimate the following equation using school-level yearly panel data:

$$Y_{i,t} = \sum_{j \in -2,2} \alpha_j m_{j,i,t} + \beta_1 y_{i,t-1} + Z_{i,t} + \gamma_i + \zeta_t + \mu_{i,t,d}, \tag{2}$$

where i represents schools, t represents years, and Y stands for either an indicator for any absence in the school-year, or the number of missed days (in log terms) in a given school-year. $m_{j,i,t}$ is an indicator variable that equals one when school i is j years away from the state election. The model also includes school and time fixed effects, $\gamma_i + \zeta_t$, where there is an indicator for each school and year. These fixed effects parameters control for unobserved national-level trends, as well as any unobserved school-specific characteristics.

The primary variable of interest is $m_{j,i,t}$ when j = 0, which indicates the year of the state election. I also include dummies for each of the two years preceding and following a state election. A negative coefficient on α_j would provide support for the hypothesis that the occurrence of a state election is associated with reduced teacher absenteeism.

Finally, I include a lag of the dependent variable, $y_{i,t-1}$ to explicitly model the temporal dependence of the data as absenteeism in one year is likely influenced by earlier absenteeism. I am also concerned about the presence of serial correlation in the data, so including a lag makes sense from a modeling perspective.

I run the analyses on two sets of outcomes: whether there is any absence in a school in a year, and the total number of absences in a school. We can think of these two sets of results as the extensive and intensive margins respectively.

RESULTS

I present results from two similar analyses. Table 2 presents a regression of the form in Equation 1 on assembly constituencies using a dummy variable for whether a school reports any absence in a school that year with school and year fixed effects. This can be considered as the *extensive* margin on absenteeism and represents a within-school-year variation on absenteeism. In election years, schools report between a three to four percent lower levels of absenteeism relative to non-election years in the same school.

In Table 3, I re-run the same specification in Equation 1, using the log number of days all teachers in the school were absent instead of the probability of any absence. Again, there is significantly less absenteeism in an election year, with schools reporting about half a day less absence in election years than in non-election years, from a mean of 5 days absent per year.

Next, I turn to the specification in Equation 2, using the same dependent variables in Table 2 and 3. Figure 1 reports whether there is any reported absenteeism in a school over the entire five year electoral cycle. The timing of absence becomes more apparent in this specification. Again, Figure 1 includes school and year fixed effects, so I am estimating within school variation from year to year with the election cycle. The strongest effects are in an election year and the year immediately prior to the election. By comparison, there is about an eight percent probability that a school will report any absence two years prior to an election, and about six percent probability that they will report any absence two years after an election. In an election year, the probability of any absence drops to less than one percent and is not significant. More importantly, the election year coefficient is statistically

Table 2: Any Absence in an Election Year

	Absent			
	(1)	(2)	(3)	(4)
Election Year	-0.046***	-0.031***	-0.051***	-0.031***
	(0.004)	(0.004)	(0.005)	(0.005)
Observations	6744913	6744913	6744913	6744913
Number of Schools	1063168	1063168	1063168	1063168
Year FE	No	Yes	No	Yes
School FE	No	No	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors clustered at the district level in parentheses.
Notes: The dependent variable in all specifications is a dummy variable that takes the value of one if the school reports any teacher absenteeism in that year. Each specification includes controls for the number of teachers in each school, a dummy for whether the school is in a rural area, and a lagged dependent variable. "Election Year" is a dummy variable that takes the value of 1 when the school is in a constituency in an election year.

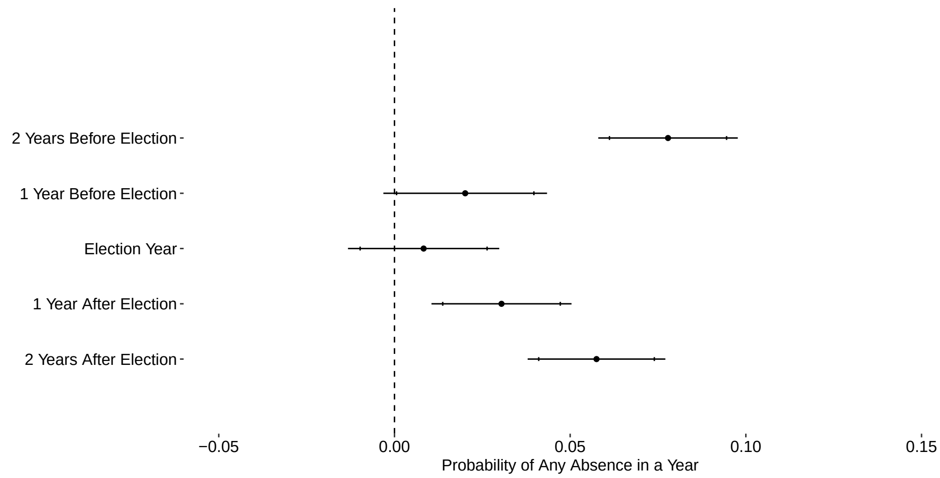
Table 3: Log Number of Days Absent in an Election Year

	Number of Absences per Teacher			
	(1)	(2)	(3)	(4)
Election Year	-0.104***	-0.071***	-0.117***	-0.073***
	(0.011)	(0.011)	(0.012)	(0.013)
Observations	6696754	6696754	6696754	6696754
Number of Schools	1060376	1060376	1060376	1060376
Year FE	No	Yes	No	Yes
School FE	No	No	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors clustered at the district level in parentheses.
Notes: The dependent variable in all specifications is the log number of days all teachers were absent from a school. Each specification includes controls for the number of teachers in each school, a dummy for rural schools, and a lagged dependent variable. "Election Year" is a dummy variable that takes the value of 1 when the school is in a constituency in an election year.

significantly different from all other coefficients in this specification other than the one year before an election (Gelman and Stern, 2006). In short, schools report nearly zero absence in an election year and the year prior to an election, and this point estimate is precisely estimated. In all other non-election years, schools report positive levels of absence ranging from between three to eight percent and these differences are different from zero and the election year estimate.

Figure 1: Any Absence in a School Year over the Electoral Cycle

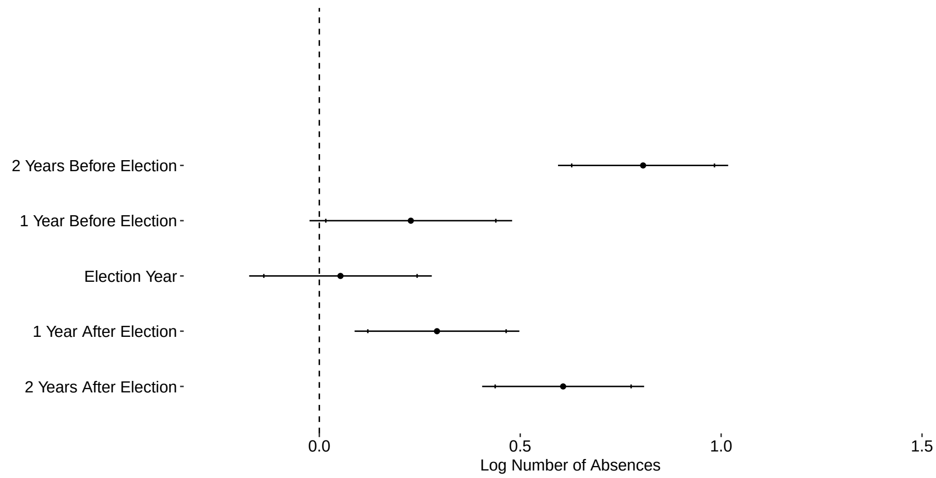


Notes: The dependent variable is a dummy variable that takes the value of one if the school reports any teacher absenteeism in that year. The regression includes controls for the number of teachers in a school, a dummy for whether the school is in a rural area, a lagged dependent variable, and year and school fixed effects. The line represents 95% confidence intervals, while the tick marks represent 90% confidence intervals. There are 8,796,007 school-year observations and 1,317,498 total schools in the sample. This figure corresponds to Column 4 in Table A1.

Figure 2 repeats the same estimates for the log number of days absent in a school year over the election cycle. The results are similar to those in Figure 1. There are higher levels of absence in non-election years, with levels of absenteeism approaching zero in election years. Again, the election year coefficient is not different from zero and statistically different from all coefficient except the year before the election.

In all specifications presented in Tables 2 and 3 and Figures 1 and 2, the coefficient on absenteeism was either lower than zero for those estimating Equation 1, or zero for those estimating Equation 2. Teacher absenteeism is lower in government schools in election years across all specifications. In the next section, I turn to teacher absenteeism in the private sector. If these results are indicative of increased political attention to the bureaucracy in election years, we should not see similar results for private schools as politicians do not exert the same level of control on private schools as they do government schools. Otherwise,

Figure 2: Log Number of Days Absent in a School Year over the Electoral Cycle



Notes: The dependent variable is the log number of days all teachers were absent from a school in that year. The regression includes controls for the number of teachers in a school, a dummy for whether the school is in a rural area, a lagged dependent variable, and year and school fixed effects. The line represents 95% confidence intervals, while the tick marks represent 90% confidence intervals. There are 8,628,707 school-year observations and 1,312,083 total schools in the sample. This figure corresponds to Column 4 in Table A2.

if results are similar, this would be suggestive of other effects specific to election years I am unable to pick-up with this data.

Absence in the Private Sector

As a placebo check on these results, I turn to absence in the private sector. In independent audits, Kremer et al. (2005) found that private school teachers are also likely to be absent, although the levels of absenteeism are much lower than government schools in the same village. With that, although absenteeism is likely to be high in private schools, too, private schools are not subject to the same monitoring that government schools are. Similar levels of absence in the private sector would raise concerns either over the quality of the data or whether we are picking up real effects in government schools. These concerns should be partially alleviated through looking at absence in the private sector. Politicians should have far less control over teacher absenteeism in the private sector as private schools do not report to elected representatives. To test this, I repeat the analysis in Equations 1, and 2 in private schools.

Table 4 reports the same results as Table 2 for private schools. While the results are similar to those in Table 2, the size of the point estimates are smaller - up to ten times smaller than those for government schools. Private schools are between half and 1/10 of a

Table 4: Any Absence in an Election Year in Private Schools

	Absent			
	(1)	(2)	(3)	(4)
Election Year	-0.006***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.002)	(0.001)	(0.002)
Observations	1153594	1153594	1153594	1153594
Number of Schools	295919	295919	295919	295919
Year FE	No	Yes	No	Yes
School FE	No	No	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors clustered at the district level in parentheses.
Notes: The dependent variable in all specifications is a dummy variable that takes the value of one if the school reports any teacher absenteeism in that year. Each specification includes controls for the number of teachers in each school, a dummy for rural schools, and a lagged dependent variable.

Table 5: Log Number of Days Absent in an Election Year in Private Schools

	Number of Absences per Teacher			
	(1)	(2)	(3)	(4)
Election Year	-0.016***	-0.009**	-0.012***	-0.006
	(0.003)	(0.004)	(0.003)	(0.004)
Observations	1244238	1244238	1244238	1244238
Number of Schools	318437	318437	318437	318437
Year FE	No	Yes	No	Yes
School FE	No	No	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors clustered at the district level in parentheses.
Notes: The dependent variable in all specifications is the log number of days any teachers were absent from a school. Each specification includes controls for the number of teachers in each school, a dummy for rural schools, and a lagged dependent variable.

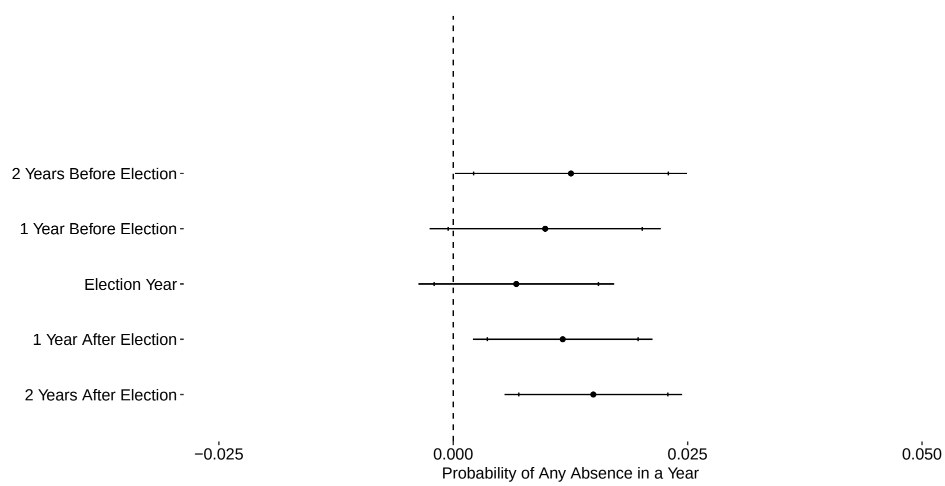
percent less likely to report absenteeism in an election year than non-election years, with this estimate precisely estimate and significant in specifications with school and year fixed effects.

Table 5 reports the same results as Table 3 for private schools. Results are similar to those in Table 3, with private schools reporting an effect size 10 times smaller than that in government schools. During elections years, absence decreased by 0.05 log days, a small but precisely estimated effect. Again, this specification includes school and year fixed effects.

Next, I turn to estimating Equation 2 for private schools, repeating the analysis from

Figures 1 and 3. This presents absenteeism over the entire electoral cycle with school and year fixed effects. While the results are similar to those in Figure 1, there are two important differences. First, the coefficient on both election years and the year before the election are indistinguishable from zero. Second, *none* of the non-election year coefficients are statistically significantly different from the election year coefficient. All coefficients are precisely estimated, suggesting that there is no difference in the levels of absence between each year over the election cycle and the probability of absence is either zero or close to zero.

Figure 3: Any Absence in a School Year over the Electoral Cycle in Private Schools

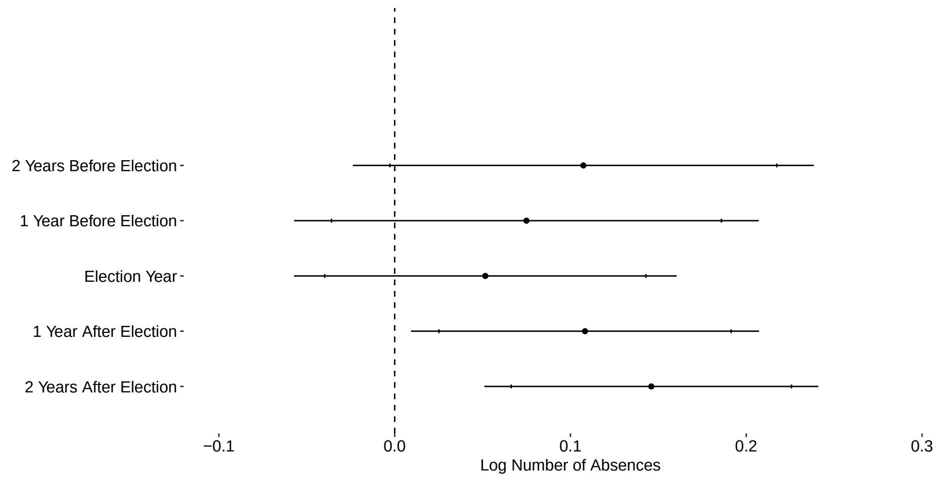


Notes: The dependent variable is a dummy variable that takes the value of one if the school reports any teacher absenteeism in that year. The regression includes controls for the number of teachers in a school, a dummy for whether the school is in a rural area, a lagged dependent variable, and year and school fixed effects. The line represents 95% confidence intervals, while the tick marks represent 90% confidence intervals. There are 1,664,926 school-year observations and 401,401 total schools in the sample. This figure corresponds to Column 4 in Table A3.

Finally, Figure 4 replicates Figure 2. In a specification that includes school and year fixed effects, the coefficients on two and one year before elections and the election year are no different from zero. Repeating the tests suggested in Gelman and Stern (2006), I find that the only coefficient that is different from the election year coefficient is the coefficient on two years after the election.

Taken together, the results from Figures 3 and Figure 4 suggest that most of the effect of absenteeism in Tables 5 and 4 is being driven by absenteeism in the two years before the election and the election year, although the effects are not consistent over the entire electoral cycle. Expanding the specifications to include and exclude school and year fixed effects in Tables A4 to A3 shows that these specifications are also sensitive to modeling

Figure 4: Log Number of Days Absent over the Electoral Cycle in Private Schools



Notes: The dependent variable is the log number of days all teachers were absent from a school in that year. The regression includes controls for the number of teachers in a school, a dummy for whether the school is in a rural area, a lagged dependent variable, and year and school fixed effects. The line represents 95% confidence intervals, while the tick marks represent 90% confidence intervals. There are 1,624,238 school-year observations and 395,020 total schools in the sample. This figure corresponds to Column 4 in Table A4.

choices. All models that do no include year fixed effects show no difference in absence over the electoral cycle or when elections are modeled as a dummy indicating the election year. Next, I turn to considering the alternative explanation that school leaders might be "cooking the books" given that the data is self-reported by school leaders.

Alternative Explanations

The major alternative alternative explanation for the observed results is that decreasing absence in the run-up to elections is not a real change in teacher absenteeism, but rather that school administrators report absence differently in election and non-election years. This would result in identical results (lower absence in elections years) through a different mechanism than I suggest here. School administrators could be "cooking the books" by systematically underreporting absence in election years. Although the material effects of this form of underreporting are vastly different, the channel through which this would happen (increased attention on schools and their management) are largely the same.

While I am unable to fully disentangle the mechanisms behind the lower absence rates, here I provide some suggestive evidence as to why absence is being driven by *actual* reduced absence as opposed to simply "cooking the books." First, it is unlikely that headmasters are referencing prior year school report cards data to complete their school report card data. From our most demanding model, that explores *within* school variation, they would have to know their exact level of absence from the previous year. If they are "cooking the books," then we should also see changes in other markers of accountability, including levels of enrollment of students. Anecdotal data suggests that completing the DISE school report cards data is a time consuming task for headmasters and they prefer to spend as little time on the activity as possible. Second, even if headmasters *are* "cooking the books," it provides support for the larger political economy story I suggest here: that of increased political pressure and attention in election years.

DISCUSSION

Combining individual school-level data from the DISE and data on election timing, I find that there is a strong and persistent electoral cycle to absenteeism in government schools. While reported rates of absenteeism are lower in this data than independent audits, approximately 13 percent of schools report some absenteeism in any given school-year, and an average of 5.5 teaching days are lost to absenteeism in each school. These numbers decline significantly in election years. The probability of any absence and the total number of days lost to absenteeism approaches zero in a government school in a constituency election year. Furthermore, these effects are not as strong, if non-existent, in private schools and are not robust to alternative modeling decisions. The results are consistent across models that only take account of election years and models that model the entire electoral cycle, as well as the inclusion of school and year fixed effects.

These results suggests that there is a strong link between bureaucratic performance, as measured by absenteeism, and democratic accountability. Teachers in government schools are more likely to show-up to work in election years. These results mirror those in Muralidharan et al. (2017) who find that there is lower absenteeism where top-down monitoring is greater. The question for policy, however, is how to extend monitoring beyond certain geographic areas or election years.

Limitations

A key question surrounding data quality is how self-reported data provided by organizations like DISE compare to independent evaluations of absence from random audits such as in Banerjee and Duflo (2006); Béteille (2009); Chaudhury and Hammer (2004); Chaudhury et al. (2006). The levels of absence found in this paper are much lower than absence found by independent evaluations of service worker absenteeism from other papers in India. Average levels of absence self-reported in the DISE dataset reach 13 percent for the *year*, far shorter than the levels of absence recorded on random spot checks in Chaudhury et al. (2006) of 25 percent on any given day.

It is important to note that this data should not be taken as a census of absence of in government schools in India. As the data is self-reported, there are strong incentives to misrepresent absence and furthermore, as DISE only asks about officially sanctioned absences, the level of unofficial absence is likely higher as we find in independent audits such as Kremer et al. (2005); Muralidharan et al. (2017). Additionally, the variable used for absence is whether there are teachers on non-teaching assignments, a specific question on whether teachers are working on official designation. While teachers are can be requisitioned for official duties, much of the absence from teachers, much that is undocumented, is not for official duties.

With this in mind, the DISE data serves as the only independent and broadly comparable source of data available to the government and broader public, and is used by the former to assess the state of schools. While the data is almost certainly biased downwards, it does have important implications for decision making as this is the dataset used by policy makers.

CONCLUSIONS

Building on previous studies of public sector worker absence in developing countries, I provide theory and evidence from a large administrative dataset on the sources of absence: the political-electoral cycle. Using an administrative dataset of over one million school over an eleven year period and an average of four elections per school, I find that teachers are far less likely to be absent in the year of a state-level election. These results are robust to a series of specifications, including year and school-level fixed effects that compare variation within schools across the entire time period. Finally, we do not see the same effect in private schools, adding support for the channel of political control of the bureaucracy.

Like other studies on the political interference of the bureaucracy (Asher and Novosad, 2017; Béteille, 2015; Gulzar and Pasquale, 2017; Kapur and Vaishnav, 2011), I find that empirical evidence of a clear channel of local level politicians interfering in service provision. Unlike Gulzar and Pasquale (2017), however, the results do not suggest the benefits of political interference, but raises the question of how to sustain political pressure in non-election years. This mirrors results in the United States where off-cycle elections result in vested interests winning power, suggesting differential attention by *voters* instead of politicians outside of high profile elections (Anzia, 2013). Teachers are present more in election years, with high levels of absenteeism in non-election years.

It is this question that the paper leaves unaddressed: how can policy makers ensure that either politicians exert the same pressure in non-election years, or teachers respond to this pressure in non-election years. The paradox is the form this pressure takes is also problematic as it is often coercive and detrimental to the provision of high quality education (Béteille, 2015; Wade, 1985). Politicians influence teacher performance through the threat of transfers, hiring, and firing, and this market is often run through middle men who do not sit in the education bureaucracy (Béteille, 2015). These networks are also embedded in larger networks of patronage that run from the local-level up to the state bureaucracy (Wade, 1985).

Bibliography

Akhmedov, Akhmed and Ekaterina Zhuravskaya. 2004. "Opportunistic Political Cycles: Test in a Young Democracy Setting." *The Quarterly Journal of Economics* 119(4):1301–1338.

Alcázar, Lorena, F. Halsey Rogers, Nazmul Chaudhury, Jeffrey Hammer, Michael Kremer and Karthik Muralidharan. 2006. "Why Are Teachers Absent? Probing Service Delivery in Peruvian Primary Schools." *International Journal of Educational Research* 45(3):117–136.

Anzia, Sarah. 2013. *Timing and Turnout: How Off-Cyle Elections Favor Organized Groups*. Chicago, IL: University of Chicago Press.

Asher, Sam and Paul Novosad. 2017. "Politics and Local Economic Growth: Evidence from India." *American Economic Journal: Applied Economics* 9(1):229–273.

Banerjee, Abhijit V. and Esther Duflo. 2006. "Addressing Absence." *Journal of Economic Perspectives* 20(1):117–132.

Béteille, Tara. 2009. Absenteeism, Transfers and Patronage: The Political Economy of Teacher Labor Markets in India PhD thesis.

Béteille, Tara. 2015. "Fixers in Indias Teacher Labor Markets: Behind the Scenes." *Asian Survey* 55(5):942–968.

Callen, Michael, Saad Gulzar, Ali Hasanain, Yasir Khan and Arman Rezaee. 2017. The Political Economy of Public Employee Absence: Experimental Evidence from Pakistan. Technical report Cambridge, MA: .

Chaudhury, Nazmul and Jeffrey Hammer. 2004. "Ghost Doctors: Absenteeism in Rural Bangladeshi Health Facilities." *The World Bank Economic Review* 18(3):423–441.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives* 20(1):91–116.

Duflo, Esther, Rema Hanna and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102(4):1241–1278.

Fagernäs, Sonja and Panu Pelkonen. 2016. "Teachers, Electoral Cycles and Learning in India.".

Gelman, Andrew and Hal Stern. 2006. "The Difference Between " Significant " and

" Not Significant " Is Not Itself Statistically Significant." *The American Statistician* 60(4):328–331.

Gulzar, Saad and Benjamin J. Pasquale. 2017. "Politicians, Bureaucrats, and Development: Evidence from India." *American Political Science Review* 111(01):162–183.

Jensenius, Francesca R. 2016. "Competing Inequalities? On the Intersection of Gender and Ethnicity in Candidate Nominations in Indian Elections." *Government and Opposition* 51(3):440–463.

Jensenius, Francesca R. and Gilles Verniers. 2017. "Studying Indian Politics with Large-Scale Data: Indian Election Data 1961–Today." *Studies in Indian Politics* 5(2):269–275.

Kapur, Devesh and Milan Vaishnav. 2011. "Quid Pro Quo: Builders, Politicians, and Election Finance in India.".

Kingdon, Geeta and Francis Teal. 2010. "Teacher Unions, Teacher Pay and Student Performance in India: A Pupil Fixed Effects Approach." *Journal of Development Economics* 91(2):278–288.

Kingdon, Geeta G. and Mohd. Muzammil. 2009. "A Political Economy of Education in India: The Case of Uttar Pradesh." *Oxford Development Studies* 37(2):123–144.

Kremer, Michael, Karthik Muralidharan, Nazmul Chaudhury, Jeffrey Hammer and F Halsey Rogers. 2005. "Teacher Absence in India: A Snapshot." *Journal of the European Economic Association* 3(2-3):658–667.

Larreguy, Horacio, Cesar E. Montiel Olea and Pablo Querubin. 2017. "Political Brokers: Partisans or Agents? Evidence from the Mexican Teachers' Union." *American Journal of Political Science* pp. n/a–n/a.

Lenz, Gabriel S. 2012. *Follow the Leader?: How Voters Respond to Politicians' Policies and Performance*. Chicago ; London: University of Chicago Press.

Moe, Terry M. 2015. "Vested Interests and Political Institutions." *Political Science Quarterly* 130(2):277–318.

Moe, Terry M. and Susanne Wiborg, eds. 2016. *The Comparative Politics of Education: Teachers Unions and Education Systems around the World*. Cambridge, United Kingdom ; New York: Cambridge University Press.

Muralidharan, Karthik, Jishnu Das, Alaka Holla and Aakash Mohpal. 2017. "The Fiscal Cost of Weak Governance: Evidence from Teacher Absence in India." *Journal of Public Economics* 145:116–135.

Neggers, Yusuf. 2018. "Enfranchising Your Own? Experimental Evidence on Bureaucrat Diversity and Election Bias in India." *American Economic Review* 108(6):1288–1321.

Pepinsky, Thomas, Jan H. Pierskalla and Audrey Sacks. 2017. "Bureaucracy and Service Delivery." *Annual Review of Political Science* 20:249–268.

Rangaraju, Baladevan, James Tooley and Pauline Dixon. 2012. The Private School Revolution in Bihar: Findings from a Survey in Patna Urban. Policy Paper India Institute New Delhi: .

Rogger, Daniel and Imran Rasul. 2013. "Management of Bureaucrats and Public Service

Delivery: Evidence from the Nigerian Civil Service.".

Rudolph, Lloyd I. and Susan H. Rudolph. 2001. "Iconisation of Chandrababu: Sharing Sovereignty in India's Federal Market Economy." *Economic and Political Weekly* 36(18):1541–1552.

Wade, Robert. 1985. "The Market for Public Office: Why the Indian State Is Not Better at Development." *World Development* 13(4):467–497.

APPENDIX

Full Results Tables

This section provides the full results tables for the figures presented in Figures 1 to 4. Those figures are represented in column 4 in each of the tables below. Columns 1 to 3 also provide robustness checks with and without year and school fixed effects. All specifications include controls for the number of teachers in the school, a dummy for rural schools, and a lagged dependent variable.

Table A1: Any Absence in a Government School over the Electoral Cycle

	Absent			
	(1)	(2)	(3)	(4)
-2 Years from Election	0.022***	0.051***	0.008	0.057***
	(0.008)	(0.008)	(0.009)	(0.008)
-1 Years from Election	-0.024***	0.010	-0.032***	0.025***
	(0.007)	(0.007)	(0.007)	(0.007)
Election Year	-0.084***	0.016**	-0.116***	0.030***
	(0.006)	(0.007)	(0.007)	(0.008)
1 Years from Election	0.004	0.027***	0.001	0.021***
	(0.007)	(0.007)	(0.007)	(0.006)
2 Years from Election	0.024***	0.069***	0.017**	0.071***
	(0.008)	(0.008)	(0.008)	(0.008)
Observations	6744913	6744913	6744913	6744913
Number of Schools	1063168	1063168	1063168	1063168
Year FE	No	Yes	No	Yes
School FE	No	No	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors clustered at the district level in parentheses.

Notes: The dependent variable in all specifications is a dummy variable that takes the value of one if the school reports any teacher absenteeism in that year. Each specification includes controls for the number of teachers in each school, a dummy for whether the school is in a rural area, and a lagged dependent variable.

Column 4 in this table corresponds to Figure 1.

Table A2: Log Number of Days Absent in a Government School over the Electoral Cycle

	Number of Absences per Teacher			
	(1)	(2)	(3)	(4)
-2 Years from Election	0.058***	0.128***	0.028	0.137***
	(0.021)	(0.020)	(0.023)	(0.020)
-1 Years from Election	-0.056***	0.036*	-0.068***	0.069***
	(0.017)	(0.019)	(0.018)	(0.018)
Election Year	-0.202***	0.051***	-0.268***	0.078***
	(0.015)	(0.017)	(0.016)	(0.020)
1 Years from Election	-0.020	0.044***	-0.021	0.037**
	(0.017)	(0.015)	(0.017)	(0.015)
2 Years from Election	0.050**	0.167***	0.036	0.170***
	(0.020)	(0.020)	(0.023)	(0.022)
Observations	6696754	6696754	6696754	6696754
Number of Schools	1060376	1060376	1060376	1060376
Year FE	No	Yes	No	Yes
School FE	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the district level in parentheses.

Notes: The dependent variable in all specifications is the log number of days all teachers were absent from a school. Each specification includes controls for the number of teachers in each school, a dummy for rural schools, and a lagged dependent variable.

Column 4 in this table correspondents to Figure 2.

Table A3: Any Absence over the Electoral Cycle in Private Schools

	Absent			
	(1)	(2)	(3)	(4)
-2 Years from Election	0.001	0.007^{**}	-0.004^{*}	0.010^{***}
	(0.003)	(0.003)	(0.002)	(0.003)
-1 Years from Election	-0.002	0.006^{**}	-0.005^{**}	0.009^{***}
	(0.002)	(0.003)	(0.002)	(0.003)
Election Year	-0.011^{***}	0.009^{***}	-0.017^{***}	0.016^{***}
	(0.003)	(0.003)	(0.003)	(0.004)
1 Years from Election	0.005^{**}	0.011^{***}	0.003	0.008^{***}
	(0.002)	(0.003)	(0.002)	(0.002)
2 Years from Election	0.005	0.016^{***}	0.002	0.016^{***}
	(0.003)	(0.004)	(0.004)	(0.004)
Observations	1258736	1258736	1258736	1258736
Number of Schools	322066	322066	322066	322066
Year FE	No	Yes	No	Yes
School FE	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the district level in parentheses.

Notes: The dependent variable in all specifications is a dummy variable that takes the value of one if the school reports any teacher absenteeism in that year. Each specification includes controls for the number of teachers in each school, a dummy for rural schools, and a lagged dependent variable.

Column 4 in this table corresponds to Figure 3.

Table A4: Log Number of Days Absence over the Electoral Cycle in Private Schools

	Number of Absences per Teacher			
	(1)	(2)	(3)	(4)
-2 Years from Election	0.003	0.015**	-0.011**	0.015**
	(0.007)	(0.007)	(0.005)	(0.007)
-1 Years from Election	-0.008	0.007	-0.014***	0.014**
	(0.005)	(0.006)	(0.005)	(0.007)
Election Year	-0.022***	0.021***	-0.036***	0.031***
	(0.006)	(0.008)	(0.006)	(0.009)
1 Years from Election	0.012**	0.023***	0.005	0.016***
	(0.005)	(0.006)	(0.005)	(0.005)
2 Years from Election	0.010	0.035***	0.001	0.032***
	(0.008)	(0.010)	(0.009)	(0.010)
Observations	1244238	1244238	1244238	1244238
Number of Schools	318437	318437	318437	318437
Year FE	No	Yes	No	Yes
School FE	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the district level in parentheses.

Notes: The dependent variable in all specifications is the log number of days any teachers were absent from a school. Each specification includes controls for the number of teachers in each school, a dummy for rural schools, and a lagged dependent variable.

Column 4 in this table correspondents to Figure 4.