September 20-22, 2018. Paper prepared for the Workshop on the Political Economy of Education at Nuffield College, University of Oxford. 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

September 7, 2018

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 panel data 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 no effects of election years on absenteeism in the private sector, lending support for a channel of political control of the bureaucracy. I argue that political control of the bureaucracy has a strong effect on service provision and bureaucratic performance, and electoral accountability focuses political attention.

*I am grateful to Francisco Lagos, Sophie Litschwartz, and Fernanda Ramirez 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, Saad Gulzar, Rabia Malik, and conference participants at the 2017 Annual Conference on South Asia have helped to greatly improve the paper. The responsibility for all errors rest solely with me.

Public sector worker absence has commonly been cited as a reason for the poor performance of developing country public services (Chaudhury et al., 2006; Alcázar et al., 2006; Callen et al., 2017). Using a variety of methods and across geographic contexts, it has been found that 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 absence costs $1.5 billion a year. In contexts of low 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 India, I argue that partisan political activities by teachers 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 coercive tools to encourage them to show-up such as the power of hiring and firing, and more importantly transfers. If desired, politicians can encourage public sector workers to work. Using a school-level dataset of all government schools in India from 2006 to 2016 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 the same election year. 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 absent, the effect is nevertheless large and consistent across a number of specifications. The results are also robust to robustness checks that explore similar effects in the private sector. 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), however, 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 the developing world, 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 potential pressures teachers can face over an electoral cycle. First, in a principal-agent framework, 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 or rewards if they do not perform. 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 — rewards for good performance and punishment for poor performance — would increase during elections and result in lower absenteeism, while the second pressure — partisan activities near elections — would result in higher absenteeism.

Principal-Agent Problems in Education: Monitoring and Transferring Teachers

Elected politicians control teachers through two channels. First, politicians control postings either directly or indirectly through putting pressure on District Education Officers (DEOs) who can re-assign teachers to favorable or less desirable teaching positions (Béteille, 2015). Second, politicians can monitor teachers directly by visiting schools and observing whether teachers are present.[1] Given this control, it is unclear whether elected politicians will exert more or less of this control in an election year. Democratic accountability would suggest that politicians would exert greater control over teachers, especially closer to an election year. On the other hand, however, teachers across the developing world often engage in non-teaching activities that are closely related, if not direct, partisan political activities.

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

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

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, 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 the power of bureaucrat transfer and assignment as a powerful source of patronage. Politicians will wield transfers as a form of punishment for non-compliant bureaucrats, evidence of which I found across Andhra Pradesh in 2013.

With regards to political pressures on public sector workers, Robert 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. Tara Béteille (2015) tests the specific mechanisms at work in Wade's (1985) claims and finds that an entire para-statal organization, in the form of *dalals* or fixers, have emerged to facilitate the ability of the state to create an internal labor market and transfer teachers.

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. While some of this work does advance a larger political agenda, teachers are also unsatisfied with this status quo and have described their role as "mere clerical workers" (Aiyar, 2018).

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 District Education Officer (DEO). Overburdened 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 where to monitor.

The literature on public sector absenteeism suggests two contradictory hypotheses. On the one hand, increased monitoring has reduced absenteeism across a number of contexts. Given the electoral cycle in India, 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.

One 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 subject to a number of pressures 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 partisan actors 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. First, I use the District Information System for Education school report cards data. Second, I rely on assembly constituency election data.

District Information System for Education School Report Cards

The primary data source used in this paper is the District Information System for Education (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 2005 to 2016. School headmasters are responsible for reporting the data to the National University of Education Planning and Administration (NUEPA) at the beginning of the academic year for the previous academic year. All registered schools are mandated to report this data and 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 urecognized 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

We also include electoral data at the assembly constituency level from 2006 to 2013.[2] Assembly constituencies are India's state-level assemblies. We 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.[3]

Through this process, I was able to match 2070820 schools year observations, or .1463958430995012 percent of the total observations in the data.[4]

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 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
Total Absences	13960360	5.54	52.88	0.00	26795.00
Absent	13960360	0.13	0.33	0.00	1.00
Log Number of Absences per Teacher	13631505	−7.81	3.64	−9.21	6.75
No. of Teachers	13631505	5.41	5.47	1.00	100.00
Private School	13372858				
Government	11000692	0.82	0.38	0.00	1.00
Private	2372166	0.18	0.38	0.00	1.00
Rural School	14113368				
Urban	1926271	0.14	0.34	0.00	1.00
Rural	12187097	0.86	0.34	0.00	1.00

We run three series of analysis. First, we 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

[2] Data was kindly provided by Francesca Refsum Jensenius and more details of the data collection process can be found in Jensenius (2016) and Jensenius and Verniers (2017).

[3] Sandip Sukhtankar has provided the ultimate public good by making his assembly constituency maps publicly available here.

[4] 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.

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, 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, γ_i + ζ_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 signifies 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 a drop in cement consumption.

Finally, I include a lag of the dependent variable, $y_{i,t-1}$ to explicitly model the temporal dependence of the data. There are strong theoretical reasons to expect that absenteeism in one year is likely influenced by earlier levels of absence. 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.

Finally, to understand the impact of political alignment and electoral competition, I estimate the following equation using constituency-level yearly panel data:

$$Y_{c,t} = \beta_1(\text{Election Year x Aligned})_{c,t} + \beta_2\text{Election Year}_{c,t} + \beta_3\text{Aligned}_{c,t} + Z_{i,t} + \gamma_i + \zeta_t + \mu_{i,t,d}, \quad (3)$$

Where c represents constituencies. Election year is a dummy that takes the value of 1 in an election year and 0 otherwise. Aligned is an indicator for whether the school is located in constituency c aligned with the party in power at the state level. Alignment is constructed based off post-election alliances as opposed to pre-election alliances as these better reflect governing coalitions. β_1 is our coefficient of interest that provides an estimate of the log number of teachers absent [percent of schools with at least one absence] in constituency c in year t.

I also estimate Equation 3 using political competition instead of political alignment to understand the effects of increasing competition on bureaucratic pressures. In the main body of the paper, I measure political competition using the margin of victory between the first and second placed candidates.[5]

For Equation 3, the dependent variables, $Y_{c,t}$ are the log number of days absent at the constituency level as well as percent of schools in the constituency that report at least one absence in that year.

RESULTS

Table 2 presents a regression of the form in Equation 1 on Vidhan Sabha (or state-level constituencies) using a dummy variable for whether a school reports any absence in a school that year and can be considered as the *extensive* margin on absenteeism. It is important to note that the coefficient on absence in the year before and the year of the election are smaller that years further from elections or after elections in all specifications. Indeed, in columns 2 and 4, the coefficient is either negative or no different from zero, suggesting lower rates of absenteeism in the year prior to an election and election years. Looking specifically at column 4, that provides the most stringent specification including year and school fixed effects for all schools in the sample, suggests that in any year, 3 and 6 percent of all schools report having a teacher absent. This number decreases to

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 less absenteeism in the year prior to an election and election years in all specifications except for column 3 that includes school but not year fixed effects. Again, in columns 1, 2

[5] In the appendix, I include a number of other measures of political competition for robustness checks including the number of effective parties and a Herfindahl index of political competition in each electoral constituency.

Table 2: Any Absence in a School Year

	Absent			
	(1)	(2)	(3)	(4)
Election Year	-0.017***	-0.033***	-0.012**	-0.032***
	(0.004)	(0.004)	(0.005)	(0.004)
Observations	8796007	8796007	8796007	8796007
Number of Schools	1317498	1317498	1317498	1317498
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.

and 4, we see rates of absenteeism that either negative or no different from zero in those two years.

Next, I turn to the specification in Equation 2, using the same dependent variables in Tables 2 and 3. Table 4 reports whether there is any reported absenteeism in a school over the electoral cycle. The timing of absence becomes more apparent in these specifications, with column 4 showing that absence decreases to zero in column 4 that includes school and year fixed effects. The strongest effects are in an election year and the year immediately prior to the election. 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.

Table 5 repeats the same exercise for the log number of days absent in a school year over the election cycle. The results are similar to those in 4. There are higher levels of absence in non-election years, with election years showing negative levels of absence in Column 2, which provides an estimate relative to the average year, and a precisely estimated zero level of absence in Column 4 that includes year and school fixed effects.

Absence in the Private Sector

As a robustness 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

Table 3: Log Number of Days Absent in a School Year

	Log Number of Absences per Teacher			
	(1)	(2)	(3)	(4)
Election Year	-0.189***	-0.376***	-0.126**	-0.367***
	(0.048)	(0.048)	(0.051)	(0.047)
Observations	8628707	8628707	8628707	8628707
Number of Schools	1312083	1312083	1312083	1312083
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.

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. Unless schools receive funding from the government, private schools are not subject to the same sanctioning mechanisms that government schools and government school teachers are.

Although not directly answering concerns over "cooking the books," if we see similar levels of absence in the private sector, we should be concerned about data quality, rather than political control over private operators. As mentioned earlier, one of the key criticism of DISE data is that it is unable to distinguish between physical absence and individuals strategically reporting lower levels of absence in election years as they know someone is paying attention.

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

Table 7 reports the same results as Table 3 for private schools. Again, results are similar to those in Table 3, with private schools reporting an effect size 10 times smaller than that

Table 4: Any Absence in a School Year over the Electoral Cycle

	Absent			
	(1)	(2)	(3)	(4)
-2 Years from Election	0.143***	0.026**	0.220***	0.078***
	(0.010)	(0.011)	(0.011)	(0.010)
-1 Years from Election	0.064***	-0.043***	0.152***	0.020*
	(0.007)	(0.009)	(0.010)	(0.012)
0 Years from Election	0.048***	-0.039***	0.118***	0.008
	(0.006)	(0.008)	(0.009)	(0.011)
1 Years from Election	0.060***	-0.010	0.119***	0.030***
	(0.006)	(0.007)	(0.010)	(0.010)
2 Years from Election	0.074***	0.017***	0.130***	0.058***
	(0.006)	(0.006)	(0.010)	(0.010)
Observations	8796007	8796007	8796007	8796007
Number of Schools	1317498	1317498	1317498	1317498
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.

in government schools. During elections years, absence decreases by between 0.004 and 0.05 log days, again, a small but precisely estimated effect. Again, specifications with year fixed effects are significant.

Next, I turn to estimating Equation 2 for private schools only, repeating the analysis from Table 4 in Table 8. Here, the most stringent specification in column 4 with year and school fixed effects, while the results mirror those in Table 4, the effect sizes are again smaller, and the point estimates in election years are not statistically significantly different from those in non-election years.

Finally, Table 9 replicates Table 5. Again, taking the strictest specification in Column 4 with year and school fixed effects, while the election year coefficient is not significant,

Table 5: Log Number of Days Absent in a School Year over the Electoral Cycle

	Log Number of Absences per Teacher			
	(1)	(2)	(3)	(4)
-2 Years from Election	1.556*** (0.107)	0.261** (0.121)	2.374*** (0.115)	0.806*** (0.108)
-1 Years from Election	0.690*** (0.078)	-0.432*** (0.101)	1.647*** (0.104)	0.228* (0.129)
0 Years from Election	0.511*** (0.065)	-0.438*** (0.082)	1.253*** (0.095)	0.053 (0.116)
1 Years from Election	0.631*** (0.062)	-0.128* (0.070)	1.254*** (0.099)	0.293*** (0.105)
2 Years from Election	0.797*** (0.067)	0.187*** (0.065)	1.390*** (0.102)	0.607*** (0.103)
Observations	8628707	8628707	8628707	8628707
Number of Schools	1312083	1312083	1312083	1312083
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.

this is no different from the coefficients on the two years before an election, as well as not statistically significantly different from the coefficient on the year after the election, either.

The Effects of Electoral Competition

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 in the run-up to elections, but rather that school administrators are reporting absence differently in election and non-election years. This would result in identical results (lower absence in elections years) but through a different mechanism than I suggest here.

Table 6: Any Absence in an Election Year in Private Schools

	Absent			
	(1)	(2)	(3)	(4)
Election Year	-0.001	-0.005***	-0.000	-0.005***
	(0.001)	(0.002)	(0.001)	(0.001)
Observations	1664926	1664926	1664926	1664926
Number of Schools	401401	401401	401401	401401
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 7: Log Number of Days Absent in an Election Year in Private Schools

	Log Number of Absences per Teacher			
	(1)	(2)	(3)	(4)
Election Year	-0.006	-0.052***	0.004	-0.053***
	(0.013)	(0.016)	(0.011)	(0.014)
Observations	1624238	1624238	1624238	1624238
Number of Schools	395020	395020	395020	395020
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.

Table 8: Any Absence over the Electoral Cycle in Private Schools

	Absent			
	(1)	(2)	(3)	(4)
-2 Years from Election	0.022*** (0.004)	0.006 (0.004)	0.032*** (0.005)	0.013** (0.006)
-1 Years from Election	0.015*** (0.002)	0.000 (0.004)	0.029*** (0.005)	0.010 (0.006)
0 Years from Election	0.012*** (0.002)	0.001 (0.004)	0.024*** (0.004)	0.007 (0.005)
1 Years from Election	0.015*** (0.002)	0.008** (0.003)	0.024*** (0.004)	0.012** (0.005)
2 Years from Election	0.014*** (0.002)	0.010*** (0.003)	0.023*** (0.005)	0.015*** (0.005)
Observations	1664926	1664926	1664926	1664926
Number of Schools	401401	401401	401401	401401
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.

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

Table 9: Log Number of Days Absence over the Electoral Cycle in Private Schools

	Log Number of Absences per Teacher			
	(1)	(2)	(3)	(4)
-2 Years from Election	0.226***	0.061	0.328***	0.107
	(0.039)	(0.046)	(0.053)	(0.067)
-1 Years from Election	0.131***	-0.023	0.296***	0.075
	(0.023)	(0.041)	(0.048)	(0.067)
0 Years from Election	0.124***	-0.000	0.241***	0.052
	(0.021)	(0.040)	(0.041)	(0.056)
1 Years from Election	0.160***	0.080**	0.244***	0.108**
	(0.024)	(0.036)	(0.041)	(0.051)
2 Years from Election	0.141***	0.104***	0.227***	0.146***
	(0.023)	(0.028)	(0.049)	(0.048)
Observations	1624238	1624238	1624238	1624238
Number of Schools	395020	395020	395020	395020
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.

provides support for the larger political economy story I suggest here: that of increased political pressure and attention in election years.

Comparison to Independent Audits

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. 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 are more likely to show-up to work and be absent for official duties 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 a ten year period and an average of two 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 off-cycle years. 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

Aiyar, Yamini. 2018. "The Politicisation of Transfers Has Undermined Teacher Accountability." *Hindustan Times* .

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.

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.".

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.

Laakso, Markku and Rein Taagepera. 1979. ""Effective" Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies* 12(3):3–27.

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.

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

Measuring Political Competition

In this section, I provide robustness checks for various ways of measuring political competition, including using the effective number of parties in a constituency as well as calculating a Herfindahl index to measure the level of competition within a constituency. I replicate the model run in Equation 3, replacing the "Aligned" variable and interaction with political competition.

Effective Number of Parties The "effective" number of parties is a measure introduced by Laakso and Taagepera (1979) to measure the relative weight of parties in an electoral constituency. The equation to measure this is:

$$N_2 = \frac{1}{\sum_{i=1}^{n} p_i^2}, \tag{4}$$

Where p_i is the fractional share of votes of the i^{th} party. This is then summed over all n parties that obtain votes. If all party vote shares are equal, N_2 is equal to the actual number of parties. If one party has close to a large majority, then N_2 will be close to one, and increase with the competitiveness of the election.

Herfindahl Index Similar to the effective number of parties, the Herfindahl index is a measure designed to measure the level of competition in a market, useful for our purposes to measure how competitive political markets are. The equation for the Herfindahl index is:

$$H = \sum_{i=1}^{N} s_i^2, \tag{5}$$

In Figure I present a scatter plot of the relation of the three measures of political competition. Panel A plots the relationship between the margin of victory and the effective number of parties, panel B plots the relationship between the margin of victory and the Herfindahl index, while panel C plots the relationship between the effective number of parties and the Herfindahl index.

ONLINE APPENDIX: NOT FOR PUBLICATION

Estimating Equations Using Logit Functions Instead of Ordinary Least Squares

In this section, I re-estimate Equations 1 and 2 using a logit function instead of ordinary least squares (OLS). OLS is preferred in the main specification for the ease of interpretation, but concerns could remain given the binary nature of the dependent variable. To test whether this affects the results, I re-estimate these two equations in Tables and .

Coding Government and Private Schools

The Indian education system has three official designation of schools: government schools that are funded and managed by the government, "government aided" or "private aided" schools that *receive* funds from the government but are privately managed,[6] and finally, fully private schools that are both financed and managed privately. As private aided school receive money from the government, they are subject to some, although not all, government regulations. For example, they must adhere to many of the provisions outlined in the 2009 Right to Education Act, including provisions on teacher pay and minimum facilities of the school. Fully private schools are exempt from these regulations.

In this section, I test whether the results are sensitive to how government schools and private schools are coded. I re-run the results presented in Tables 2 to Table 9 in two ways. First, I code private aided schools as government schools, and then I leave them out altogether. The results presented in the main body of the paper codes them as private schools.

[6] Similar to charter schools in the United States. From here on, I will refer to them as "private aided" as they are more commonly referred to in India.

Dealing with Zero Absence Values

Given the self-reported nature of the data, many schools report zero absence in multiple years. I also use a logged value of absence as there are extreme right tails to the data with some schools reporting very high levels of absence. To deal with zero logs, I assign schools that report zero absence a value of 10 percent of the minimum average level of absence with a non-zero value.