The unintended consequences of accountability: Quasi-experimental evidence from policing in Pakistan

Hamna Ahmed[*] Dareen Latif[†] Kate Vyborny[‡][§]

First version: November 7, 2021

This version: December 30, 2022

Abstract

We investigate data manipulation in the principal-agent relationship using administrative microdata from the police in Pakistan. We find that police are *less* likely to register citizens' crime reports in areas where political accountability pressure is higher. We report a series of tests to rule out the possibility that this result is explained by increased crime prevention. The results are consistent with officials deciding not to register some reports, under pressure to keep crime statistics low. As a new crime reports tracking database is rolled out, improving observability, the number of registered crimes increases.

JEL codes: D72, D73, H40

[*]Lahore School of Economics

[†]Lahore School of Economics

[‡]Duke University

[§]We gratefully acknowledge Punjab Police for providing access to the main dataset used in this paper. We appreciate helpful comments from Bocar Ba, Aditi Bhowmick, Jennifer Doleac, Nirvikar Jassal, Erzo Luttmer, and workshop participants at Duke University, the Workshop on the Political Economy of South Asia, NOVAFRICA-Universidade Nova de Lisboa, the Nordic Conference in Development Economics 2021, Lahore School ADE 2021, the Indian Statistical Institute 2021, Oxford-CSAE 2022, and Royal Economic Society 2022. This study was approved under Duke University IRB protocol 2021-0439.

1 Introduction

Principal-agent relationships are a central feature of complex organizations such as firms and bureaucracies. If the principal imperfectly observes the agent's performance, accountability pressure could backfire. This can lead to reallocation of effort from outcomes unobserved by the principal to those observed (multitasking) (Holmstrom, 1979; Holmstrom and Milgrom, 1991). Another possibility is that the agent more directly *manipulates* the signal to the principal, for example by falsifying data on performance.

We investigate this issue in the context of a core state function: policing. Quantitative and qualitative reports from various contexts suggest that senior officials exert pressure on police to reduce crime rates, and that this may backfire by incentivizing police to ignore citizens' crime reports instead of registering them, or downgrade them to lower level infractions (Eterno and Silverman, 2012; Eterno *et al.*, 2016; Yu and Zhang, 1999; Eckhouse, 2022; Banerjee *et al.*, 2021). We use large-scale, detailed administrative microdata from the police force in Pakistan to document such data manipulation, in the form of police officers choosing not to register or to downgrade crimes reported by citizens. We examine the effect of (1) an increase in the ability of the principal (politician) to impose consequences on the agent (police officer), (*accountability pressure*); and (2) an improvement in the principal's ability to observe the agent's activity (*observability*).

To isolate (1), we examine changes in police behavior before and after an election in which the party in power changed. We posit that senior officials in areas with a higher level of *political alignment* with the government (defined as the vote share for the governing party) have a greater ability to exert *accountability pressure* on police, i.e. they can impose stronger incentives (such as promotions) on officials. This may occur for two reasons: first, these areas are represented by officials who are members of the governing party and can thus influence decisions by the provincial chief executive. Second, the governing party may have an incentive to invest more effort in areas with a stronger base of support (Cox and Mccubbins, 1986). We use variation in the vote share of the governing party in the catchment of each police station before and after each election to identify the effect of accountability pressure on how crime reports are filed and handled, holding constant unobservable differences in the characteristics of each police station catchment area. We first demonstrate that there is no differential pre-trend in crime registration in the months before the election. At the the time of the election, the power of politicians in different constituencies to put pressure on the police bureaucracy changes; crime registration responds dramatically. Stations in catchment areas which are politically aligned with the government, and are thus subject to more accountability pressure, reduce the number of crime reports that are registered as official crimes. Our results suggest that a constituency experiencing an average increase would see a drop of six percent in the number of officially registered crimes. Because registering a report as an official crime is the first step required to move forward to investigation and arrest, a reduction in crime registrations has potential for significant consequences for justice system outcomes.

We report a series of tests to rule out the possibility that these results represent a true decrease in crime due to increased crime prevention among police stations facing greater accountability pressure. First, we use the data from a crime report database, in which citizens' initial *reports* are recorded by a professional receptionist at the front desk of the police station before being assigned to a police officer (who de facto decides whether to *register* each report as an official crime). This gives us unique leverage on the empirical challenge of systematic measurement error commonly faced in research using administrative crime data. We observe that the police do not register the majority of crime reports despite the legal requirement that they do so: 74% of violent crime and 67% of non-violent crime reports recorded by the receptionist are not officially registered. This observation of initial crime reports recorded in the database by the receptionist, yet not registered and pursued by police, allows us a uniquely precise measure of under-registration, which is typically inferred indirectly in other studies. This data structure allows us to further isolate the effect of accountability pressure. We find that when political accountability pressure increases, a crime report is *less* likely to be registered.

Second, the crime reports database also includes non-criminal loss reports, which do not require further action from the police and are not included in crime statistics monitored by senior officials; this allows us to test whether police officers "downgrade" crime reports into such loss reports. For example, a mugging should be recorded as a crime, but could instead be recorded as a loss of property and ID card. This allows the citizen reporting the mugging to officially request a new ID card and potentially make an insurance claim, but does not require the police to pursue the case further or result in an increase in crime statistics. We find that an increase in political accountability pressure leads to a shift in composition of citizen reports from property crime to non-criminal loss reports.

Third, if increased resources, better enforcement and a subsequent reduction in crime explained our aggregate results, we should expect to see political accountability pressure leading to increased staffing levels or faster response times. There is no change in these outcomes.

Taken together, these results are consistent with police becoming *less* responsive in registering crime reports by citizens, rather than a reduction in the true level of crime. The increased accountability pressure backfires because of the principal's limited ability to observe the agent's activity. This reduction in crime registration may have damaging consequences. If police decline to register or miscategorize incidents reported by citizens, rather than moving forward with registration, this not only makes the crime statistics misleading, it also means that these crimes cannot be pursued, as crime registration is required for further investigation, arrests, or court filing.

We examine the potential mechanisms for the decrease in crime registration. The effects are significantly stronger for police officers who were promoted under the current government at the time of the report. This could be because these promotions proxy for a pre-existing linkage with the political party in power, increasing the potential for politicians to exert pressure on officers to keep crime levels low.

We then examine how a shift in the principal's ability to observe the agent's activity affects the agent's actions. To answer this, we use registered crime data from an extended time period and examine the impact of the roll-out of the crime reports tracking database. We find that the overall number of registered crimes is flat in the months leading up to the implementation of the new database in each station, but starts to increase after implementation. This suggests that improved observability effectively reduces the problem of under-registration of crime reports. We estimate that the implementation of the new database leads to a one third increase in the number of crimes registered. Interacting the effect of the rollout and political accountability pressure suggests that the rollout of the new database attenuates the effect of accountability pressure on under-registration of crime.

We also examine the incidence of effects. Both effects - the negative consequence of political accountability pressure decreasing crime registration, and the positive effects of the new database increasing crime registration - are more pronounced in poorer areas of the city.

We make three key contributions to the empirical literature on principal-agent problems in public services. First, we document the phenomenon of agents manipulating the signal to the principal in the context of policing. An extensive empirical literature investigates problems of gaming in public services, including both multi-tasking (in which the agent shifts effort from an unincentivized to an incentivized activity, such as reducing effort in teaching non-tested subjects) and signal manipulation (in which the agent attempts to directly change the signal seen by the principal, for example teachers assisting students in cheating on a test). These behaviors have been extensively documented in empirical work on education (Miller and Babiarz, 2014; De Philippis, 2015; Martinelli *et al.*, 2018; Dumont *et al.*, 2008; Johnson *et al.*, 2015; Hong *et al.*, 2018; Fryer and Holden, 2013; Feng Lu, 2012; Jacob, 2005; Glewwe *et al.*, 2010; Neal and Schanzenbach, 2010; Jacob and Levitt, 2003; Lavy, 2009; Dee *et al.*, 2019; Battistin *et al.*, 2017; Figlio, 2006; Cullen and Reback, 2006; Figlio and Getzler, 2006; Diamond and Persson, 2016), healthcare (Dranove *et al.*, 2003; Werner and Asch, 2005; Frank *et al.*, 2000; Oxman and Fretheim, 2009; Gupta, 2021), road safety (Fisman and Wang, 2017; Kelley *et al.*, 2020), population policy (Suárez Serrato *et al.*, 2019), environmental policy (Ghanem *et al.*, 2020), and in the military (Acemoglu *et al.*, 2020). In addition to weakening the principal's ability to hold the agent accountable, such manipulation can be directly welfare reducing. For example, re-classification of poorly performing students as special needs to avoid their inclusion in high-stakes tests may have serious consequences for their future educational opportunities. Similarly, a reduction in crime registration in response to pressure to keep crime statistics low has real consequences: registration is required for further investigation and arrests to take place. Therefore, quantifying the extent of under-registration and understanding the impact of public policies on it is a priority. However, while a number of studies have documented such patterns qualitatively (Eterno and Silverman, 2012; Eterno *et al.*, 2016; Yu and Zhang, 1999), quantitative evidence is scarce. Existing literature in the economics of crime suggests large gaps between crimes experienced by victims and officially registered by the police (Iyer *et al.*, 2012; Miller and Segal, 2016; Amaral and Bhalotra, 2021; Kavanaugh *et al.*, 2019; Lakhtakia, 2021). However, this gap may be due to a combination of factors including victims' reluctance to report crimes due to stigma or fear of reprisal, as well as police responses such as low effort or signal manipulation. We contribute to a nascent literature quantitatively documenting police data manipulation through under-registration more precisely through "mystery shoppers" (Banerjee *et al.*, 2021), reclassification records (Eckhouse, 2022), or comparison between administrative and crowdsourced data (Cook and Fortunato, 2022). We build on this work by leveraging two administrative data sources, one comprising crime reports by citizens and one comprising registered crimes, to allow us to precisely identify crime reports that are not registered.

Second, we demonstrate that political pressure increases data manipulation. In this respect, we contribute to the literature examining the interaction between politicians and bureaucracy in determining delivery of services to citizens (Besley *et al.*, 2021; Finan *et al.*, 2015), particularly regarding politician efforts to ensure preferential treatment from the bureaucracy for the governing party's constituencies or supporters. A substantial literature finds evidence of preferential treatment for the governing party's constituencies or supporters through service delivery or regulation (Hodler and Raschky, 2014; Burgess *et al.*, 2015; Hsieh *et al.*, 2011). A few recent studies more directly pinpoint the mechanisms through which politicians influence the delivery of state services through interaction with the function of state bureaucracy (Asher and Novosad, 2017; Gulzar and Pasquale, 2017; Callen *et al.*, 2020a,b). A few studies find negative consequences of political alignment with the governing party through a mechanism of patronage exchanges between politicians and public servants (Callen *et al.*, 2020b; Das and Sabharwal, 2017). We build on this literature by documenting that political pressure on bureaucrats backfires because of limited observability, leading to a reduction in registered crimes and thus worse public service to citizens. Specifically, we provide evidence that bureaucrats promoted by the party in government drive our estimated effects. In this respect we build on previous work on the personnel economics of the state in developing countries (Finan *et al.*, 2015; Wade, 1985; Zwart, 1994; Finan *et al.*, 2015; Khan *et al.*, 2019; Xu and Burgess, 2018; Iyer and Mani, 2012; Banerjee *et al.*, 2021; Akhtari *et al.*, 2022; Lehne *et al.*, 2018).

Finally, we contribute to the literature examining the impact of policies or interventions for improving the principal's ability to observe agent effort or outcomes in delivery of public services such as infrastructure and transport (Olken, 2007; Calvo *et al.*, 2019; Kelley *et al.*, 2020), healthcare (Bjorkman and Svensson, 2009; Dhaliwal and Hanna, 2017; Callen *et al.*, 2020a) education (Reinikka and Svensson, 2011; Duflo *et al.*, 2012; Borcan *et al.*, 2017), or agricultural extension (Dal Bó *et al.*, 2021). A few recent studies examine such interventions in the context of policing (Banerjee *et al.*, 2021; Rivera and Ba, 2019; Ba, 2020; Eckhouse, 2022). These interventions either increase observability, or a combination of observability and pressure. Banerjee *et al.* (2021) test several interventions in the context of policing in Rajasthan, India. They find that improving observability through the threat of random audits by decoy citizens is effective at ensuring crime reports are registered; introducing community observers to police stations has no similar effect. Eckhouse (2022) studies the rollout of "metrics management" systems in US states, which bundle both pressure (in the form of quantitative targets) and observability (in the form of data collection). She finds an increase in data manipulation as measured by the reclassification of rape reports as "unfounded". Cook and Fortunato (2022) find that police are more likely to under-report lethal force in U.S. states with legislatures that have lower capacity in terms of pay, staffing and duration of legislative sessions. This legislative capacity measure could proxy for either increased pressure on police, increased observability because of higher capacity to collect data, or some combination of both. Our key contribution to this literature is to study the impact of both *pressure* from the principal and *observability* by the principal and separate the two in the same setting. We show that while political accountability pressure increases data manipulation, enhanced observability through the new crime reports database decreases it - demonstrating the potential for a scalable technology to mitigate the problem of observability in principal-agent settings.

The remainder of the paper proceeds as follows. Section 2 describes the context, the administrative data we use, and an overview of how this data allows us to test our hypotheses; Section 3 presents the empirical strategy and results for our analysis of accountability pressure; Section 4 presents the empirical strategy and results on increased observability; Section 5 discusses incidence of the effects; and Section 6 concludes.

2 Context and data

2.1 Lahore Police

The scope of our study is the district of Lahore, Punjab, Pakistan, a district comprising a major urban center and the surrounding rural area, with a total population of 11 million (Pakistan Bureau of Statistics, 2017). Lahore is one of 36 districts in the province of Punjab; the police force in Lahore is administered at the provincial level, by a department of the provincial bureaucracy, the Punjab Police. A total of 2,200 police officers are deployed across 83 police stations in Lahore district. Each police station has a designated catchment area; all citizen crime reports of events in that catchment area must be reported in the corresponding police station.

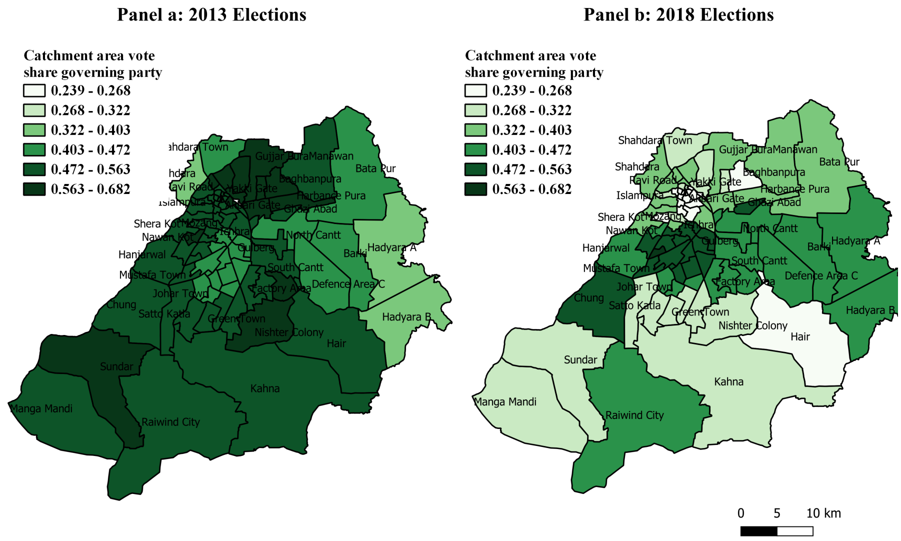
2.2 Politicians' influence over police

Because the police force is administered at the level of the province, the key elected officials that hold the police accountable are Members of the Provincial Assembly (MPAs; hereafter, we refer to MPAs as politicians). These politicians are elected in single member districts in first-past-the-post multiparty elections. Pakistan operates under a parliamentary system, thus the Chief Minister, an MPA from the governing party elected by the MPAs, is the ultimate principal to whom the police answer.

Over the two election cycles covered in our sample period, there was a change in the governing party at the national as well as provincial level; the Pakistan Muslim League-N held power from 2013, while the Pakistan Tehreek-e-Insaaf took power in 2018. Figure 1 illustrates the percentage of vote share for the winning party by constituency in the 2013 and 2018 elections.

Provincial politicians, particularly those from the governing party, play a role in legislation, budget allocation and in influencing the postings and transfers of all provincial government officials, including the police. Politicians appoint senior police chiefs and administrators. District police heads are appointed by the Chief Minister (the governing party MPA who is the political leader of the province). District heads in turn appoint police station staff including station chiefs. However, it is common practice for politicians to refer and recommend their favoured candidates (at any level) to the district head for appointment or transfers, or conversely

Figure 1: Winning Party Vote Share



Source: Authors' calculations using data from Election Commission of Pakistan (a,b).

to block the appointment or transfer of candidates (Siddiqi *et al.*, 2014; Malik and Qureshi, 2020). If a district police head does not oblige, governing party politicians can use their clout with the Chief Minister to get the district police head transferred elsewhere (potentially to an undesirable posting, e.g. in a remote or challenging location). Therefore, while individual politicians are not officially responsible for the postings, transfers and promotions of police station staff, *governing party* politicians can exert influence indirectly through their party connection with the Chief Minister.

The political leadership of the governing party at the provincial level, including the Chief Minister, periodically calls high level meetings with senior police officials to hold them accountable for official crime statistics (a press photo of one such meeting is shown in Figure A1). For instance, the Chief Minister of another province, Sindh, recently chaired a high level meeting on the law and order situation in which he stressed that "I am establishing a separate unit to monitor crime situations at Chief Minister House, therefore *crime rate must be seen coming down in all the police zones and regions of the province*" (Express Tribune (2021), emphasis added by authors). These meetings may include discussion of complaints of MPAs from the governing party in response to concerns about their specific constituencies. Politicians also directly intervene in the police stations in their constituency to pressure police to intervene in specific cases (Siddiqi *et al.*, 2014; Malik and Qureshi, 2020).

2.3 Crime registration as an outcome of interest

In this section, we lay out the logic of our key outcomes of interest, i.e. the registration of crime. We define each part of the process as we describe it; Table A1 also lists the definitions of key terms used throughout the paper.

The key outcome of the crime reporting process is the *registration* of a reported crime. In official terminology, registration implies the filing of a First Information Report (FIR), an official part of the police record; however, throughout the paper we simply refer to this as "registration". Registration is required for any further steps to be taken, including arrest, investigation and court proceedings. Thus, registration is the key initial outcome for citizens seeking justice system response to a crime. Legally, police are required to register *all* crime reports in serious categories (such as violent crime). However, de facto many reports are not registered and police officers in fact exercise discretion over whether to register the crime. Senior officials in the Punjab police have publicly expressed concern over the failure of police officers to register crime reports (Dunya News, 2021). Cheema *et al.* (2017), a team including a lead Punjab police official along with academic collaborators, report on a 2016 representative cross-sectional survey of households in Lahore, before the rollout of the reforms we study; their findings demonstrate the importance of crime registration as an outcome for citizens and the difficulties citizens face in getting to this stage. Only seven percent of respondents who reported crimes to the police were able to get their reports registered. Of those who were unsuccessful, 60% attributed this to police "unwillingness" to proceed with the registration process. When respondents overall were asked how they feel the police should be reformed, improving and simplifying the crime registration process was the fourth most common citizen response (after improving their attitude to the public, insulating them from political pressure, and reducing corruption). This reflects similar experiences from other South Asian contexts; the "mystery shoppers" Banerjee *et al.* (2021) sent to police stations in India only succeed in getting their reported crimes registered in 48% of cases. Qualitative studies of policing in Pakistan (Siddiqi *et al.*, 2014) and across South Asia (Sanjay Patil, 2008) confirm that the problem of refusal to register crimes is a first order concern for citizen access to justice in the region. Cheema *et al.* (2017) therefore proposed a reform agenda for Punjab in 2017, stating that "improving the FIR registration process must be a core objective of police reforms".

In addition to the extensive margin outcome of whether a crime is registered at all, there is also potential for intensive margin variation in crime registration, through *downgrading* of reported crimes to lower level offenses. Eckhouse (2022) gives a comprehensive review of qualitative evidence on the real world consequences of downgrading in a U.S. setting. For example, she discusses the case of a serial rapist in New York evading detection despite an escalating pattern of attacks, in part due to the fact that officers downgraded the rape cases to misdemeanours such as criminal trespassing, and thus the pattern of rapes was not detected for some time.

Thus we treat both the registration of a crime and the classification of crimes in police records as meaningful outcomes of interest in our study.

2.4 Crime reporting and registration before and after the CMS crime report tracking database

To address the concern of under-registration of crime, in 2017, Punjab Police launched a new crime report tracking database, the Complaints Management System (CMS). This electronic database is used to record, track and monitor the progress and actions taken against each citizen crime report (Punjab Information Technology Board, 2017; Rizvi, 2017). Table 1 provides an overview of how citizen crime reports are handled by the police before and after the introduction of the new database.

In response to a crime, a citizen makes a report to the police station corresponding to location of the incident. Both before and after the rollout of the new database, upon entering the police station, a citizen is first received by a professional receptionist or "station assistant". These receptionists are not recruited or trained for policing jobs but are recruited separately as administrative professionals for performing clerical duties within the organization. Crucially, these positions fall outside the police hierarchy and are not subject to the transfer or promotion decisions used to incentivize police officers.

Once the citizen makes his or her report, it must then be recorded. Before the introduction of the new database, crime reports were logged by a station clerk, a uniformed officer in the regular police force, who has been recruited, trained and posted in the same way as other police officers. Station clerks are typically two ranks below the station head, and their relationship with the station head has implications for retaining the existing posting, their subsequent promotion and future career path. The station clerk was supposed to record each crime report in a manual logbook. He then prepares a duplicate of the crime report on paper and attests it with a signature and official seal of the police station as a record for the citizen.

In contrast, once the new database is rolled out, crime reports are recorded by the professional receptionists described above, who enter each report in the electronic CMS database. This results in a system generated receipt that is printed and handed to the citizen with a tracking number to be used for subsequent follow-ups. The citizen also receives a system-generated text message confirming the receipt of his/her crime report by the police. Figure A2 shows an image of the front desk process once the CMS database is rolled out. Unlike the crime registration process, the reporting database appears to cover most cases reported: Cheema *et al.* (2017) report from a survey after the new database was rolled out that 70% of incidents were recorded by police.

Thus, while the staffing of the station remains the same with the introduction of the new database, the task of logging the initial crime report is shifted between the staff: from a uniformed police officer within the police hierarchy, to an a professional receptionist outside the police hierarchy.

Once a crime report has been recorded, an investigating officer conducts a follow-up inquiry to establish the authenticity of the reported offense. The findings of this preliminary inquiry are reported to the station head (Station House Officer), who makes the final decision on whether follow-up action is to be taken.

Throughout the period we study, the number of *officially registered crimes* is a key indicator tracked and used by senior police officials and politicians to assess the prevalence of crime and to gauge performance of the police in each area. In contrast, before the rollout of the new database, initial *crime reports* by citizens were neither reported to nor monitored by superiors. A record of such crime reports could only be checked by reviewing the logbook on an in-person monitoring visit to the police station. Thus, police officials facing scrutiny for high crime rates likely had an incentive to avoid registering initial crime reports, or to reclassify crime reports as non-crime reports (such as recording a mugging as a loss of wallet and ID card), which do not require the police to pursue the case further or result in an increase in the registered crime statistics monitored by senior officials.

After the new database rolled out, senior officers have real time access to information both for any individual crime report, as well as aggregate statistics on crime report figures. Figure A3 shows screenshots of the dashboard interface and the overview a senior officer has of a specific crime report, with a more detailed report available for download. We hypothesize that the rollout makes it more difficult or risky for police officers to avoid registering a crime, because the database provides an accessible trail of information about the officer's actions that superiors can access with ease at any point in time during or after the case.

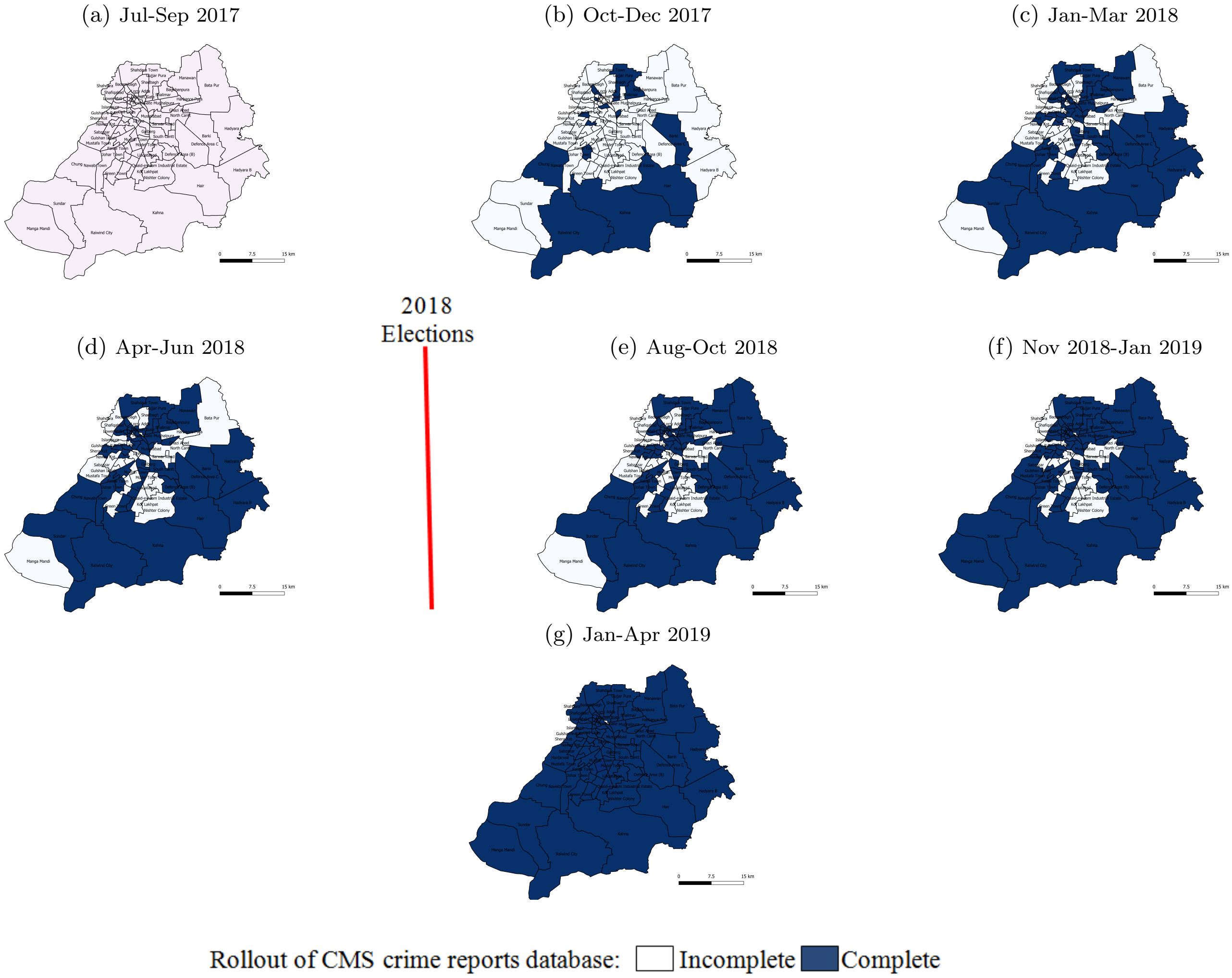
Table 1: Overview of crime reporting and registration process pre and post CMS crime report tracking database

Pre CMS	Post CMS
Citizen makes a report to the police station corresponding to location of crime / event.	
A civilian front-desk staffer outside the police recruitment and promotion system receives the citizen.	
A uniformed officer within police recruitment and promotion system records the report.	The civilian front-desk staffer outside the police recruitment and promotion system records the report.
The report is recorded in a paper logbook.	The report is recorded electronically into the CMS database and a tracking number is issued.
Complaint is followed up by a preliminary probe. Findings are reported to the station head, who decides whether to register the crime by filing a First Information Report – an official document required for arrest or further action.	
In-person station visit to access hard copy records required for monitoring crime reports	Senior officers have real-time access to an electronic database of crime reports with a summary dashboard.
Statistics on registered crimes (FIRs) by area are discussed at regular high level meetings.	

The new database was officially activated in 2017, but its practical implementation was rolled out to some police stations before others for logistical reasons (Figure 2). Figure 3 shows the timeline. In Section 4, we will compare the stations with earlier and later rollout of the new database and show that they are similar in terms of observable characteristics and trends in the outcome variable.

Note that none of the administrative data in our setting are released to the public at any geographic level below the level of the district (the entire area we study comprising 11 million people). The data from the new database were not released to the public at all over the time period we study. Thus, citizen or media monitoring of aggregate crime levels are not expected to be a mechanism for the effects we study. Therefore in our setting, we anticipate that politicians do not have an incentive to pressure police in their constituencies to manipulate crime registration numbers downward in the interest of improving the image of their performance to the public. Such changes in an individual consistency would be expected to have a negligible impact on the aggregate district-level figures released to the public. Rather, politicians are likely motivated to improve the perception of the voters in their constituencies based on media reports of individual cases, public experiences of crime and police performance. Politicians and the senior officers who answer to them may thus exert pressure on police based on registered crime information shared in detailed internal briefings. Thus in our setting, we consider both elected politicians and senior officers as principals, trying to hold field-level police officers (agents) accountable for keeping crime levels low. Before the new database rolls out, they do so by monitoring registered crime statistics; after it rolls out, they can supplement this with checks on crime reports from the dashboard. We will provide further empirical evidence for this point in Section 4.4.

Figure 2: Rollout of the new crime report tracking database



Notes: We classify a station as completing the rollout of the new database when 95% of crimes in the crime registration data appear in the new database.

2.5 Data

We use two main sources of administrative data from the Punjab police: microdata on all 135,000 registered crimes in Lahore district between January 2016 - December 2020; and microdata from all 690,000 crime and loss reports in the crime report tracking database from January 2017-August 2019.[1]

Tables 2 - 3 present descriptive statistics. Our primary outcome is the number of *registered* crimes in a given station, category and month. For each registered crime, we observe the date, the police station catchment area; and the category of offense. We complement this with data from the crime report database, which covers all citizen *crime reports* logged by police in the district of Lahore from the inception of the database in 2017 through the end of 2019. For each crime report that is recorded in the database, we observe the same variables observed in the registered crimes data (date, the police station catchment area, the category of offense), as well

[1] We exclude from analysis database records corresponding to other types of reports or requests: 2,927 requests for police “character verification” certificates, i.e. certificates from the police indicating that a person has no criminal record, used for applications for government employment and visas; and 45 records corresponding to complaints lodged against the police.

as the officer who handled the case and the response time, i.e. the duration between when the crime report is lodged and the first subsequent follow-up contact by the investigating officer. We also observe whether each crime report was registered.

The raw data include information on the category and offense of each entry. In our specifications using aggregated data, we aggregate to the level indicated as Level 3 in Table 2. In the crime report tracking database, approximately half of the records are loss reports, which are not eligible for registration. Therefore, for most of our analysis, we focus on crime reports and exclude these loss reports from the analysis.[2]

We use the crime report tracking database in two ways in our analysis. First, we use data from within the new database, which allows us to identify crime reports that were not registered, to help pinpoint the impact of political accountability pressure on the under-registration of crime. In all such specifications analyzing administrative data taken from within the crime report tracking database itself, we restrict the dates covered to the period when the database rollout covered most of the district; we also present robustness of these analyses to restricting analysis to only stations where the new database had been rolled out throughout the period we study. For this analysis, we restrict the sample to the years covered in both the new database and registered crimes data after database coverage in the district was complete. Figure 3 shows the timeline of data coverage. While the database may not capture all cases where a citizen attempted to report a crime, it clearly captures a large number of such cases: Table 2 shows that three quarters of violent crime reports in the database are not registered, including crimes as serious as murder. Two thirds of non-violent crime reports are not registered.

Second, we examine the effects of the database rollout on crime registration, covering the period both before and after its rollout. In these specifications, the data analyzed do not come from the new database but rather from the database of *registered crimes*. In these cases we are examining the impact of the database rollout itself; thus we extend the time period for our estimation sample, and include a baseline period before the database was rolled out.

The 83 police station catchment areas and 35 electoral constituencies in Lahore district do not correspond geographically. To assign the electoral outcomes to each given police station catchment area, we conduct a spatial join between the two GIS shapefiles and calculate the proportion of each catchment area that falls in each electoral constituency. We then assign a weighted average:

$$VoteShare_{kt} = \sum_{c} \frac{OverlapArea_{ckt}}{TotalArea_k} \frac{GoverningVotes_{ct}}{TotalVotes_{ct}}$$

where $OverlapArea_{ckt}$ is the area in the catchment area of police station k that falls in constituency c, $TotalArea_k$ is the total area in the catchment area of police station k. $GoverningVotes_{ct}$ is the number of votes cast in constituency c for the governing party in the last election, and $TotalVotes$ the total number of votes cast in the constituency.

[2] Approximately 14% of reports lack information on categorization and are labeled simply as "other", likely due to error or haste by personnel entering the data. We include these in our main analysis but exclude them when presenting heterogeneity by category.

Figure 3: Timeline

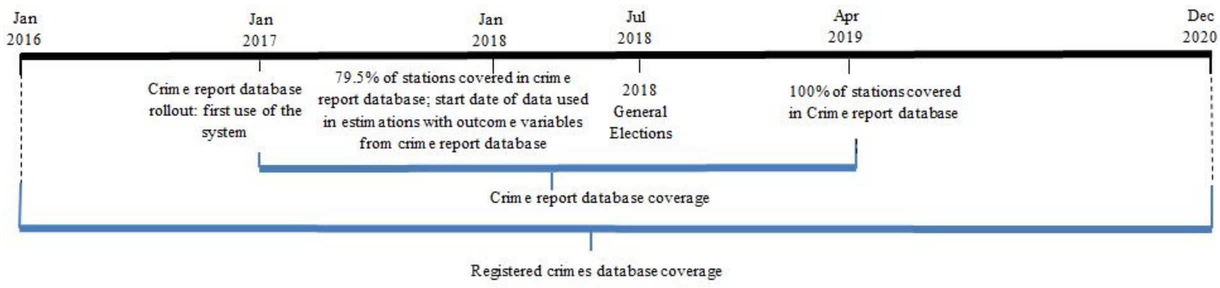


Table 2: Report Categories and Descriptive Statistics from crime tracking database and registered crime data

Categories	(1) Percentage of registered crimes database	(2) Percentage of crime reports database	(3) Percentage that received registered crime in Crime reports database	(4) Response time (hours) in Crime reports database
		(a) Crime		
1. Violent	*9.09*	*8.06*	*26.39*	*73.02*
2. Crime against person	8.56	5.74	19.52	71.67
3. Hurt and torture	1.85	3.49	7.67	79.54
3. kidnapping and illegal confinement	1.26	1.44	39.56	59.87
3. Gender abuse miscellaneous	4.07	0.51	20.09	69.78
3. Murder & attempted murder	1.51	0.20	75.87	31.51
3. Domestic violence	0.002	0.10	27.40	58.46
2. Crime against property	0.53	2.32	43.42	76.35
3. Dacoity (armed roberry)	0.53	2.32	43.42	76.35
1. Non-violent	*60.23*	*29.47*	*32.60*	*75.31*
2. Crime against person	0.34	4.12	5.33	84.69
3. Extortion & threat	0.14	3.84	2.99	86.50
3. Accidents	0.20	0.27	38.40	58.98
2. Crime against property	43.83	21.83	29.92	82.72
3. Minor theft	24.25	12.18	32.85	81.09
3. Financial crime	3.97	6.36	13.47	88.53
3. Vehicle theft	7.48	2.57	48.51	80.41
3. Property	2.88	0.41	3.25	93.61
3. Burglary	5.88	0.31	32.48	80.15
2. Local and special laws	15.71	3.36	90.64	13.48
3. Local and special laws (e.g. Electricity Act, Anti Narcotics Act, Arms Ordinance Act etc.)	15.71	3.36	90.64	13.48
2. Other offense	0.60	0.13	92.09	15.78
3. Other offense (e.g. Over speeding, Sectarianism etc.)	0.60	0.13	92.09	15.78
1. Misc./Unspecified	*29.22*	*15.49*	*38.95*	*56.37*
2. Local and special laws	29.36	1.47	94.06	8.40
3. Local and special laws (e.g. Kite Flying Act, Price Control Act, Local Government Act etc.)	29.36	1.47	94.06	8.40
2. Unspecified Other offense	0.29	14.00	33.13	61.38
3. Unspecified Other offense (e.g. miscellaneous, other crime etc.)	0.29	14.00	33.13	61.38
	(b) Loss and insurance reports			
1. Loss and insurance reports	*N/A*	*47.02*	*N/A*	*13.88*
N	134,873	688,487	365,818	688,487

Note: Table shows data for our main estimation sample, i.e. the years between January 2018 to April 2019: the years covered in both the crime reports database and registered crimes data after database coverage in the district was at least 95%, and at least 80% of stations were fully covered.

Table 3: Descriptive Statistics (aggregated data)

	Mean	SD	Min	Max
Crime reports database by station-crime type-month				
Number of crime reports	27.32	67.61	0	759
Number of crime reports (Inverse hyperbolic sine)	2.22	1.95	0	7
Observations	25,232			
Registered crime database by station-crime type-month				
Number of registered crimes	5.64	11.62	0	186
Number of registered crimes (Inverse hyperbolic sine)	1.33	1.43	0	6
Observations	23,904			
Vote share database by station				
(2013) Pre-period political alignment	0.52	0.06	0	1
(2018) Post period political alignment	0.37	0.09	0	1
Observations	83			

2.6 Hypotheses

This rich data structure allows us to test for data manipulation by police officers under pressure to keep crime statistics low.

Note that the principal (politician and senior bureaucrats) observes *registered* crimes, but does not observe crimes reported to the police or other police activity to address these reports. Thus, if the principal rewards or punishes the agent based on the crime levels he observes, police officers face an incentive not to register all crimes reported to them ("under-registration"), and to register crimes as non-crime loss reports, which do not appear in the registered crimes statistics ("downgrading").

Thus, we hypothesize that when a politician has greater power to put accountability pressure on the police force, this will result in (1) a decrease in *registered* crimes, (2) a decrease in the proportion of crime reports that are registered, and (3) an *increase* in non-crime loss reports. We present our empirical strategy and results for these hypotheses in Section 3.

The new database allows principals greater ability to observe crimes reported to the police and other police activity to address these reports. Thus we hypothesize that the rollout of the database will reduce police officers' incentive to under-register and downgrade crime reports. Thus we hypothesize that (4) the rollout of the new database is associated with an increase in registered crimes, and (5) an attenuation of the relationship between political pressure and crime registration. We present our empirical strategy and results for these hypotheses in Section 4.

3 Impact of change in principal's power to hold agents accountable

3.1 Empirical strategy

We exploit variation in governing party vote share in the station catchment area over two election cycles to assess the impact of varying political accountability pressure on registered crime. The vote share of the catchment area proxies for accountability pressure for two reasons. First, catchment areas with a higher governing party vote share have a larger proportion of their area represented by governing party legislators (note that catchment area and constituencies have overlapping boundaries, as discussed in Section 2.5). Second, catchment areas with a higher governing party vote share represent more "core" areas, which are expected to be a target for governing party attention (Cox and Mccubbins, 1986).[3] There are two sources of variation in the independent variable: (i) a change in governing party, and (ii) changes in vote share of the parties in each constituency.

For this analysis, we use both registered crime data and crime reports from the crime report tracking database. Therefore, our estimation sample consists of data between January 2018 to April 2019, i.e. the years covered in both the new database and registered crimes data after database coverage in the district was at least 95% overall, and at least 80% of stations were fully covered by the database. Figure 3 shows the timeline of data coverage.

We examine effects of political accountability pressure on the number of registered crimes aggregated by crime category j, station k and month t. We estimate:

$$RegisteredCrimes_{jkt} = \beta_0 + \beta_1 Voteshare_{kt} + \alpha_j + \eta_k + \mu_t + \epsilon_{jkt} \tag{1}$$

We report results for two ways of aggregating the level of registered crime, $RegisteredCrimes$ in category j, station k, in month t: first, the rate per 1,000 population; and second, using an Inverse Hyperbolic Sine transformation to the number of registered crimes. On the right hand side, $VoteShare$ is a continuous variable representing vote share for the governing party in station catchment area k; this varies by time t as data span two election cycles. We include fixed effects for crime report category (α), station (η) and month (μ). We cluster standard errors at the police station level, and additionally report Conley (1999) spatially clustered standard errors in main results tables. Seventy-two percent of stations experienced a change of at least five percentage points in the vote share of the governing party in their catchment areas, and 63% experienced a change of at least ten percentage points in vote share. Thus the variation in our independent variable is substantial across most clusters in the data.

Our identifying assumption is that crime registration does not change at the time of the election differentially in areas with a greater change in vote share for the governing party, for other reasons than the change in political pressure on the police.

To examine pre-trends and visualize the results, we also estimate an event study version of 1:

$$RegisteredCrimes_{jkt} = \beta_0 + \sum_{s \neq -1} \beta_s \Delta VoteShare_k \times I[t = s] + \alpha_j + \eta_k + \mu_t + \epsilon_{jkt} \tag{2}$$

Where s indexes months before / after the 2018 election. The coefficients of interest are β_s, which are the counterpart to β_1 in 1; these estimates capture effects over each month before and after the election. Note that

[3] In Section 3.4 we show a similar pattern of results when we use a specification that captures the posting of governing party legislators, but not when we examine "swing" constituencies, following the main alternative mechanism discussed in the theoretical literature, that elected officials target areas that are more closely contested for attention and support Lindbeck and Weibull (1987).

$\Delta VoteShare_k$, the change in vote share at the time of the election, is not included because the specification includes catchment area fixed effect η_k. Note that 1 - 2 do not represent a staggered treatment event study (Goodman-bacon, 2021), since there is a single point in time at which the independent variable changes: the 2018 election.

Our second specification uses the microdata from the crime reports database at the individual crime report level i to study the effect of governing party vote share on the probability that an individual crime report is registered and the response time. We estimate:

$$Registered_{ijkt} = \beta_0 + \beta_1 VoteShare_{kt} + \alpha_j + \eta_k + \mu_t + \epsilon_{jkt} \tag{3}$$

Where $Registered_{ijkt}$ is an indicator for whether crime report i in crime category j, station k and month t was officially registered.

We also aggregate data at station k-month t level to study the effect of political accountability pressure on staff inputs (measured by police active in station k, month t).

3.2 Results

In this section, we present results on the impact of a change in political accountability pressure on registered crime. Columns 1-2, Table 4 show the results of Equation 1. Fewer crimes are registered in areas that see an increase in governing party vote share, where accountability pressure is higher. The point estimate in Column 2 suggests that if a police station's catchment area goes from no support to complete support for the governing party, the station would register 36 percent less crime. However, since no area actually shifts between zero to 100 percent vote share, we rescale the point estimate based on the mean absolute value change in vote share, i.e. 16 percentage points (Table 3). This suggests an effect size of 6 percent less crime registration. Figure A4, Panel (a), breaks down these results by violent and nonviolent crime; the results are driven by nonviolent crime.

3.3 Ruling out improved crime prevention

Could the results presented so far suggest that politicians simply work to improve services for areas that supported them, consistent with the "pork barrel politics" literature, thus resulting in a real decrease in crime? Here we present a series of tests, each of which is inconsistent with this interpretation.

Crime reports are less likely to be officially registered: If improved policing resources and prevention of crime explained our main result (in columns 1-2, Table 4), we would expect to see a crime report to be more likely to be registered. If improved public services in other domains (e.g. improved social services), we should expect no change in the probability that a crime report would be registered. In fact, we find that areas with a mean absolute change in vote share ($\approx$16 percentage points) have a 2.2 percentage point lower probability of registering a given report, or a 7% decrease compared to the sample mean (column 3, Table 4).[4]

No change in response times: If improved policing and crime prevention explained our main result (in column 1, Table 4), we would expect to see a reduction in response time. However, we there is no effect on the speed of police response to reports (column 3, Table 4).

No change in staffing inputs: If improved policing and crime prevention explained our main results, we would expect to see an increase in staffing inputs. To test for this, we construct a measure of police active in each station-month based on the administrative records. There is no effect on staffing levels (column 4, Table 4).

"Downgrading": Non-crime loss reports increase as property crimes reported decrease: To test for downgrading, we examine whether crime reports shift from minor property crimes to property loss reports. We hypothesize that property crimes have the potential to be downgraded to loss reports; therefore, we construct a new estimation sample composed of minor property crimes (eligible for crime registration) and loss reports (not eligible). We then estimate an interacted version of Equation 1:

$$\begin{aligned} ReportedCrimes_{jkt} = \beta_0 &+ \beta_1 VoteShare_{kt} \times MinorPropertyCrime_j \\ &+ \beta_2 VoteShare_{kt} \times LossReport_j + \alpha_j + \eta_k + \mu_t + \epsilon_{jkt} \end{aligned} \tag{4}$$

Areas with a mean absolute change in vote share register fewer minor property crime reports and correspondingly more property loss reports (Figure 4). This suggests substitution from property crime reports to non-crime loss reports, consistent with downgrading rather than improved crime prevention in areas with higher support for the governing party.[5]

3.4 Other robustness checks

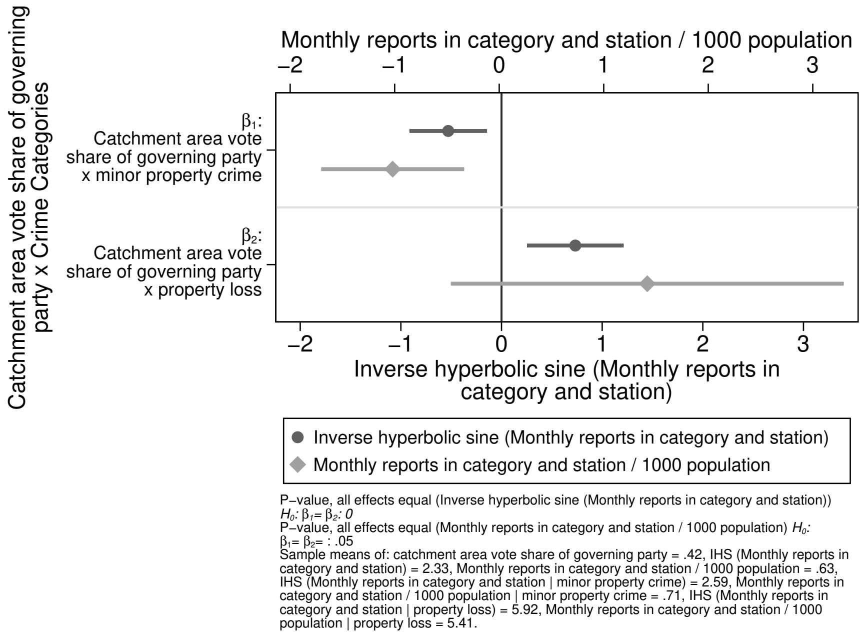
We present event study versions of our estimates (Equation 2) for all four outcome variables i.e. number of crimes registered, probability of registration, response time and staff inputs in Figure A7. Each event study figure also shows the difference-in-differences estimate (Equation 1) for reference. We observe no significant pre-trends in any of the outcomes, and the results are similar across specifications.

One potential concern is that the rollout of the new crime database could affect the measurement of the crime registration outcomes. To address this, we further restrict the sample to the subset of 66 stations in which the new database had already rolled out by January 2018 (7 months before the election). Table A2 shows the results, which are similar to our main estimates.

[4] To rule out the possibility that the decrease in registered crime is not because the police in governing party areas become more responsive in initially recording e.g. more minor citizen crime reports, with a smaller proportion of the aggregate being registered, we estimate Equation 2 with the number of crime reports as the dependent variable. The results vary between different constructions of the dependent variable. The results for registered crimes per capita show no impact on crime reporting in the database, while the inverse hyperbolic sine results do show a reduction (Figure A5).

[5] Figure A6 shows that this result is robust to excluding a three month window around the election.

Figure 4: Effect of political pressure on "downgrading" property crime reports into loss reports



Notes: Coefficient estimates and 90% confidence intervals are shown. Unit of observation is crime report category j by police station k by month t (N = 3,984). The dependent variables are IHS (inverse hyperbolic sine) of the number of reports per report category per police station catchment area per month and number of reports per 1000 population per report category per police station catchment area per month. Coefficient estimates are for the interaction terms between catchment area vote share of the governing party (in 2013 and 2018) and indicator variables equal to 1 for the respective category j; *property loss* and *minor property crime*. All estimates include fixed effects for station, month and broader crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment area.

To address the potential concern that election related disruption could drive our results, we re-estimate Equation 1 dropping a three month window around the election. The results are similar to our main estimates (Table A3).

Our results are also robust to various alternative methods of aggregation, including aggregating all time periods to pre / post election following Bertrand *et al.* (2004) (column 3, Table A4) and aggregating at station and month level (column 4, Table A4).

We also estimate a version of 1 in which we redefine the independent variable as the proportion of the police station catchment area which falls in constituencies in which the governing party won the seat, rather than the weighted vote share; the results are similar to our main estimates (Table A5). [6]

3.5 Mechanisms: Political connections and promotions

One potential mechanism through which the political environment may affect work done by the police is through connections between political parties and individual officers. Politicians are thought to advance the posting, promotion and favorable transfer of individuals connected with them, in exchange for those officers' cooperation in the politician's wishes in police work. To investigate this, we examine heterogeneous effects by the political

[6] We do not see a similar pattern when we examine the alternative hypothesis, that attention is targeted to "swing" or closely contested areas (Lindbeck and Weibull, 1987), by estimating 1, replacing the independent variable with the vote margin.

Table 4: Effect of Political Pressure

	(1) IHS (Monthly registered crimes in category and station)	(2) Monthly registered crimes in category and station / 1000 population	(3) Probability of crime report registration	(4) Response time (Hours)	(5) Number of officers in station and month
Catchment	-0.254***	-0.047*	-0.139**	0.760	1.547
area vote share	(0.087)	(0.024)	(0.064)	(7.538)	(1.833)
governing party	{0.059}	{0.016}	{0.037}	{3.987}	{1.199}
Observations	22576	22576	365818	688487	1328
Sample Mean	1.406	0.130	0.335	43.41	14.31
Obs. unit	cat-station-month	cat-station-month	crime-report-level	crime-report-level	station-month
Category FEs	yes	yes	yes	yes	
Station FEs	yes	yes	yes	yes	yes
Month FEs	yes	yes	yes	yes	yes
Clustering	station	station	station	station	station

Notes: *, **, *** denote significance at 10%, 5% and 1% respectively. Unit of observation is crime report category j by police station k by month t in columns 1 and 2, where data are aggregated to categories shown as Level 3 in Table 2. Unit of observation is individual crime report for the sub-sample of crime categories eligible for registration in column 3, individual crime/loss report for the overall sample including categories not eligible for registration in column 4, and number of officers active in station k by month t in column 5. The dependent variable is IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month in column 1, number of registered crimes per 1000 population per report category per police station catchment area per month in column 2, the probability of a crime report being registered in column 3, time taken (in hours) by police to first respond to a crime report in column 4 and number of officers active in station and month in column 5. All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2) except for those in column 5 where fixed effects are for station and month. Standard errors are in parenthesis and clustered by police station catchment areas. Spatially clustered standard errors are given in braces.

timing of officers' career progression within the police.

Because we do not have access to direct staff records, we impute promotions from the crime reports database. After cleaning and deduplicating officer names, we observe 1,912 unique officers who appear in at least three different months over the two year period, corresponding to 98% of records in the database.[7] When an officer appears in a database record with a higher rank than his previous records, we infer that he has been promoted as of the date of the new record. We exclude officers who are already senior the first time they are observed in the dataset and therefore are not observed when promoted. We use this approach to categorize officers into groups based on their promotion under the first (PML-N) or second (PTI) government in the period we study, and take this as a proxy for their connection with that political party. We then convert these to time-varying indicators for officers promoted under the current government at time t, under the opposing government, or never promoted. Thus for crime reports before the election, an officer who will in the future be promoted by the second government is tagged as an officer allied with the opposition. For crime reports after the election, the same officer will be tagged as promoted by the government. We then interact these with Equation 3, including officer fixed effects:

[7] This is slightly lower than the official statistic of 2,200 officers posted in the district, likely because not all officers handle crime reporting.

$$
\begin{aligned}
Registered_{ijkpt} = \beta_0 &+ \beta_1 VoteShare_{kt} \times PromotedGovt_{pt} \\
&+ \beta_2 VoteShare_{kt} \times PromotedOpposition_{pt} + \beta_3 VoteShare_{kt} \times NotPromoted_{pt} \\
&+ \alpha_j + \eta_k + \mu_t + \psi_p + \epsilon_{jkpt}
\end{aligned} \tag{5}
$$

As before, the outcome of interest is the probability of a crime report i handled by officer p being registered. Figure 5 shows the results. The effect of political accountability pressure reducing crime registration is driven by officers who are promoted under the *current* government at the time of the crime report. This could be because these promotions proxy for a pre-existing linkage with the political party in power, increasing the potential for politicians to pressure officers on crime registration rate. Alternatively, promotion could be a reward for keeping registered crime rates low. In either case this result is consistent with a mechanism of political involvement in the bureaucratic promotion process.

One potential concern with this test is that officers promoted at an earlier date in the sample period have different career incentives to register crime, for example if they expect not to be considered for promotion again. Note that this is already addressed by the symmetry of the test. Those who are promoted by the first government in power are considered as "promoted by government" pre-election, and "promoted by opposition" after the election; the converse also holds for those promoted by the second government in power. As a further robustness check, we also replicate this test focusing attention on officers who are promoted within a shorter duration: three months before and three months after the election. This ensures that this group of officers are similar in terms of the timing of their career progression, and simply differ in terms of the government in power when they were promoted. Figure A8 shows the results, which are similar in magnitude and statistical significance to our main results.

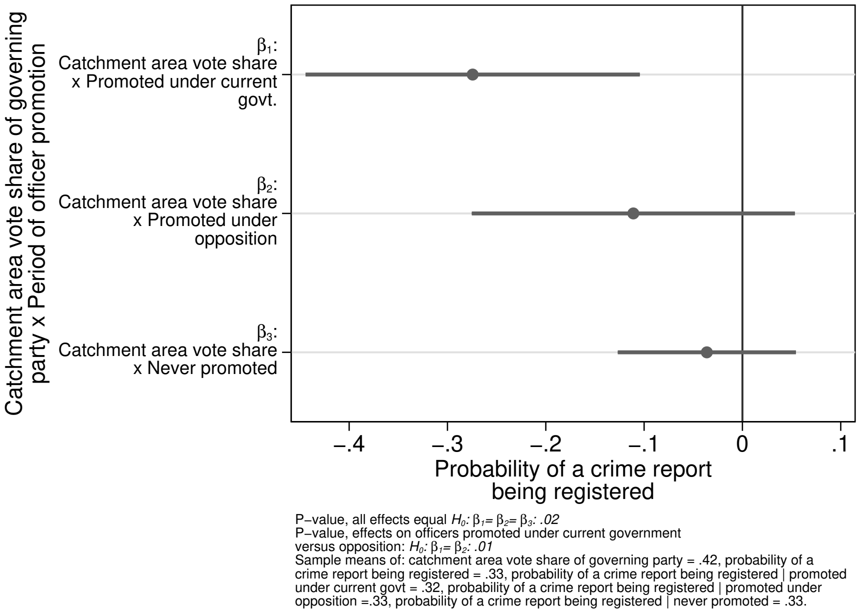
4 Impact of an increase in observability

4.1 Empirical strategy

To examine the impact of increased observability, we use registered crime data from an extended time period and examine how roll-out of the new crime report tracking database affects registered crimes. As discussed in section 2.4, prior to the database rollout, registered crime levels were monitored by senior officials, but crime *reports* could only be monitored at great time cost by reviewing the manual logbook through an in-person police station visit. In contrast, the new database affords senior officials real-time access to information on crime reports and followup action taken by police officers.

In this part of the analysis, we are examining the impact of the database rollout itself; thus, the data analyzed do not come from the new database but rather from the database of *registered crimes*. We extend the time period for our estimation sample, and include a baseline period before the database was rolled out.

Figure 5: Heterogeneous Effects of Political Pressure on Crime Registration by Officer Promotion



Notes: Figure shows coefficients from Equation 5. Coefficient estimates and 90% confidence intervals are shown. Unit of observation is an individual crime report. The dependent variable is the probability of a crime report being registered (N = 365,818). Coefficient estimates are for the interaction terms between catchment area vote share of the governing party (in 2013 and 2018) and indicator variables equal to 1 for respective category of officers; those promoted by the first government in power pre-election considered as *promoted under current government*, those promoted after the election considered as *promoted under opposition* and those who were *never promoted*. All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment area.

Table A6 compares the characteristics of stations where the new database was rolled out earlier for logistical reasons (prior to October 2017) versus later (October 2017 onwards). Areas where coverage was completed early appear very similar on observables to those where it was completed later. Most importantly, we do not observe a pre-trend in our main outcome variable, the number of registered crimes (Figure 6).

To visualize geographical variation in the timing of rollout of the new database, we map take-up of the system by station catchment area in each quarter since its launch (Figure 2). We hypothesize that coverage must be mostly complete for the new database to discourage under-registration; if a station is still testing out the system, officers would be less likely to anticipate being held to account for missing data. Thus, we consider a station catchment area to be fully covered once at least 95% of the registered crimes data are observed in the database:

$$StationCovered_{kt} = \frac{CrimesRegistered_{kt}^{ReportDatabase}}{CrimesRegistered_{kt}^{RegisteredCrimesData}} > 95\%$$

Figure 2 shows the rollout of the system based on this threshold.

To investigate how adoption of the new crime report tracking database affects police response to crime, we use a standard difference-in-differences (DD) event study design (shown by specification 6).We study effects of adopting the new system on the number of crimes registered in category j, station k and month t; we estimate:

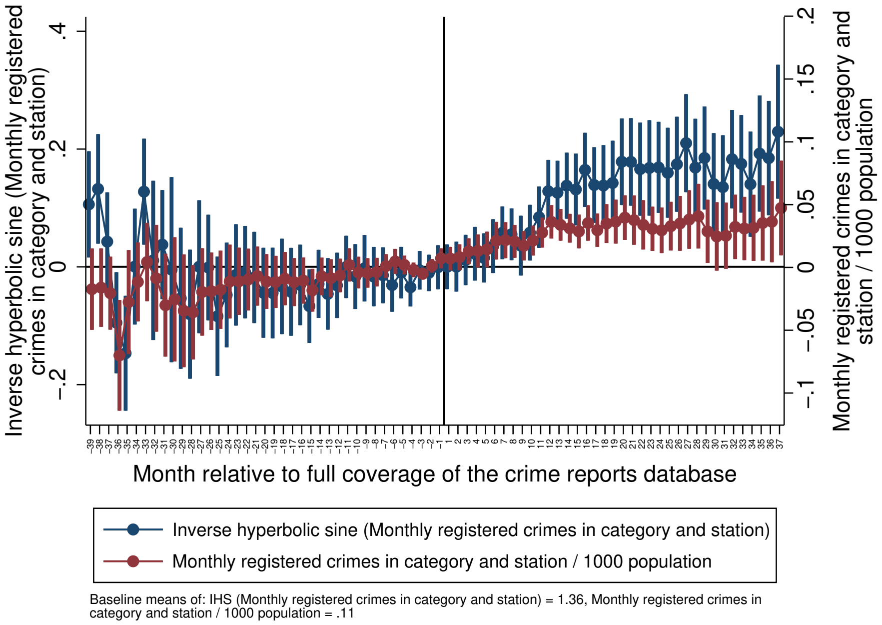
$$RegisteredCrimes_{jkt} = \sum_{s \neq -1} \beta_s I[t = s] + \alpha_j + \eta_k + \mu_t + \epsilon_{jkt} \tag{6}$$

Where s indexes months before / after station j reaches 95% coverage. We control for crime report category α, station η and month μ fixed effects; we cluster errors at the police station level. Because this is a staggered rollout difference-in-differences setting, we also present robustness to Callaway and Sant'Anna (2021) and Borusyak *et al.* (2021) estimators.

4.2 Results

The results of Equation 6 are shown in Figure 6. The X-axis shows the months before / after complete coverage of the electronic database, while the Y-axis shows the impact on crime registration. There is no detectable pre-trend before the new crime report tracking database was introduced. When the new database is implemented, registered crime each month increases. The number of crimes registered per category per month is approximately one third higher after the rollout of the new system. This is consistent with greater difficulty in under-registering crimes after the new system has been implemented. Figure A9 disaggregates the effects of the CMS database by violent and non-violent crime. The new system increases registration of both violent and non-violent crimes, with a proportionally larger effect on non-violent crimes.

Figure 6: Effect of Crime Report Tracking Database Coverage on Number of Registered Crimes



Notes: Coefficient estimates and 90% confidence intervals are shown for Equation 6. Unit of observation is crime report category j by police station k by month t for the sub-sample of crime reports in categories eligible for registration (N = 84,660). The dependent variables are IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month and number of registered crimes per 1000 population per crime report category per police station catchment area per month. Data are aggregated to categories shown as Level 3 in Table 2. Coefficient estimates are for I_{t+s}, which denotes the number of months before/after coverage of the system is complete i.e. 95% or more of registered crimes appear in the electronic system of a station k. Estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment area.

4.3 Robustness checks

To address the potential sensitivity of our results to how we define full coverage of the new database, we vary the cut-off used from 95% to 90%. The results are similar (Figure A10a). We also show robustness to alternative approaches to aggregating the data, collapsing at the station-month level instead of crime report category-station-month level (Figure A10b).

The roll-out of the new database took place over several months (Figure 2); thus there is potential for staggered treatment timing to bias the results (Goodman-bacon, 2021). To address this problem, Figure A10c shows results estimated using the Callaway and Sant'Anna (2021) approach; while those estimated using the (Borusyak *et al.*, 2021) approach are shown in Figure A10d. The results are similar to the main estimates.

4.4 Interaction of pressure and observability

Finally, we examine how accountability pressure and observability interact. We test how the rollout of the new crime report tracking database interacts with the effect of political accountability pressure. This also presents another opportunity to confirm our interpretation of the results as under-registration of crime reports. If the effect of political pressure on registered crimes represents increased effort and a true decrease in crime, we should expect better observability to intensify the political pressure effect as politicians can better track police officer effort. If in fact it represents a reduction in *registration* of crime, we should expect improved observability to attenuate the effect as it becomes more different to under-register crimes without detection.

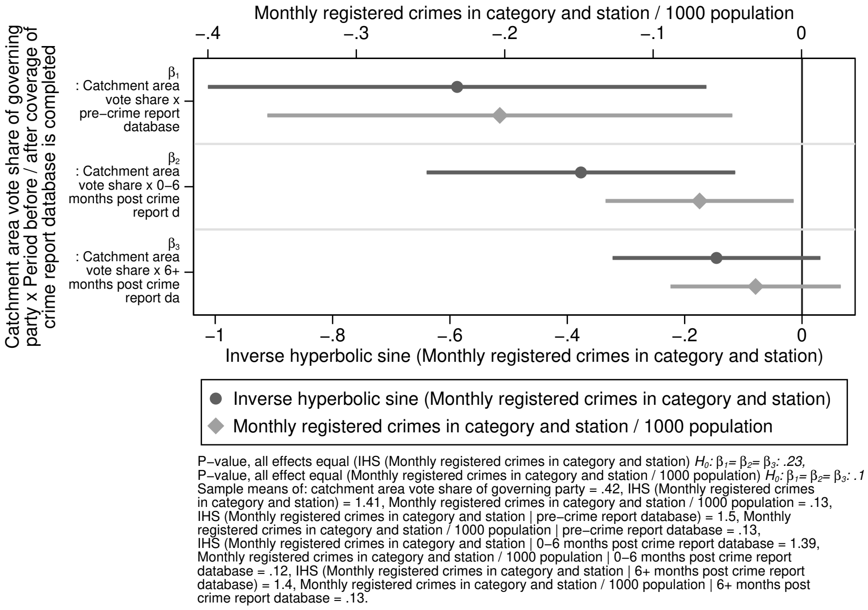
We estimate a version of Equation 1 interacted with the rollout of the new database:

$$
\begin{aligned}
RegisteredCrimes_{jkt} = \beta_0 + \sum_{s \neq -1} \gamma_s I[t = s] \\
&+ \beta_1 VoteShare_{kt} \times I[s \leq 0] \\
&+ \beta_2 VoteShare_{kt} \times I[0 < s \leq 6] \\
&+ \beta_3 VoteShare_{kt} \times I[6 < s] \\
&+ \alpha_j + \eta_k + \mu_t + \epsilon_{jkt} \qquad (7)
\end{aligned}
$$

Where again s indexes months before / after station j reaches 95% coverage. Figure 7 shows the results. The pattern of results suggests that the effect of political accountability pressure is attenuated as the new crime report tracking database is rolled out. The magnitude of the political pressure effect is less than half as large six months after the new database rolls out in a station. The results are imprecise, but the difference is on the margin of statistical significance for the model in levels of crime registration per capita (p = 0.1). This pattern again confirms the interpretation that our political pressure results do not represent a true decrease in crime in governing party areas.

The results from the rollout of the new database overall, as well as the interaction effects, help to further confirm the interpretation that our effects are driven by politicians and senior police officials attempting to hold police accountable for performance, rather than colluding with police to keep crime figures low. As discussed in Section 2.4, the administrative data are not publicly released at any level lower than the district and the data from the new database are not released to the public at all. If senior officials were colluding to keep figures low, we would not see an increase in crime registration when the database increases observability by senior officials.

Figure 7: Interaction of Political Pressure and Crime Report Tracking Database Coverage



Notes: Coefficient estimates and 90% confidence intervals are shown for Equation 7. Unit of observation is crime report category j by police station k by month t for the sub-sample of crime reports in categories eligible for registration (N = 22,576). Data has been aggregated to categories shown as Level 3 in Table 2. The dependent variables are IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month and number of registered crimes per 1000 population per crime report category per police station catchment area per month. Estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment area.

5 Incidence

We have shown that political accountability pressure leads to under-registration of crime, and that the crime report tracking database mitigates this problem. We now explore the incidence of effects. We construct a wealth index at the neighborhood level following the inverse covariance weighting approach proposed in Anderson (2008) using data from a 2010 survey of 18,000 households sampled from all two hundred neighborhoods (union councils) of Lahore.[8] We divide police station catchment areas into terciles of this index, and repeat our two main analyses by terciles. Panel A of Figure 8 shows coefficient estimates for interaction terms between wealth tercile dummies and governing party vote share on the number of registered crimes. Panel B of Figure 8 shows event study estimates for coverage of the electronic database on number of registered crime disaggregated by neighbourhood terciles. While results vary somewhat between different constructions of the dependent variable, the results suggest stronger effects in poor neighborhoods. The effect of political accountability pressure is larger for poorer neighborhoods; in the model with levels of registered crime per capita, we can reject equality of effects (p ¡ 0.00). The effect of the new database rollout on crime registration is also significantly larger in the poorest neighborhoods. This could be the case because even in the absence of effective accountability pressure from above, wealthier households have greater ability to navigate the bureaucracy successfully and directly pressure police to handle their reports.

6 Discussion

In this paper we use rich administrative microdata on crime in Pakistan to investigate how police respond to changes in political accountability pressure, using variation in political alignment of neighborhoods to the governing party over time. We find that police are robustly less likely to register citizen crime reports in areas that supported the governing party. Since registration of a report as an official crime is the first step required to move forward to investigation and arrest, a reduction in crime registrations has potential for significant consequences for justice system outcomes.

We rule out the possibility that the decrease in registered crime is explained by increased prevention effort and a true reduction in crime incidence: logged crime reports are less likely to be pursued, staffing levels and response times do not change, and property crime reports are "downgraded" to non-criminal loss reports. Instead, results are consistent with a mechanism of officials deciding not to register some crime reports, under pressure to keep crime statistics low.

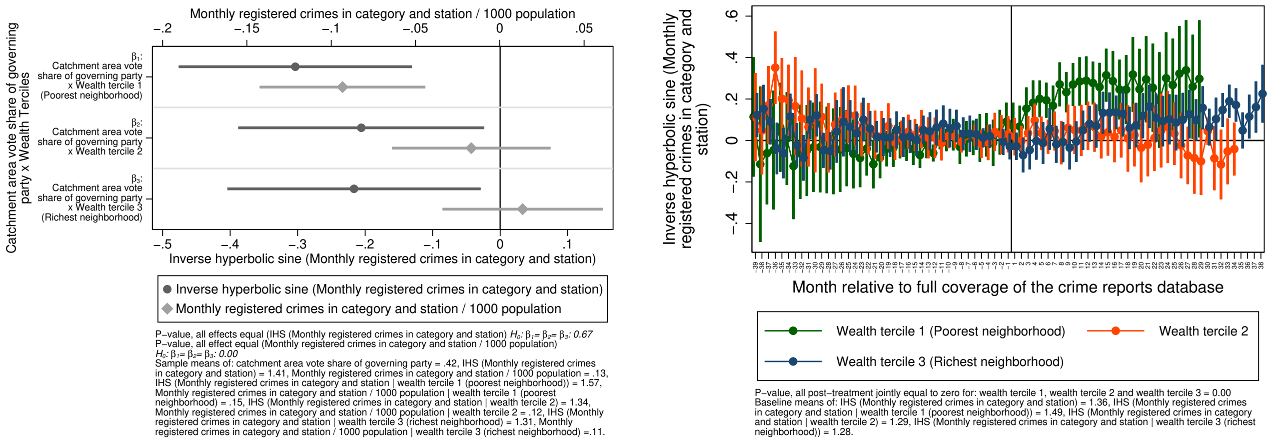
We use the rollout of a new crime reports tracking database to study how under-registration of crime changes as principals can better observe of police handling of citizen crime reports. We find that as observability

[8]The data source is the JICA Lahore Urban Transport Master Plan Household Integrated Survey, covering 18,000 households across Lahore District. The wealth index is generated using the Inverse Covariance Index (ICW Index) package developed by Bouguen and Varejkova (2020) following Anderson (2008), using a list of variables including indicators for whether the house is owned, rented, leased or other, the area of the plot, the total number of rooms within the house, availability and expenditure on utilities including electricity, gas, sewage, solid waste collection, telephone and internet, and monthly expenditure on transportation in household; ownership of air conditioner, washing machine, refrigerator, television, radio, computer, and mobile and landline phones; and indicators for self-reported bands of monthly household income.

Figure 8: Heterogeneity by Neighborhood Wealth

(a) Effect of political pressure

(b) Coverage of the crime reports database



Notes: Coefficient estimates and 90% confidence intervals are shown. Unit of observation is crime report category j by police station k by month t for the sub-sample of crime categories eligible for registration (N = 22,576 (8a), 84,660 (8b)). Data has been aggregated by Level 3 in Table 2. The dependent variables are IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month and number of registered crimes per 1000 population per crime report category per police station catchment area per month in panel (a) and only the former in panel (b) whereas the latter is shown in appendix figure A11. In panel (a), coefficient estimates are for the interaction terms between catchment area vote share (percentage) of the governing party (in 2013 and 2018) and a set of indicator variables equal to 1 for the respective neighborhood wealth tercile; *Wealth tercile 1 (poorest neighborhood)*, *Wealth tercile 2* and *Wealth tercile 3 (richest neighborhood)*. In panel 8b coefficient estimates are for I_{t+s}, which denote the number of months before/after coverage of the electronic system is complete (i.e. 95% or more of registered crimes appear in the electronic system of a station k) disaggregated by neighborhood wealth terciles (green denotes *Wealth Tercile 1 (Poorest neighborhood)*, orange denotes *Wealth Tercile 2* while blue denotes *Wealth Tercile 3 (Richest neighborhood)*). All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). A version of Panel (b) estimated with crimes per 1,000 population as the dependent variable is in Figure A11. Standard errors are clustered by police station catchment area.

improves, the number of registered crimes increases dramatically. This suggests the potential for such scalable technologies to mitigate the unintended consequences of accountability pressure, improve public services and ultimately enhance state capacity.

References

Acemoglu, D., Fergusson, L., Robinson, J., Romero, D. and Vargas, J. F. (2020). The perils of high-powered incentives: Evidence from Colombia's false positives. *American Economic Journal: Economic Policy*, **12** (3), 1–43.

Akhtari, M., Moreira, D. and Trucco, L. (2022). Political Turnover, Bureaucratic Turnover, and the Quality of Public Services. *American Economic Review*.

Amaral, S. and Bhalotra, S. (2021). Gender , Crime and Punishment : Evidence from Women Police Stations in India.

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, **103** (484), 1481–1495.

Asher, S. and Novosad, P. (2017). Politics and local economic growth: Evidence from India. *American Economic Journal: Applied Economics*, **9** (1), 229–273.

Ba, B. A. (2020). Going the Extra Mile: the Cost of Complaint Filing, Accountability, and Law Enforcement Outcomes in Chicago.

Banerjee, A., Chattopadhyay, R., Duflo, E., Keniston, D. and Singh, N. (2021). Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy, and Training. *American Economic Journal: Economic Policy*, **13** (1), 36–66.

Battistin, E., De Nadai, M. and Vuri, D. (2017). Counting rotten apples: Student achievement and score manipulation in Italian elementary Schools. *Journal of Econometrics*, **200** (2), 344–362.

Bertrand, M., Duflo, E. and Mullainathan, S. (2004). How much should we trust difference-in-differences estimates? *Quarterly Journal of Economics*, **119** (1).

Besley, T. J., Burgess, R., Khan, A. and Xu, G. (2021). Bureaucracy and development. *NBER Working Paper 29163*.

Bjorkman, M. and Svensson, J. (2009). Power to the people: evidence from a randomized field experiment on community-based monitoring in Uganda. *Quarterly Journal of Economics*, (May).

Borcan, O., Lindahl, M. and Mitrut, A. (2017). Fighting corruption in education: What works and who benefits? *American Economic Journal: Economic Policy*, **9** (1), 180–209.

Borusyak, K. (2021a). DID_IMPUTATION: Stata module to perform treatment effect estimation and pre-trend testing in event studies. Statistical Software Components, Boston College Department of Economics.

— (2021b). EVENT_PLOT: Stata module to plot the staggered-adoption diff-in-diff ("event study") estimates.

—, Jaravel, X. and Spiess, J. (2021). Revisiting Event Study Designs : Robust and Efficient Estimation. pp. 1–54.

Bouguen, A. and Varejkova, T. (2020). ICW_INDEX: Stata module to aggregate the variables included in the varlist into an index. Statistical Software Components, Boston College Department of Economics.

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A. and Padró i Miquel, G. (2015). The Value of Democracy: Evidence from Road Building In Kenya. *American Economic Review*, **105** (6), 1817–1851.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, **225** (2), 200–230.

Callen, M., Gulzar, S., Hasanain, A., Khan, M. Y. and Rezaee, A. (2020a). Data and policy decisions: Experimental evidence from Pakistan. *Journal of Development Economics*, **146**.

—, — and Rezaee, A. (2020b). Can political alignment be costly? *Journal of Politics*, **82** (2), 612–626.

Calvo, E., Cui, R. and Serpa, J. C. (2019). Oversight and efficiency in public projects: A regression discontinuity analysis. *Management Science*, **65** (12), 5651–5675.

Cheema, A., Hameed, Z. and Shapiro, J. N. (2017). Victimization, Citizen Engagement, and Policing in Lahore. *Institute of Development and Economic Alternatives Working Paper*, **1**, 1–65.

Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, **92** (1), 1–45.

Cook, S. and Fortunato, D. (2022). The Politics of Police Data: State Legislative Capacity and the Transparency of State and Substate Agencies. *American Political Science Review*, pp. 1–16.

Cox, G. W. and Mccubbins, M. D. (1986). Electoral Politics as a Redistributive Game. *Journal of Politics*, **48** (2), 370–389.

Cullen, J. B. and Reback, R. (2006). Tinkering Toward Accolades: School Gaming under a Performance Accountability System. *NBER Working Paper 12286*.

Dal Bó, E., Finan, F., Li, N. Y. and Schechter, L. (2021). Information Technology and Government Decentralization: Experimental Evidence From Paraguay. *Econometrica*, **89** (2), 677–701.

Das, S. and Sabharwal, G. (2017). Whom are you doing a favor to? Political Alignment and Allocation of Public Servants.

De Philippis, M. (2015). Multitask Agents and Incentives : The Case of Teaching and Research for University Professors. *CEP Discussion Paper No 1386*, (1386).

Dee, T. S., Dobbie, W., Jacob, B. A. and Rockoff, J. (2019). The causes and consequences of test score manipulation: Evidence from the New York regents examinations. *American Economic Journal: Applied Economics*, **11** (3), 382–423.

Dhaliwal, I. and Hanna, R. (2017). The devil is in the details: The successes and limitations of bureaucratic reform in India. *Journal of Development Economics*, **124**, 1–21.

Diamond, R. and Persson, P. (2016). The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests. *National Bureau of Economic Research Working Paper Series*, **No. 22207**.

Dranove, D., Kessler, D., McClellan, M. and Satterthwaite, M. (2003). Is more information better? The effects of "Report Cards" on health care providers. *Journal of Political Economy*, **111** (3), 555–588.

Duflo, E., Hanna, R. and Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review*, **102** (4), 1241–1278.

Dumont, E., Fortin, B., Jacquemet, N. and Shearer, B. (2008). Physicians' multitasking and incentives: empirical evidence from a natural experiment. *Journal of Health Economics*.

Dunya News (2021). IG Punjab Rao Sardar Ali Khan First Press Conference.

Eckhouse, L. (2022). Metrics Management and Bureaucratic Accountability: Evidence from Policing. *American Journal of Political Science*, **66** (2), 385–401.

Election Commission of Pakistan (a). GENERAL ELECTIONS 2013 RESULT.

Election Commission of Pakistan (b). GENERAL ELECTIONS 2018 RESULT.

Eterno, J. A. and Silverman, E. B. (2012). *The Crime Numbers Game: Management by Manipulation*. Taylor & Francis Group.

—, Verma, A. and Silverman, E. B. (2016). Police manipulations of crime reporting: Insiders' revelations. *Justice Quarterly*, **33** (5), 811–835.

Express Tribune (2021). Crime against women, children intolerable.

Feng Lu, S. (2012). Multitasking, Information Disclosure, and Product Quality: Evidence from Nursing Homes. *Journal of Economics and Management Strategy*, **21** (3), 673–705.

Figlio, D. N. (2006). Testing, crime and punishment. *Journal of Public Economics*, **90** (4-5), 837–851.

— and Getzler, L. S. (2006). Accountability, Ability and Disability: Gaming the System? *NBER Working Paper 9307*.

Finan, F., Olken, B. A. and Pande, R. (2015). The Personnel Economics of the State. *Handbook of Field Experiments*, (December).

Fisman, R. and Wang, Y. (2017). The distortionary effects of incentives in government: Evidence from China's "death ceiling" program. *American Economic Journal: Applied Economics*, **9** (2), 202–218.

Frank, R. G., Glazer, J. and McGuire, T. G. (2000). Measuring adverse selection in managed health care. *Journal of Health Economics*, **19** (6), 829–854.

Fryer, R. G. and Holden, R. T. (2013). Multitasking, Dynamic Complementarities, and Incentives: A Cautionary Tale.

Ghanem, D., Shen, S. and Zhang, J. (2020). A censored maximum likelihood approach to quantifying manipulation in china's air pollution data. *Journal of the Association of Environmental and Resource Economists*, **7** (5), 965–1003.

Glewwe, P., Ilias, N. and Kremer, M. (2010). Teacher incentives. *American Economic Journal: Applied Economics*, **2** (July), 481–488.

Goodman-bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, **225** (March).

Gulzar, S. and Pasquale, B. J. (2017). Politicians, bureaucrats, and development: Evidence from India. *American Political Science Review*, **111** (1), 162–183.

Gupta, A. (2021). Impacts of performance pay for hospitals: The readmissions reduction program. *American Economic Review*, **11** (4), 1241–1283.

Hodler, R. and Raschky, P. A. (2014). Regional Favoritism. *Quarterly Journal of Economics*, pp. 995–1033.

Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, (1970), 74–92.

— and Milgrom, P. (1991). Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *The Journal of Law, Economics, and Organization*, **7** (special_issue), 24–52.

Hong, F., Hossain, T., List, J. A. and Tanaka, M. (2018). Testing the Theory of Multitasking: Evidence From a Natural Field Experiment in Chinese Factories. *International Economic Review*, **59** (2), 511–536.

Hsieh, B. C.-t., Miguel, E., Ortega, D. and Rodriguez, F. (2011). The Price of Political Opposition: Evidence from Venezuela's Maisanta. *American Economic Journal: Applied Economics*, **3** (April), 196–214.

Iyer, L. and Mani, A. (2012). Traveling agents: political change and bureaucratic turnover in India. *Review of Economics and Statistics*, **94**, 723–739.

—, —, Mishra, P. and Topalova, P. (2012). The Power of Political Voice: Women's Political Representation and Crime in India. *American Economic Journal: Applied*, **4** (4), 165–193.

Jacob, B. and Levitt, S. (2003). Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics*, (August), 843–877.

Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, **89** (5-6), 761–796.

Johnson, R. M., Reiley, D. H. and Muñoz, J. C. (2015). "The war for the fare": How driver compensation affects bus system performance. *Economic Inquiry*, **53** (3), 1401–1419.

Kavanaugh, G., Sviatschi, M. and Trako, I. (2019). Women Officers, Gender Violence and Human Capital: Evidence from Women's Justice Centers in Peru. p. 88.

Kelley, E. M., Lane, G. and Schonholzer, D. (2020). Monitoring in Target Contracts : Theory and Experiment in Kenyan Public Transit.

Khan, A. Q., Khwaja, A. I. and Olken, B. A. (2019). Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings. *American Economic Review*, **109** (1), 237–270.

Lakhtakia, S. (2021). Law Makers or Law Breakers ? The Impact of Criminal Politicians on Local Crime. (May).

Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, **99** (5), 1979–2011.

Lehne, J., Shapiro, J. N. and Vanden Eynde, O. (2018). Building connections: Political corruption and road construction in India. *Journal of Development Economics*, **131** (November 2016), 62–78.

Lindbeck, A. and Weibull, J. W. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice*, **52** (3), 273–297.

Malik, N. and Qureshi, T. A. (2020). A Study of Economic, Cultural, and Political Causes of Police Corruption in Pakistan. *Policing: A Journal of Policy and Practice*.

Martinelli, C., Parker, S. W., Pérez-Gea, A. C. and Rodrigo, R. (2018). Cheating and incentives: Learning from a policy experiment. *American Economic Journal: Economic Policy*, **10** (1), 298–325.

Miller, A. R. and Segal, C. (2016). Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence Escalation. *Review of Economic Studies*.

Miller, G. and Babiarz, K. S. (2014). Pay-for-Performance Incentives in Low- and Middle-Income Country Health Programs. *Encyclopedia of Health Economics*, pp. 457–466.

Neal, D. and Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics*, **92** (2), 263–283.

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, **115** (2), 200–249.

Oxman, A. D. and Fretheim, A. (2009). Can paying for results help toachieve the Millennium Development Goals? A critical review of selected evaluations of results-based financing. *Journal of Evidence-Based Medicine*, **2** (3), 184–195.

Pakistan Bureau of Statistics (2017). *Population & Housing Census*. Tech. rep., Pakistan Bureau of Statistics, Islamabad, Pakistan.

Punjab Information Technology Board (2017). Digital Punjab: Enhancing Public Services Through Technology.

Reinikka, R. and Svensson, J. (2011). The power of information in public services: Evidence from education in Uganda. *Journal of Public Economics*, **95** (7-8), 956–966.

Rios-Avila, F., Sant'Anna, P. and Callaway, B. (2021). CSDID: Stata module for the estimation of Difference-in-Difference models with multiple time periods.

Rivera, R. and Ba, B. A. (2019). The Effect of Police Oversight on Crime and Allegations of Misconduct: Evidence from Chicago. *SSRN Electronic Journal*.

Rizvi, J. (2017). Transforming Policing Through Technology. *MIT Technology Review*.

Sanjay Patil (2008). *Feudal Forces: Reform Delayed*.

Siddiqi, M. U. A., Butt, K. M. and Afzaal, M. (2014). Politicized Policing in Pakistan: A Constructivist Study of Problems of Policing in Lahore. *Journal of Political Science*, **XXXII** (December 2014), 3–25.

Suárez Serrato, J. C., Wang, X. Y. and Zhang, S. (2019). The limits of meritocracy: Screening bureaucrats under imperfect verifiability. *Journal of Development Economics*, **140** (May), 223–241.

Wade, R. (1985). The market for public office: Why the Indian state is not better at development. *World Development*, **13** (4), 467–497.

Werner, R. M. and Asch, D. a. (2005). The Unintended Consequences of Publicly Reporting Quality Information. *JAMA : the journal of the American Medical Association*, **293** (10), 1239–1244.

Xu, G. and Burgess, M. B. R. (2018). Social proximity and bureaucrat performance: Evidence from India. *NBER Working Paper 25389*.

Yu, O. and Zhang, L. (1999). The under-recording of crime by police in China: a case study. *Policing: An International Journal of Police Strategies & Management*.

Zwart, F. d. (1994). *The Bureaucratic Merry-Go-Round: Manipulating the Transfer of Indian Civil Servants*. Amsterdam University Press.

Appendix A Supplementary material

A.1 Additional Tables

Table A1: Definition of terms

Term	Definition
Report	Occurs when a citizen reports a crime or loss to the police. Before the rollout of the new crime reports database, these were recorded only in hard copy registers at the front desk of police stations, which are not centralized or digitized. After the rollout of the new crime reports database, these are recorded in the database (see Table 1).
Registration of crime (as a First Information Report or FIR)	Occurs when police process an official record of a crime. The process for an official crime registration stays the same throughout the period we study (see Table 1). The numbers of registered crimes by category are monitored by senior officials throughout the period we study.
Loss report	A specific type of report, which is classified by police as a non-criminal loss (such as a lost wallet). These are not eligible to be registered as a crime with an FIR.
Crime reports database (CMS or "complaints management system")	New database rolled out over the period we study, which is used to track and allow senior officials to immediately access details about any report, its progress and whether it was registered.

Table A2: Effect of Political Pressure on Likelihood of Crime Registration: Robustness to Sub-sample of Stations with Early Coverage of the Crime Database

	Probability of crime report registration	
	(1)	(2)
	Overall	Stations with early coverage of the electronic system
Catchment area	-0.139**	-0.133*
vote share	(0.064)	(0.071)
governing party	{0.037}	{0.042}
Observations	365818	280558
Sample Mean	0.335	0.343
Obs. unit	crime-report-level	crime-report-level
Category FEs	yes	yes
Station FEs	yes	yes
Month FEs	yes	yes
Clustering	station	station

Notes: *, **, *** denote significance at 10%, 5% and 1% respectively. Unit of observation is individual crime report for the sub-sample of crime categories eligible for registration in both columns with the difference that in column the overall sample has been used for estimation; in column 2 we restrict the sample to police stations where the new database had rolled out by January 2018. The dependent variable is the probability of a crime report being registered in both columns. All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are in parenthesis and clustered by police station catchment areas. Spatially clustered standard errors are given in braces. Data has been aggregated by Level 3 in Table 2 for columns 1-3.

Table A3: Robustness Check: Effect of Political Pressure - Excluding three month window around election

	(1)	(2)	(3)	(4)	(5)
	IHS (Monthly registered crimes in category and station)	Monthly registered crimes in category and station / 1000 population	Probability of crime report registration	Response time (Hours)	Number of officers in station and month
Catchment area	-0.308**	-0.064**	-0.189**	3.631	1.438
vote share	(0.124)	(0.030)	(0.088)	(7.863)	(2.445)
governing party	{0.086}	{0.024}	{0.054}	{4.571}	{1.747}
Observations	14110	14110	203360	388905	747
Sample Mean	1.405	0.133	0.356	38.700	14.213
Obs. unit	cat-station-month	cat-station-month	crime-report-level	crime-report-level	station-month
Category FEs	yes	yes	yes	yes	
Station FEs	yes	yes	yes	yes	yes
Month FEs	yes	yes	yes	yes	yes

Notes: *, **, *** denote significance at 10%, 5% and 1% respectively. The estimation sample excludes the three month window around the election. Unit of observation is crime report category j by police station k by month t in columns 1 and 2, individual crime report for the sub-sample of crime categories eligible for registration in column 3, individual crime/loss report for the overall sample including categories not eligible for registration in column 4, and number of officers active in station k by month t in column 5. The dependent variable is IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month in column 1, number of registered crimes per 1000 population per report category per police station catchment area per month in column 2, the probability of a crime report being registered in column 3, time taken (in hours) by police to first respond to a crime report in column 4 and number of officers active in station and month in column 5. All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2) except for those in column 5 where fixed effects are for station and month. Standard errors are in parenthesis and clustered by police station catchment areas. Spatially clustered standard errors are given in braces. Data has been aggregated by Level 3 in Table 2 for columns 1 and 2.

Table A4: Effect of Political Pressure on Registered Crime: Alternative Levels of Data Aggregation

	(1) IHS (Monthly registered crimes in category and station)	(2) Monthly registered crimes in category and station / 1000 population	(3) IHS (Monthly registered crimes in category and station)	(4) IHS (Monthly registered crimes in category and station)
Catchment area	-0.254***	-0.047*	-0.318***	-0.562***
vote share	(0.087)	(0.024)	(0.087)	(0.140)
governing party	{0.059}	{0.016}	{0.060}	{0.097}
Observations	22576	22576	2822	1328
Sample Mean	1.406	0.130	1.505	5.142
Obs. unit	cat-station-month	cat-station-month	cat-station-prepost	station-month
Category FEs	yes	yes	yes	
Station FEs	yes	yes	yes	yes
Month FEs	yes	yes		yes
Pre-post period FEs			yes	

Notes: *, **, *** denote significance at 10%, 5% and 1% respectively. Unit of observation is crime report category j by police station k by month t in columns 1 and 2, crime report category j by police station k by pre/post 2018 elections period in column 3 and police station k by month t in column 4. The dependent variable is IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month in all columns except for in column 2 where it is number of registered crimes per 1000 population per crime report category per police station catchment area per month. All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2) except for in column 3 where instead of month, pre and post election period fixed effects have been included and column 4 where fixed effects are only for station and month. Standard errors are in parenthesis and clustered by police station catchment areas. Spatially clustered standard errors are given in braces. Data has been aggregated by Level 3 in Table 2 for columns 1-3.

Table A5: Effect of Political Pressure on Registered Crime: Alternative definition of independent variable

	(1) IHS (Monthly registered crimes in category and station)	(2) Monthly registered crimes in category and station / 1000 population	(3) Probability of crime report registration	(4) Response time (Hours)	(5) Number of officers in station and month
(%) Catchment area	-0.044***	-0.009***	-0.021**	0.699	0.145
in constituencies with	(0.017)	(0.003)	(0.010)	(1.593)	(0.338)
governing party MPAs	{0.012}	{0.003}	{0.006}	{0.824}	{0.216}
Observations	22576	22576	365818	688487	1328
Sample Mean	1.406	0.130	0.335	43.409	14.306
Obs. unit	cat-station-month	cat-station-month	crime-report-level	crime-report-level	station-month
Category FEs	yes	yes	yes	yes	
Station FEs	yes	yes	yes	yes	yes
Month FEs	yes	yes	yes	yes	yes
Clustering	station	station	station	station	station

Notes: *, **, *** denote significance at 10%, 5% and 1% respectively. Unit of observation is crime report category j by police station k by month t in columns 1 and 2, individual crime report for the sub-sample of crime categories eligible for registration in column 3, individual crime/loss report for the overall sample including categories not eligible for registration in column 4, and station k by month t in column 5. The dependent variable is IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month in column 1, number of registered crimes per 1000 population per report category per police station catchment area per month in column 2, the probability of a crime report being registered in column 3, time taken (in hours) by police to first respond to a crime report in column 4 and number of officers active in station and month in column 5. Coefficient estimates are for the percentage of catchment area in constituencies with governing party MPAs between the 2013 and 2018 electoral period. All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are in parenthesis and clustered by police station catchment areas. Spatially clustered standard errors are given in braces. Data has been aggregated by Level 3 in Table 2 for column 1.

Table A6: Balance between stations with early and late coverage of the crime reports database

Variable	(1) 0 = late implementer Mean/SE	(2) 1 = early implementer Mean/SE	T-test Difference (1)-(2)
Wealth index	0.003	0.003	-0.001
	(0.002)	(0.001)	
Years at location	27.307	26.238	1.070
	(0.979)	(1.147)	
Population in 2009 (1000s)	47.466	47.179	0.288
	(0.995)	(1.272)	
Population density in 2009 (1000 people per sq km)	17.842	22.671	-4.829
	(3.302)	(3.572)	
Proportion high school+	0.413	0.403	0.009
	(0.017)	(0.013)	
Proportion employed	0.627	0.638	-0.011
	(0.008)	(0.005)	
Household income (1000 PKR per month)	24.336	23.376	0.959
	(1.011)	(1.038)	
Area (sq km)	10.100	14.777	-4.678
	(3.158)	(3.086)	
PML-N vote share 2013	0.503	0.525	-0.022*
	(0.008)	(0.008)	
PTI vote share 2013	0.360	0.325	0.034
	(0.019)	(0.014)	
IHS (Number of FIRs Registered, Baseline)	1.357	1.242	0.115
	(0.057)	(0.053)	
N	32	51	

Notes: The value displayed for t-tests are the differences in the means across the groups. Unit of observation is the police station catchment area. Robust standard errors shown. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level. Principal component index of household assets.[1] Average number of years individuals have lived at a location.[2] Proportion high school+ includes total number of graduate, undergraduate, college and high school students.[3] The vote share for PML-N in the 2013 General Elections.[4] The vote share for PTI in the 2013 General Elections.[5] Inverse hyperbolic sine transformation of number of registered crimes in a station at baseline.[6]

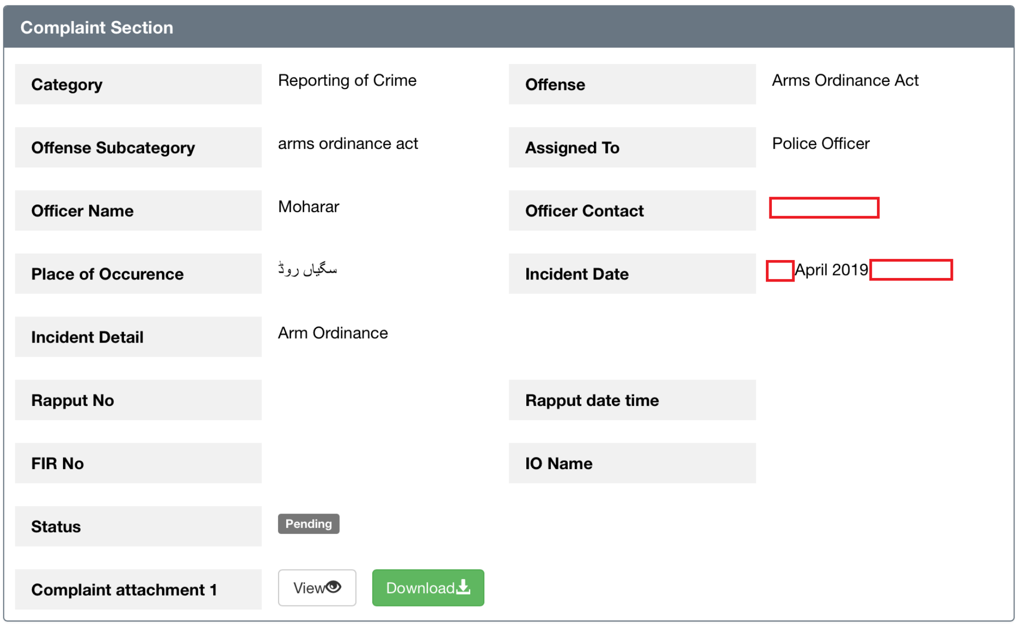
A.2 Additional Figures

Figure A1: The Chief Minister of Punjab, chairing a meeting at police headquarters in Lahore to get an update on the law and order situation. Source: Daily Times, November 8, 2018

Figure A2: Police Station Assistant enters the crime report into new database. Photo source: Government of Punjab

Figure A3: Screenshot of dashboard from new CMS database

(a) Database dashboard

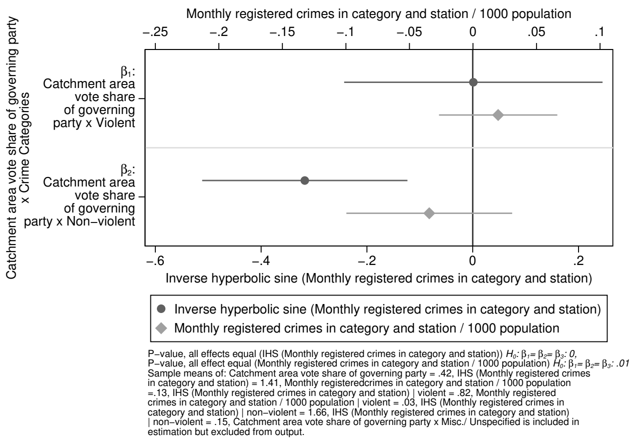


(b) View of individual report status with downloadable detailed report

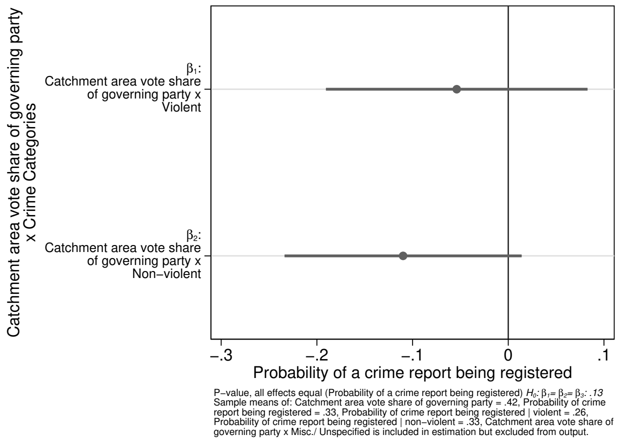
Source: Government of Punjab

Figure A4: Effect of Political Pressure by Crime Type

(a) IHS (Monthly registered crimes in category and station) & monthly registered crimes in category and station / 1000 population)

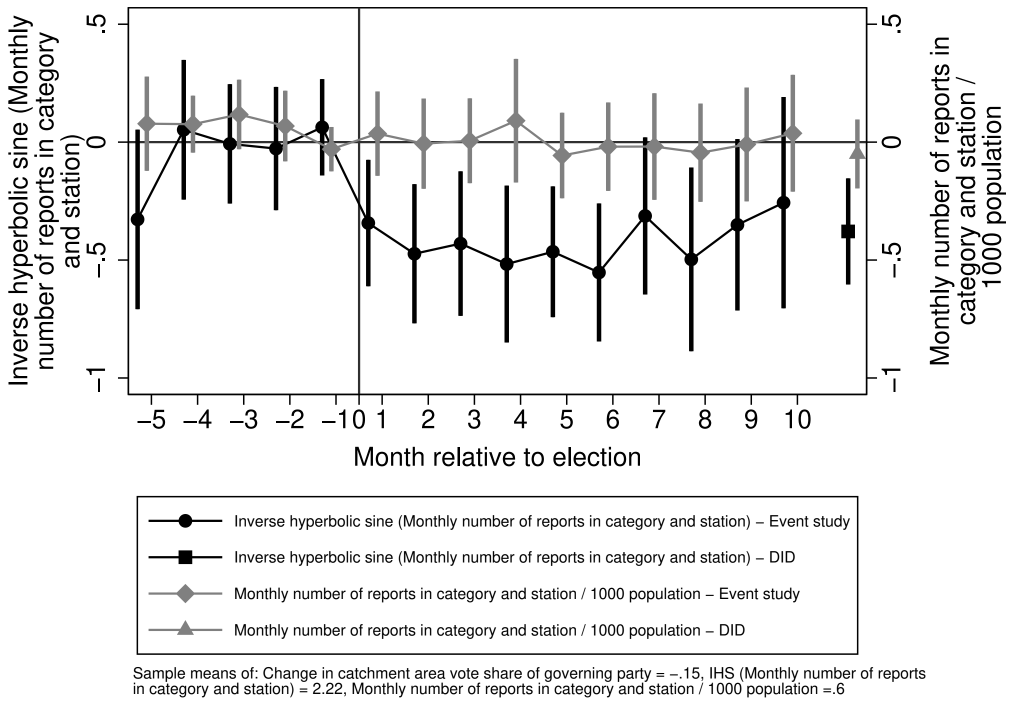


(b) Probability of a crime report being registered



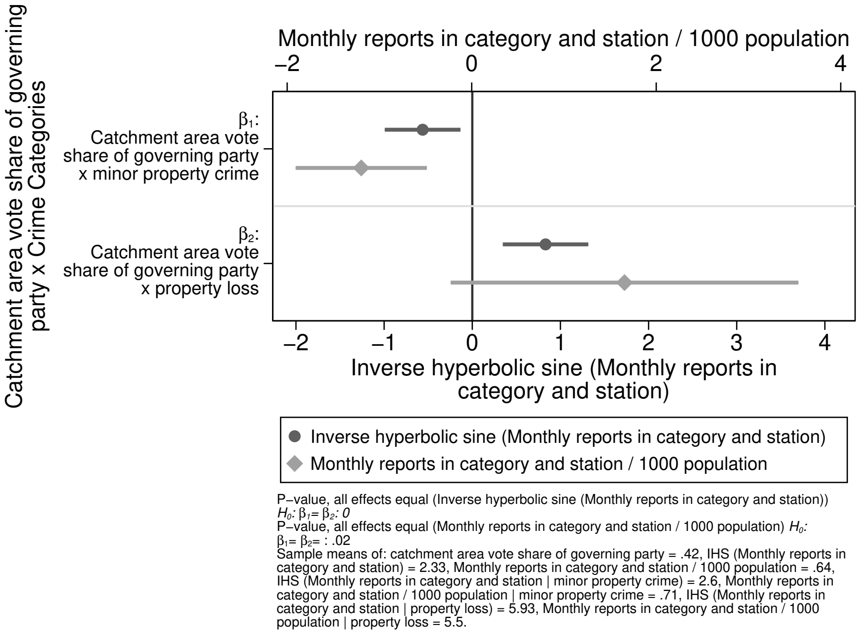
Notes: Coefficient estimates and 90% confidence intervals are shown. Unit of observation is crime report category j by police station k by month t in panel A4a (N = 22,576), and individual crime report in panel A4b (N = 365,818), for the sub-sample of crime categories eligible for registration. The dependent variables are IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month and number of registered crimes per 1000 population per crime report category per police station catchment area per month in panel A4a, and the probability of a crime report being registered in panel A4b. Coefficient estimates are for the interaction between catchment area vote share (percentage) of the governing party (in 2013 and 2018) and a set of indicator variables equal to 1 for the respective crime report category j; *violent* and *non-violent*. All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment areas. In panel A4a data has been aggregated by Level 3 in Table 2.

Figure A5: Effect of Political Pressure on Crimes Reported



Notes: Coefficient estimates and 90% confidence intervals are shown. Unit of observation is crime report category j by police station k by month t (N = 23,904). The dependent variables are IHS (inverse hyperbolic sine) of the number of crime reports per crime report category per police station catchment area per month and number of registered crimes per 1000 population per crime report category per police station catchment area per month. Coefficient estimates are for interaction terms between catchment area vote share (percentage) of the governing party (in 2013 and 2018) and a set of indicator variables, I_t, equal to 1 for the respective month t under study. All estimates include fixed effects for crime report category, police station catchment area and month (as shown by Level 1 in Table 2). The month prior to the elections is the reference period. Standard errors are clustered by police station catchment area. Data has been aggregated by Level 3 in Table 2.

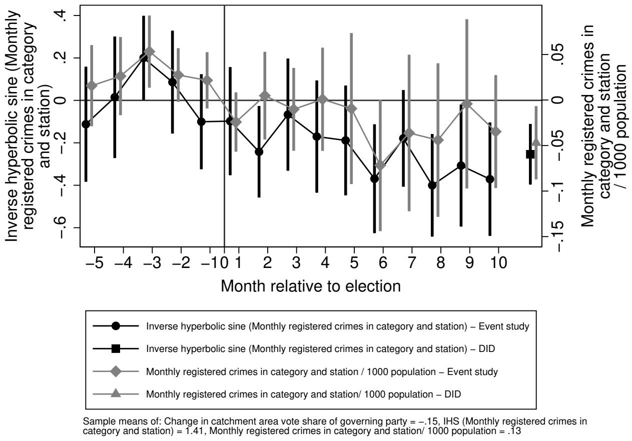
Figure A6: Robustness Check: Effect of political pressure on “downgrading” property crime reports into loss reports - excluding three month window around election



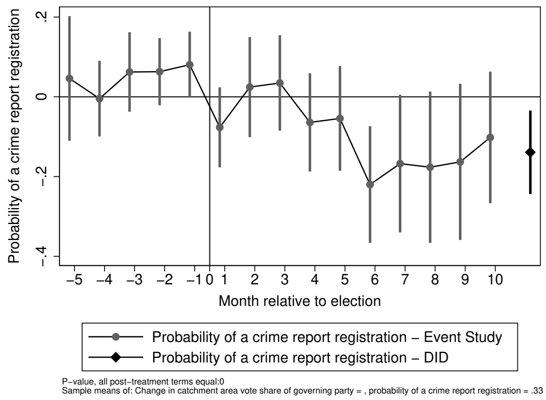
Notes: Coefficient estimates and 90% confidence intervals are shown. The estimation sample excludes the three month window around election. Unit of observation is crime report category j by police station k by month t (N = 3,237). The dependent variables are IHS (inverse hyperbolic sine) of the number of reports per report category per police station catchment area per month and number of reports per 1000 population per report category per police station catchment area per month. Sample includes categories aggregated to categories shown as Level 3 in Table 2; estimation sample includes only data for categories within *property loss* and *minor property crime*. All estimates include fixed effects for station, month and broader crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment area.

Figure A7: Event Studies: Effect of Political Pressure

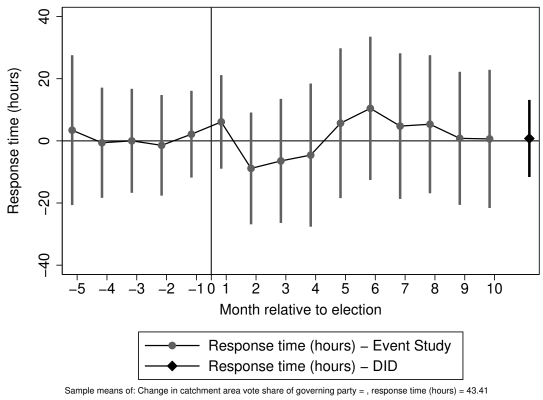
(a) IHS (Monthly registered crimes in category and station) & Monthly registered crimes in category and station/ 1000 population



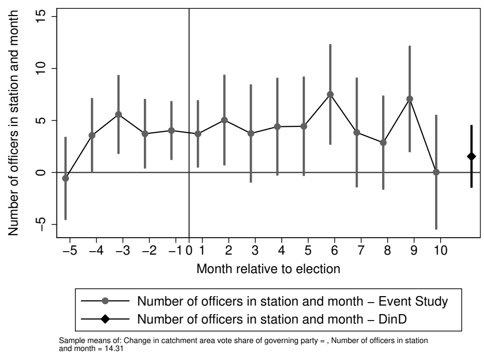
(b) Probability of crime report registration



(c) Response time

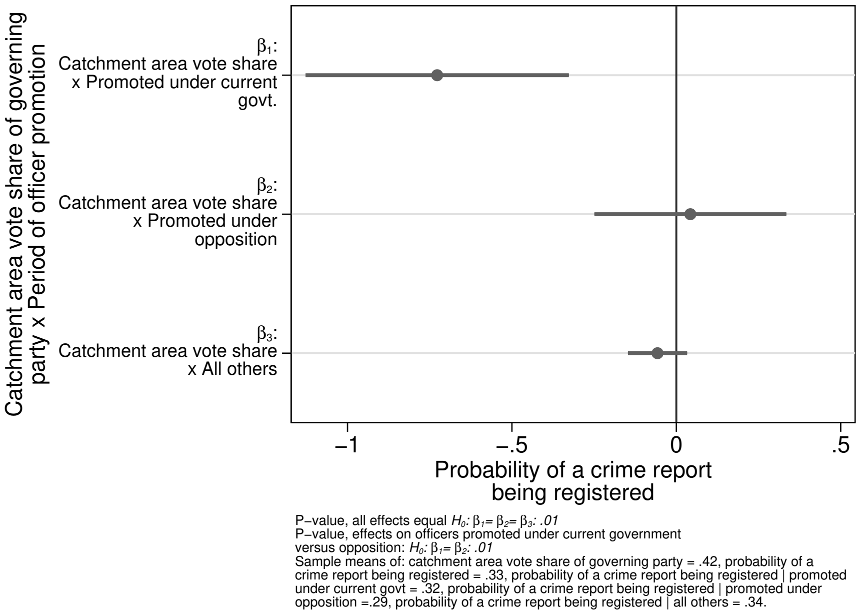


(d) Staff inputs



Notes: Coefficient estimates and 90% confidence intervals are shown for the event study versions of Equations 1 and 3. The unit of observation is crime report category j by police station k by month t for the sub-sample of crime reports in categories eligible for registration (N = 22,576) in panel A7a, individual crime report for the sub-sample of crime categories eligible for registration (N = 365,818) in panel A7b, individual crime report for the overall sample including categories not eligible for registration (N = 688,487) in panel A7c, and number of officers active in station k by month t (N = 1,328) in panel A7d. The dependent variables are IHS (inverse hyperbolic sine) of the number of registered crimes per crime report category per police station catchment area per month and number of registered crimes per 1000 population per crime report category per police station catchment area per month in panel A7a, the probability of a crime report being registered in panel A7b, time taken (in hours) by police to first respond to a crime report in panel A7c and number of officers active in station and month in panel A7d. Coefficient estimates are for the interaction terms between catchment area vote share of the governing party (in 2013 and 2018) and I_t, a set of indicator variables equal to 1 for the respective month t under study. All estimates include fixed effects for crime report category (as shown by level 1 in Table 2), police station catchment area and month except for in panel A7d where fixed effects are for station and month. The month prior to the elections is the reference period. Standard errors are clustered by police station catchment area. In panel A7a data has been aggregated by level 3 in Table 2.

Figure A8: Robustness Check: Heterogeneous Effects of Political Pressure on Crime Registration by Officer Promotion - including only promotions in a three month window before and after election

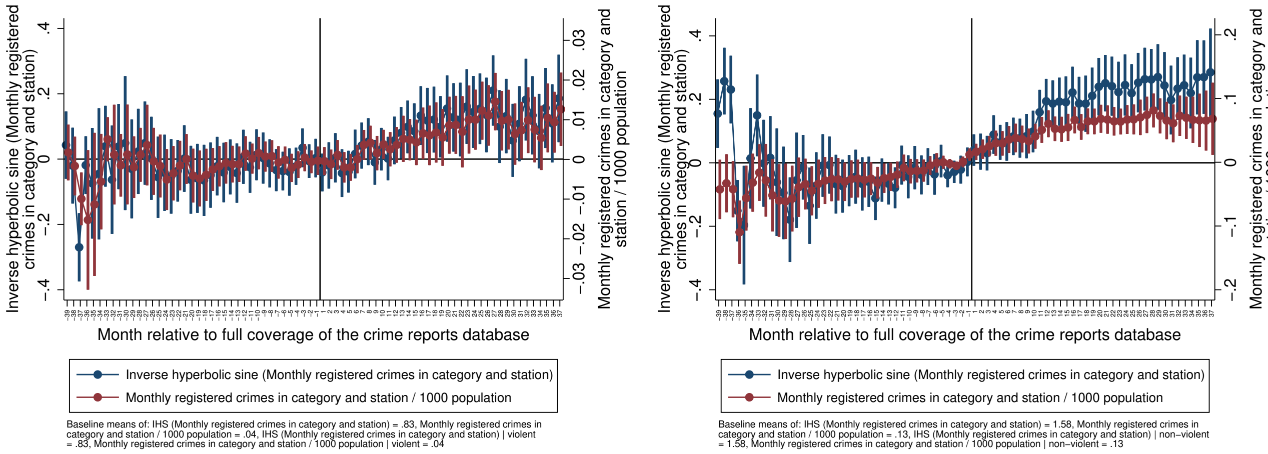


Notes: Figure shows coefficients from Equation 5. The estimation sample excludes the three months around election. Coefficient estimates and 90% confidence intervals are shown. Unit of observation is an individual crime report (N = 231,239). The dependent variable is the probability of a crime report being registered. Coefficient estimates are for the interaction terms between catchment area vote share of the governing party (in 2013 and 2018) and indicator variables equal to 1 for respective category of officers; those promoted by the first government in power pre-election considered as *promoted under current government*, those promoted after the election considered as *promoted under opposition* and *all others* including those who were never promoted. All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment area.

Figure A9: Effect of Electronic System Coverage on Number of Registered Crimes by Crime Type

(a) Violent

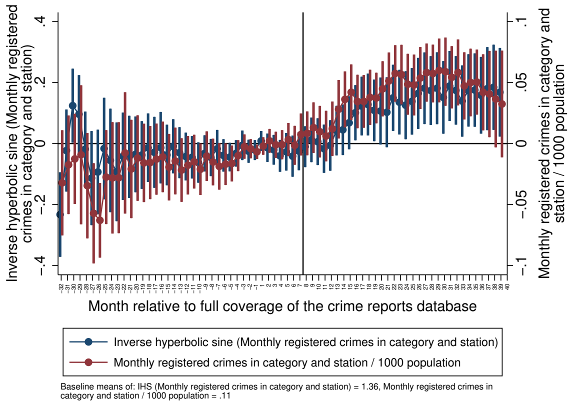
(b) Non-violent



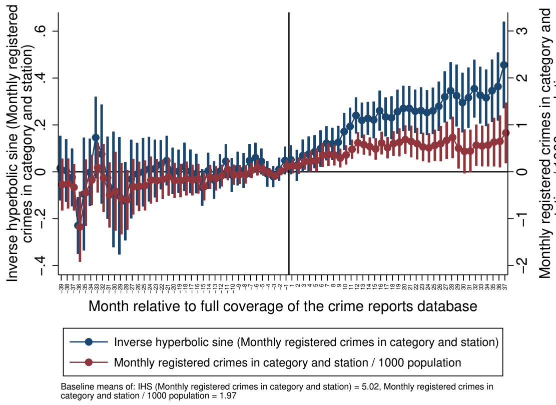
Notes: Coefficient estimates and 90% confidence intervals are shown. Unit of observation is crime report category j by police station k by month t for the sub-sample of crime reports (eligible for registration) in the *violent* category (N = 29,880) in panel A9a and in *nonviolent* category (includes property crime, N = 44,820) in panel A9b. The dependent variable are IHS (inverse hyperbolic sine) of the number of registered crimes per report category per police station catchment area per month and number of registered crimes per 1000 population per report category per police station catchment area per month. Coefficient estimates are for I_{t+s}, which denotes the number of months before/after coverage of the system is complete i.e. 95% or more of FIRs appear in the electronic system of a station k. Estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment areas. Data has been aggregated by Level 3 in Table 2.

Figure A10: Robustness Tests for the Effect of Electronic System Coverage

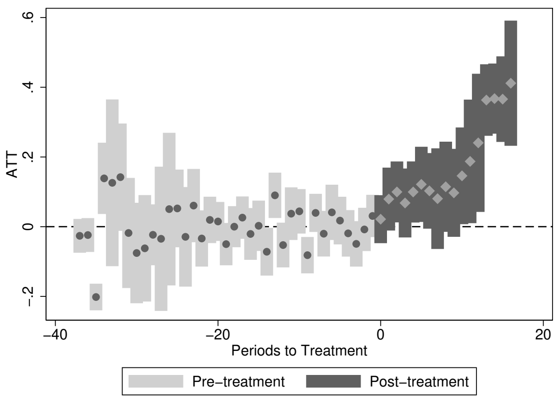
(a) Alternative threshold
(90% rather than 95%)



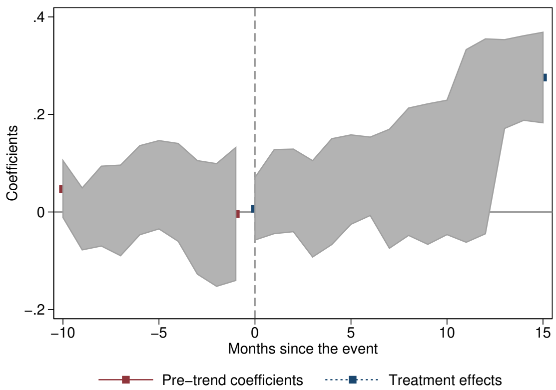
(b) Alternative Aggregation
(station-month level)



(c) Callaway and Santana Estimation

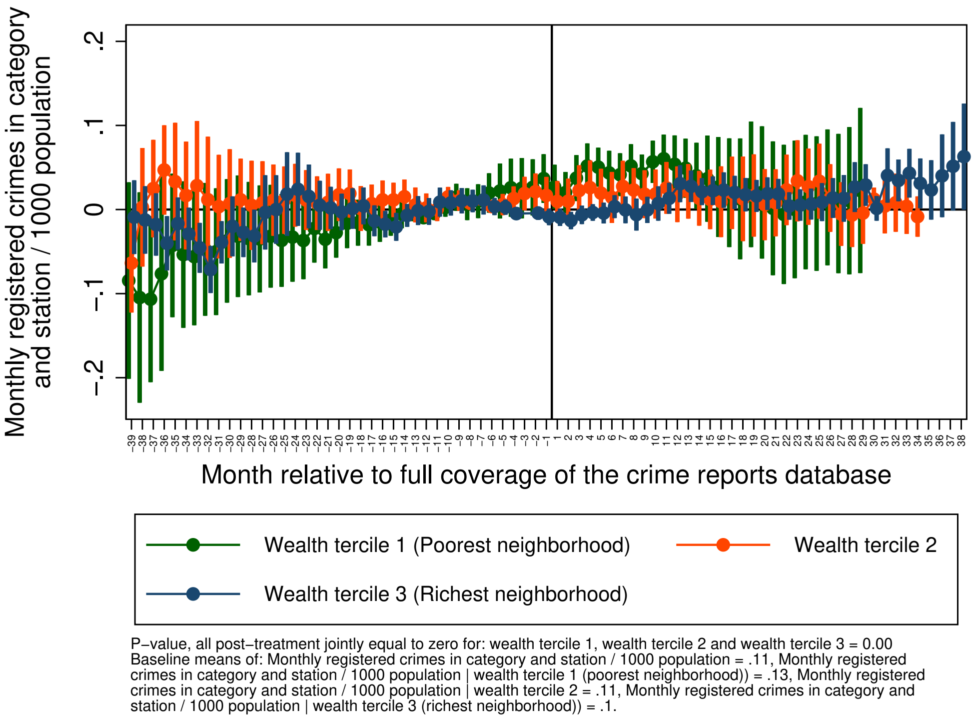


(d) Borusyak Estimation



Notes: Coefficient estimates and 90% confidence intervals are shown for Equation 6 in all panels. Unit of observation is crime report category j by police station k by month t for the sub-sample of crime reports in categories eligible for registration (N = 84,660) in panels A10a and only police station k by month t for the same sub-sample (N = 4,980) in panels A10b, A10c and A10d. The dependent variables are IHS (inverse hyperbolic sine) number of registered crimes per report category per police station catchment area per month and number of registered crimes per 1000 population per report category per police station catchment area per month in panels A10a and A10b whereas only the former in panels A10c and A10d. Coefficient estimates are for I_{t+s}, which denotes the number of months before/after coverage of the system is complete i.e. 95% or more of FIRs appear in the electronic system of a station k in all panels except for in panel A10a where an alternative 90% threshold is used i.e. the system is complete when 90% or more of registered crimes appear in the electronic system of a station k. Estimates in panel A10c have been estimated using the csdid package developed by Rios-Avila *et al.* (2021) implementing the estimator proposed by Callaway and Sant'Anna (2021). Estimates in panel A10d have been estimated using the eventplot and didimputation packages developed by (Borusyak, 2021b,a) implementing Borusyak *et al.* (2021). All estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2) except for A10b where only station and month fixed effects have been included. Standard errors are clustered by police station catchment areas. Data has been aggregated by Level 3 in Table 2.

Figure A11: Heterogeneity by Neighborhood Wealth - Registered Crimes Per 1000 Population



Notes: Coefficient estimates and 90% confidence intervals are shown. Unit of observation is crime report category j by police station k by month t for the sub-sample of crime categories eligible for registration (N = 84,660). This is an alternate version of fig 8b in which the dependent variable is number of registered crimes per 1000 population per crime report category per police station catchment area per month. Coefficient estimates are for I_{t+s}, which denote the number of months before/after coverage of the electronic system is complete (i.e. 95% or more of registered crimes appear in the electronic system of a station k) disaggregated by neighborhood wealth terciles (green denotes *Wealth Tercile 1 (Poorest neighborhood)*, orange denotes *Wealth Tercile 2* while blue denotes *Wealth Tercile 3 (Richest neighborhood)*). Estimates include fixed effects for station, month and crime report category (as shown by Level 1 in Table 2). Standard errors are clustered by police station catchment area. Data has been aggregated by Level 3 in Table 2.