

Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy, and Training[†]

By ABHIJIT BANERJEE, RAGHABENDRA CHATTOPADHYAY, ESTHER DUFLO,
DANIEL KENISTON, AND NINA SINGH*

Management matters for firms, but what practices are optimal in hierarchical government organizations? And can skilled managers identify them? A large-scale randomized trial conducted with the police of Rajasthan, India, tested four interventions recommended by senior police officers: limitations of transfers, rotation of duties and days off, increased community involvement, and on-duty training. Field experience motivated a fifth intervention: “decoy” visits by enumerators to register cases, incentivizing staff to improve service. Only training and decoy visits had robust impacts; others were poorly implemented and ineffective. Management reforms can improve policing, but even skilled leaders struggle to identify the optimal interventions. (JEL H76, J24, J45, K42, M53, O17)

The role of management has taken a central place in the study of organizations, both private firms as well as government bureaucracies. Following the seminal work by Bloom and Van Reenen (2007) (henceforth “BVR”), a literature has emerged arguing that the quality of management has an independent role in explaining firm productivity (management as technology). In a proof-of-concept experiment in India, Bloom et al. (2013) (henceforth “BEMMR”) demonstrate that by providing specific, customized advice, management consultants can improve practices and

*Banerjee: Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 (email: banerjee@mit.edu); Chattopadhyay: IIM Calcutta, Diamond Harbour Rd, Joka, Kolkata, West Bengal 700104, India (email: rchattopadhyay@gmail.com); Duflo: Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 (email: eduflo@mit.edu); Keniston: Louisiana State University, 501 South Quad Drive, Baton Rouge, LA 70808 (email: dkeniston@lsu.edu); Singh: Rajasthan Police, Lalkothi, Jaipur, Rajasthan 302015, India (email: ninasingh89@gmail.com). Matthew Shapiro was coeditor for this article. We thank Clement Imbert, Selvan Kumar, Dhruva Kothari, Neil Buddy Shah, and Pankaj Verma for their outstanding contributions as research assistants on this project. We are indebted to M.K. Devarajan and M.L. Lather for their insights and comments, and A.S. Gill, former Director General of Police for Rajasthan Police, for initiating and constantly supporting this research. The findings and opinions expressed in this paper are those of the authors and do not in any way engage the responsibility of the Rajasthan police. The paper was written in complete independence and the Rajasthan police, as an institution, did not weigh in on the analysis of the data or the drafting of the report. All the data collected in the course of this project is available for download and public use (including, but not limited to, replication of our results) at <https://doi.org/10.7910/DVN/OC35YC> (Banerjee et al. 2011). This study is registered in the AEA RCT Registry with the unique identifying number: AEARCTR-0001362 (Banerjee et al. 2016) and it has been approved by the MIT Committee on the Use of Humans and Experimental Subjects, protocol #1108004638.

We are grateful for funding from the Will and Flora Hewlett Foundation and the UNODC. The Hewlett Foundation was not involved in the drafting of this paper.

[†]Go to <https://doi.org/10.1257/pol.20190664> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

profits even in large and successful firms operating in a competitive environment. We investigate what kind of management reforms can work in a very different type of institution—the police department of the state of Rajasthan, India—by conducting a large-scale experiment, covering 162 police stations serving several million citizens.

This paper contributes to two open questions in the literature. This first is whether changes in management practices can have a causal effect, holding constant the quality of managers themselves, guidance from external consultants, and other dimensions of human capital. This question is particularly salient for the police, a rigidly hierarchical organization, and perhaps an unlikely setting for standard management principles to apply. Here we find some positive results, although, as we discuss in depth below, our conclusions are mixed. The second question, tightly linked to the first, is whether existing firm managers are able to map the broad principles of management best practices onto specific reforms that will increase organizational productivity. On this question, our findings are more negative. Although the reforms we studied were selected by highly experienced senior police officers, guided by eminent National Police Commission reports, several failed to show any impact and were poorly implemented by station-level officers.

The question is still open as to whether the broad principles of good management identified by Bloom and Van Reenen (2007) apply to public organizations or government bureaucracies. Bloom et al. (2010) and Bloom et al. (2015) find that in public hospitals and educational institutions, respectively, general management practices adapted to the setting are strongly associated with better outcomes. Another strand of research focusing specifically on the role of incentives in noncommercial settings (Di Tella and Schargrodsky 2003, Olken 2007, Dhaliwal and Hanna 2017) also finds positive effects on performance. A different perspective comes from Rasul and Rogger (2018), who inquire whether the BVR measures of management quality explain productivity differences across Nigerian government departments. Their results show that (middle) managerial autonomy, one of BVR's main measures of management quality, has a positive effect on performance. In contrast, the provision of incentives, which is associated with good performance in BVR, turns out to be *negatively* correlated with performance on project implementation in the Nigerian bureaucracy. Earlier studies of incentives in policing also found negative effects on outcomes. Baicker and Jacobson (2007) show that providing financial incentives to police departments from drug seizures leads to a decrease in effort in other areas, and Prendergast (2001) argues that subjecting the LA police to external oversight led to a sharp reduction in their effort to fight crime.

A “universalist” interpretation of the negative effect of incentives is that strong incentives are simply inappropriate in some government jobs. The seminal paper by Holmstrom and Milgrom (1991) highlights the problem of multitasking: the nature of a civil servant's job may be such that he has to attend to multiple differing objectives, some more easily measured than others. Incentivizing civil servants based on performance will then attract attention away from tasks that are not easily measured, potentially with counterproductive effects. This may be particularly true for the police, since criminals have the best information on police enforcement effort, but the least incentive to complain (Prendergast 2001). Alternatively, if achieving

performance outcomes requires collaboration in large teams, then free riding may dominate any positive incentive effects, as found in the British bureaucracy by Burgess et al. (2017).

An alternative “contingency” perspective is that the problem is not with incentives per se but the specific way they are implemented in these government departments: perhaps the managers interviewed by Rasul and Rogger (2018) did not know how to design incentives in a way that avoids the multitasking pitfalls. Conceptualizing effective reforms within an organization may be particularly challenging if, as in the case of the police, existing managers strongly subscribe to a particular worldview (Tripsas and Gavetti 2000). If managers cannot always be relied upon to choose the incentives that work best, then the effect of incentives is mediated through the competence and worldviews of the managers, and the appropriate action is not to abandon incentives, but to give managers the kind of specific advice that helps them choose the right incentive system.

The police reform literature illustrates these two perspectives. This literature (recently reviewed in Gramckow et al. 2016 and Hope 2015) consistently identifies a set of broad “best practice” goals of police reform—community policing, increased training, increased accountability, etc.—and documents a series of case studies of how various police forces applied these principles. Yet the details of each reform program are extremely heterogeneous, from increasing wages in China to wholesale replacement of the police force in Georgia (Kakachia and O’Shea 2012). Thus, a primary conclusion of the literature, taken as a whole, is that the practical details of optimal police reforms are highly context dependent with many possible choices. This puts tremendous pressure on senior officers to select the right interventions, and indeed the literature documents many failed reforms.

This paper sheds light on these issues through five randomized experiments carried out in partnership with police department in the Indian state of Rajasthan, the first RCT on police reform in a developing country. Guided by recommendations from high-level police commission reports, senior police officers selected four reforms designed to improve the behavior of police officers with the general public and crime victims. The first was a freeze on transfers of police staff from station to station. The second entailed training police staff, with a strong emphasis on “soft skills” such as sensitivity and communication. The third was a reform of intrastation personnel management to improve working conditions: rotation of duty assignments and guaranteed days off. The fourth was to install civilian observers at police stations in an effort to improve community monitoring and understanding of the police. The fifth policy, which emerged during the research process, was designed to increase incentives for police to register crimes that were reported to them: surveyors posing as crime victims (henceforth referred to as “decoys”) were sent to police stations to register imaginary crimes. The decoys then revealed their role, giving the police a sense that they were being monitored.

To evaluate these reforms, we collected data through two rounds of surveys including police interviews, surprise visits by surveyors to police stations, and a large-scale representative crime victimization survey—the first of its kind in India.

In a few areas, the reforms showed positive results. Even the limited incentives inherent in the decoy visits increased the probability of case registration and officer

politeness with complainants. The “soft skills” part of the training experiment also worked: trained station staff were more polite with crime victims, and victims were more satisfied with the police. In contrast, the other three interventions were less successful, and the evidence suggests they were not implemented properly. This is perhaps not surprising since these programs in effect *reduced* the autonomy of the middle manager (the police station chief, in this case), violating one of the broad management principles found to be important in the literature. Therefore, the chiefs of the police stations had every reason to try to undermine them.

Taken together, these results argue against the view that managers can readily identify the interventions that work best in their context. The broad management principles that apply in private firms or in hospitals may very well work in the police context too, but the management needed help to identify the specific interventions that would help translate these into concrete, implementable programs. Outsiders, whether consultants or researchers, may have a role in revealing weaknesses in current practices and the space to design new ones. Of course, we have only five interventions that we are comparing; it could be the case that some unidentified force caused the interventions we suggested to work, while three of the four chosen by the police did not.

I. Program Objectives and Design

A. Policing and Police Reform in India

As in many developing countries—e.g., parts of Africa (Opolot 1992), the Philippines (Varona 2010), etc.—the Indian police largely retained the organization and rules established by the former colonial power. Indeed, the Police Act of 1861 is still in effect 72 years after independence. As might be expected of a colonial law, the emphasis of the Police Act is on keeping the populace quiescent, in part through intimidation, rather than helping them with their problems (Kolsky 2010).¹ Moreover, the military-style hierarchical organization, with its rigid pay scale and promotion schedule and no guaranteed leave, allows essentially no scope for formal incentives or disincentives. Since 1977, a succession of police reform commissions have recommended wide-ranging changes, but they have very rarely been implemented.

The recent history and current position of the Rajasthan State Police are broadly consistent with that of police across India. In 2007, the total force consisted of 70,767 personnel, divided among 711 police stations with 76 percent constables—the lowest rank—and only 0.2 percent members of the elite Indian Police Service (IPS) cadre that supplies most of the senior officers. Crime rates were comparable with other Indian states: Rajasthan ranked fourth out of 28 states in total crimes serious enough to be investigated without court order, but ranked twenty-fifth in murder rates.

¹This is especially the case because the force is extremely overstretched. According to 2011 UNODC data, India has 129 police per capita, versus 227 in the United States. See <https://dataunodc.un.org/data/crime/Police%20personnel>.

This long history of estrangement from the public was apparent in the data collected in the baseline of our study. The representative household survey data that we collected (it turns out, for the first time in India) showed that only 29 percent of crime victims had bothered trying to register the crime, often because they felt that the police would not make an effort to assist them. They were probably correct. Field surveyors posing as regular citizens trying to report incidents were sent away without being able to file a report 52 percent of the time. The lack of trust is further demonstrated by the 53 percent of baseline survey respondents who claimed that law-abiding citizens fear the police.

B. Police Reform Experiment

This alienation of the public from the police was becoming increasingly costly in a time of rising terrorism, when it was important to get citizens to report anything that they felt might be suspicious. Against this backdrop, in 2005 the Director General of the police in Rajasthan, A.S. Gill, took the unusual step to request our team (made up of Nina Singh, then a senior officer in the Rajasthan Police and an author of this article, and researchers from J-PAL) help the department identify a set of reforms that could both improve the public perception of the police and their effectiveness through randomized control trials.

The main interventions for this experiment were chosen by the senior state-level police officers after discussion with the research team and were strongly influenced by the reports of the various police commissions, which were cowritten by retired senior police officers.² Within that set of options, choices were constrained by our requirement that the unit of implementation be a police station (rather than something larger) to ensure enough independent units to randomize across. With a view toward potentially scaling up the interventions throughout Rajasthan and elsewhere, the police leadership ruled out any reforms that would require a large increase in budget outlay or manpower, or require changes in the fundamental processes used by the police (changes in the pay scale, performance bonuses, etc.). Senior officers argued that, based upon their experience, interventions to reduce corruption, bribe taking, or the ability to handle violent law and order situations would also be politically sensitive and unlikely to be successful.

Throughout the project planning phase, a consistent theme in selecting optimal reforms was that station-level police staff, from the chief to the constables, face perverse incentives in their day-to-day jobs. The clearest example lies in the choice to register a crime report: since local police are evaluated based upon crime rates, which are measured by case registration, refusing to register a crime would reflect well on a police station's crime statistics. Similarly, because station-level officers were transferred to new locations with high frequency, they had little incentive to make long-term investments in developing trust and positive relations with the local community. Both senior state-level officers and constables argued that police station

²The freeze on transfers, guaranteed time off, and training interventions came directly from the Padmanabhaiah Committee on Police Reforms, constituted largely of active and retired senior police officials, while the community observer program was developed in response to the committee's recommendations on community policing.

chiefs engaged in favoritism: giving desirable assignments and allowing vacations for a small group of preferred personnel while overworking others.

This perspective highlights the negative aspects of managerial autonomy: that managers' incentives and preferences may not be aligned with those of the larger organization. From this viewpoint, allowing greater autonomy may worsen the effect of perverse incentives. Yet, in the context of the Indian police, replacing these negative incentives with positive, overt incentives seemed an intractable problem. There was no prospect of organizing a national crime victimization survey to provide an independent measure of crime, and providing explicit financial or promotional incentives to police staff would have required major legal changes. Thus, the reforms adopted focused on improving police performance through removing a clear source of negative incentives (rapid transfers), and more generally on subjecting local police chiefs to more control and monitoring.

The five main interventions were:

- (i) **Freezing of Transfers:** In the Indian police, transfers are one of the very few incentives instruments in the hands of supervisors. Partly as a result, transfers are very frequent: one-third of all police officers were transferred in our control stations over a period of eighteen months. While nominally under the control of the police, transfers are frequently imposed by politicians (who are able to apply pressure at the relevant level), often allegedly for partisan or corrupt motives (National Police Commission 1980, Centre for Media Studies 2005, Wade 1985). Unchecked transfers were seen as both bad for the morale of the police and a disincentive for them to invest in acquiring local knowledge, which in turn might affect their police work and relationship with the population. The intervention specified *that for personnel posted to all police stations selected for this intervention (except the control group) all administrative transfers would be frozen for period of at least two years. Exceptions could be made for well-documented cases of police misconduct.*
- (ii) **Weekly Day Off and Duty Rotation System:** Indian police procedural regulations provide little guidance on management of human resources at the police station level. In theory, police staff remain on duty for 24 hours a day, seven days a week, and all assignment of leave and duties (such as clerk constable, night patrolling, etc.) is at the discretion of the police station chief. Such an ad hoc system is susceptible to accusations of favoritism in assigning desirable duties, and many officers report feeling overworked and burnt out (Sidhu 2003). Replacing this with a transparent and fair system of work allocation might lead to less jockeying for positions, more predictability, better informed officers, and higher overall productivity. The intervention specified *that duties at the police station level were to be allocated to staff on a rotating basis according to a written schedule, posted in a public place. The entire staff in selected police stations (except the station chief) received one day off every eight days. In smaller police stations, where the shortage of manpower might be more acute, the station chief had the option of extending the work period up to fifteen days.*

- (iii) **Community Observers:** Recognizing the lack of police accountability vis-à-vis the public, and the limited communication between the police and the public, several of the Indian police reform commissions have recommended the creation of civilian oversight institutions for the police. They have also emphasized the need to communicate to the public a “clear understanding of the limitations and constraints within which the police has to function” (NPC 61.46). Such experiments have been tried in several developed countries, including the post-Soviet states (Caparini and Marenin 2005) and South Africa (Bruce and Neild 2005) but this was the first time in Rajasthan. The intervention specified *that two volunteers drawn from a (long) list of potential volunteers would spend about three hours in the police station during peak operating hours, on rotating basis. The observers’ sole task was to watch the activities within the police station and become familiar with the duties, procedures, and challenges faced by the police.*
- (iv) **In-Service Training Program:** A lack of skills, both professional (investigation methods, etc.) and “soft skills” (mediation, communication) have frequently been recognized as significant barriers to effective policing.³ Constables are recruited mainly on physical characteristics (height, chest measurements, etc.), and only education through class 10 is required. After the initial training of 12 months for constables or two years for sub-inspectors, further training is only given at the time of promotion, and this does not include any soft skills component. Under this intervention *randomly selected police personnel were given training in at least one of two modules:*
- Professional/Investigation Skills: 292 investigating officers (inspectors, sub-inspectors and assistant sub-inspectors) were trained for six days at the Rajasthan Police Academy with inputs on improving investigation procedures, such as field techniques and documentation, with emphasis on scientific techniques.
 - Soft Skills: 1,541 police personnel of all ranks were trained for three days on improving attitude with the public with inputs on “soft skills,” such as communication, mediation, stress management, motivation, team building, leadership, attitudinal change, etc.
- This training was rolled out across the state of Rajasthan in groups of two to three districts in the months after the other reform interventions had begun.
- (v) **Decoy Crime Victims:** To check whether our interventions improved police behavior and willingness to register cases, enumerators posing as regular citizens attempted to register complaints at the police station.⁴ These visits were

³ See “Rebuild India’s Police Forces.” *The Hindu*, September 9, 2011. <https://www.thehindu.com/opinion/editorial/rebuild-indias-police-forces/article2439809.ece>.

⁴ The crimes chosen were not representative of the entire range of possible crimes; rather, we chose scenarios where the crime was relatively petty, which is where the tendency to underregister may be the greatest: stolen mobile phone (29 percent), stolen suitcase (25 percent), stolen bicycle (15 percent), stolen railway ticket (10 percent), harassment of female relative (9 percent), victim of dangerous driving (8 percent), domestic violence (2 percent), burglary (1 percent).

unannounced and the police learned only ex post that they had been assigned to a decoy visit, either when the field officer revealed himself in order to prevent the registration of a false case, or (if the police refused to register the case) when a supervisor called shortly after the visit. Thus, the station personnel's knowledge of the intervention did not depend on their behavior during the decoy.

Early in the project, our team observed that the police were animatedly discussing decoy visits after they happened. This suggested that each actual decoy visit would reinforce the message that these visits were happening, and that knowledge would potentially have an effect on behavior (even though police staff had been assured that the monitoring was not linked to a specific punishment). *We therefore decided to transform the decoy visits into an intervention. We randomized the order in which the decoy visits were carried out and used police staff behavior at subsequent visits as the outcome. All police stations were assigned to be visited by five decoy crime victims, and a randomly chosen subset were assigned to a sixth visit.*

Consistent with the viewpoint maintained during selection of the reforms, the first three reforms all reinforced the authority of the state-level police and/or reduced the autonomy of the police station chiefs. By freezing transfers, station chiefs could not petition to alter the composition of their staff, nor could they escape a potentially unpleasant posting themselves. The duty rotation intervention moved the assignment of tasks within the police station out of the discretion of the police chief and required treated stations to create and post a calendar of duties in which all staff were to be rotated to all duties. Similarly, the weekly day off (which was integrated in the duty rotation) intervention mandated that each member of the police staff, except the police chief, should have a rotating day off. Under this new system, the police chief had no explicit role in assigning duties or approving vacations, weakening his authority to demand compliance and work from his subordinates at any day and any time. Finally, the community observer's presence in the police station was intended to encourage the staff to be on "best behavior," which may also have reduced the set of actions the police chief could take (at least in the observer's presence).

A second common element of all the interventions was their very low cost. The freeze on transfers, duty rotation, and community observer programs had zero financial cost, although there might have been a short period in which constables' productivity suffered due to learning new tasks and some up-front staff time required to recruit the community observers. The main cost of the weekly day off was in terms of availability of police staff—in principle, personnel would have to increase by 14.3 percent to compensate for the new 1 day off every 8 days in larger treated police stations. In practice, senior officers claimed that the actual effect would be much smaller—potentially zero—partly because many constables were already receiving substantial informal time off, and because instituting such a system might increase productivity during the rest of the week. The most expensive intervention was the UNODC-funded training, which cost roughly \$23,000 for trainers, plus the opportunity cost of one week's time away from police stations for officers and three days

for constables. Interestingly, despite its higher cost, this training intervention was the only one adopted and scaled up after the intervention ended.

Unlike the actual consultants in BEMMR, the researchers had very little interaction with mid-level managers and provided no direct guidance. With the exception of the decoy program, which was run by the research team, the primary implementation of the reforms was performed through the police hierarchy, with overall supervision done at the level of state police headquarters. The district superintendents of police supervised the implementation in their respective districts and, except when specified otherwise by the experiment, police station chiefs were responsible for carrying out the interventions in their stations. Thus, any effects observed are due to changes in management policies, rather than changes in the managers themselves.

C. Evaluation Strategy

Partly to ensure statistical power and partly to make sure that the experiments took place at scale, the interventions covered large parts of the state: 162 police stations across 11 districts of Rajasthan were included in the sample, with districts chosen to be geographically representative of the state of Rajasthan. The sample represents a substantial fraction of the state: over 20 percent of the police stations in the state and a population of almost 8 million people. Not only was this scale necessary to get adequate power, it also gives meaning to success (or failure) in implementation; this was not a “gold-plated,” small-scale experiment, but a full-scale rehearsal of what the reforms would look like if they were to be broadly implemented.

Police stations were randomly assigned (by the research team, using a computerized randomization) to different reform strategies, using stratification to ensure a balance on geographical area (district), criminality (number of cases registered in 2005), and urban/rural environment. These variables are included as controls in all regressions performed in this study. Within each police station, staff members were randomly selected for the training intervention with the total percentage of staff trained in each station randomized at 0 percent, 25 percent, 50 percent, 75 percent, or 100 percent of station personnel. This allows us to detect whether the effects of training are different when few or almost all the officers are trained: is it sufficient to train a few police officers, who will then train their colleagues? Or is a critical mass necessary to get any impact?

Surveyors visited sample police stations to collect data on a relatively regular basis. To assess a possible Hawthorne effect—the possibility that the police stations participating in the program might change their responses or behavior simply because they are being observed by outside evaluators—an additional group of police stations was also randomly selected to serve as the “pure control.” These police stations were not informed of the project, had no staff trained, and were never visited by investigators until the final endline survey, at which point they were surveyed like all other stations. The Hawthorne effect can be estimated by comparing the control and “pure control” groups.

Table 1 describes how the police stations were allocated to each evaluation group, with the number of police stations in each group in column 1, and the program

TABLE 1—PROGRAM DESIGN

Group	Number of police stations (1)	Program elements			Percent staff trained				
		Community observer (2)	Weekly day off and duty rotation (3)	No transfer (4)	0	25	50	75	100
1 All interventions	35	Yes	Yes	Yes	7	7	7	7	7
2 Community observer + no transfer	25	Yes	No	Yes	5	5	5	5	5
3 Weekly day off/duty rotation + no transfer	25	No	Yes	Yes	6	5	5	4	5
4 No transfer	40	No	No	Yes	8	8	8	8	8
5 Control	25	No	No	No	8	5	4	4	4
6 Pure control	12	No	No	No	12	0	0	0	0
Total stations with program element:	162	60	60	125	46	30	29	28	29

elements implemented in that group in columns 2–4. As shown in column 4, wherever the community observer or weekly day off/duty rotation interventions were attempted, transfers were also frozen. This was done for two reasons: first, because long tenure of staff is considered necessary to get acquainted with the area and people of the police station jurisdiction and see if this has any effect on their performance; and second, due to concerns about attrition during the study. This implies that the impact of duty rotation and community observers (or the combination of the two in the all interventions stations) will be seen on top of the impact of the freezing of transfers.

Finally, to evaluate the impact of the decoy intervention, we randomly selected the order in which each police station would be visited by a surveyor pretending to register a crime. As shown in Table 2, decoy visits were rolled out in overlapping rounds, so that, for example, the final police stations of the third round were receiving their decoy visits during the same period as the first stations of the fourth round. This implies that, conditional on calendar time, the number of past decoys performed is randomly assigned for any given police station. All police stations were assigned five or six decoy visits, although three stations were visited only four times due to implementation issues. The scripts used to attempt to register cases, as well as the identity of the surveyors, were also randomized. Surveyors only attempted to register a decoy case in police stations they had never previously visited.

II. Data

A series of surveys and interviews were conducted to establish baseline levels of police performance and baseline public perception of the police, as well as to estimate the effectiveness of the police implementation at the endline. These included traditional household surveys, surveys of police personnel, as well as random, unannounced visits to all police stations to collect data on program

TABLE 2—DECOY CRIME VICTIM VISITS

Survey month	Number of visits	Percent visits in which police suspected decoy	Decoy visit number					
			1	2	3	4	5	6
1	148	20.9	148	0	0	0	0	0
2	162	15.4	2	137	23	0	0	0
3	131	27.5	0	13	118	0	0	0
4	154	18.8	0	0	9	144	1	0
5	80	16.3	0	0	0	5	75	0
6	75	13.3	0	0	0	1	69	5
7	38	21.1	0	0	0	0	2	36

implementation. These rich datasets are one of the outcomes of the project; they are available for public use at Banerjee et al. (2011).⁵

Visit of Decoy Victim Surveyors.—As mentioned above, the “decoy” intervention was originally conceived as a source of objective data on officer behavior. Immediately after attempting to register a case, the surveyor completed a short form recording his success or failure, and the attitudes and actions of the police. The field officer also recorded their impression on whether or not they thought the police officer may have suspected them to be a decoy.

Public Surveys.—The other main source of data was a household survey that recorded both general perceptions of the police and the experiences of those individuals who happened to be victims of crimes in the previous year. In total, 5,895 households were interviewed: in the spring of 2007, a baseline was conducted with 2,454 households randomly chosen from the 2004 voters list. In the autumn of 2008, the households that could be located again were reinterviewed along with 3,441 newly chosen households.

The crime-related surveys were broadly based upon the International Crime and Victimization Survey from the UN’s Interregional Crime and Justice Research Institute (UNICRI), customized to the Indian context. Within each household, three modules were administered (at baseline and endline). First, a general crime-screening questionnaire administered to the head of the household, inquired whether any member of the household had been a victim of a crime in the previous year. If any victims were reported, a crime victim survey recorded much greater detail about the crime. Finally, an opinion survey, administered to a randomly chosen adult in the household, asked about their contact with the police, and perceptions of police performance and integrity. Like any crime victimization survey, these questionnaires do not capture “victimless crimes” such as gambling or drug use. Also, no attempt was made to measure the incidence of domestic violence because of concerns about accuracy and decreased respondent cooperation.

⁵ See <https://doi.org/10.7910/DVN/OC35YC>.

Case Investigation Grading.—To measure the impact of the training (and potentially other interventions) on the quality of the actual police work, 982 case files were randomly selected from the project police stations and sent to a group of retired senior police officers for grading in two rounds, half before the training and half post-training. The retired officers filled out a detailed confidential report about each case in which they graded the performance of the officer on his actions at the crime scene, whether scientific techniques were used, and the care with which the evidence was collected and documents in the case file were prepared.

Due to high turnover of police staff—despite the freeze of transfers that was supposedly in operation—the sampling of the case investigation grading is not representative of the sample of officers being trained. Rather, we sampled at endline to ensure representativeness of officers working in the police stations at that point, and we will compare the quality of the investigation at the police station level. This comparison will therefore not give us the total effect of the training (since a trained investigator may now be working in a control training station), but it will shed light on the channel through which police training improves victim and public satisfaction.

III. Results

A. Crime and Policing in Rajasthan: Baseline Status

The baseline data collected for the project provides further evidence on the context and background for the interventions and the study. In particular, it confirms many of the problems discussed in the introduction, which motivated our choices of interventions. Table 3 presents some descriptive statistics with highlights from the surveys at baseline.

Crime.—The detailed crime survey provides the first large-scale representative estimates of the crime rate in Rajasthan, since India does not have a standard crime victimization survey. We find that 1.7 percent of individuals were victims of a crime in the past year, and that 5.9 percent of households had at least one member who was a victim of a crime. This estimate excludes crimes of domestic violence, which were not surveyed because of accuracy concerns. Theft was by far the most common, constituting 37.9 percent of all reported crimes, followed by burglary at 16.6 percent and assault at 12 percent.

Most crime victims never report their incidents to the police. Only 29 percent of the crime victims we surveyed stated that they had visited a police station in an attempt to report the crime. This rate is somewhat lower than most developed countries: 42 percent of all US crime victims reported their case to police (NCVS 2009–2010). More broadly, the average reporting rate over a set of 33 countries, but for a more limited group of 5 property crimes, is 47 percent (van Kesteren et al. 2000), and the average likelihood of reporting a crime over a larger set of 10 crimes in 17 industrialized countries is 49 percent (van Dijk, van Kesteren, and Smit 2007).

Among those who attempted to report their crimes, 17 percent did not succeed in registering a case due to police resistance. The comparison of crime surveys and official police records confirms that per capita crime rates are substantially higher

TABLE 3—SUMMARY STATISTICS

	Control (1)	Any treatment (excludes decoy) (2)	Difference (3)
<i>Panel A. Baseline opinion survey respondents (observations = 2,436)</i>			
Crime victim in household	0.0547 (0.228)	0.0610 (0.239)	0.0063 (0.015)
Ever met police	0.0844 (0.278)	0.0915 (0.288)	0.0071 (0.018)
Ever visited police station	0.0943 (0.293)	0.0846 (0.278)	−0.0097 (0.017)
Ever arrested	0.0223 (0.148)	0.0133 (0.115)	−0.0091 (0.010)
<i>Panel B. First round decoy survey (observations = 150)</i>			
Case not filed	0.520 (0.510)	0.616 (0.488)	0.0960 (0.110)
Police very polite to decoy	0.200 (0.408)	0.144 (0.353)	−0.0560 (0.087)
Police rude to decoy	0.120 (0.332)	0.192 (0.395)	0.0720 (0.074)
<i>Panel C. Baseline victim satisfaction</i>			
Victim satisfied or very satisfied (observations = 141)	0.348 (0.487)	0.356 (0.481)	0.008 (0.118)
Case registered with police (observations = 373)	0.311 (0.467)	0.301 (0.460)	−0.010 (0.066)
<i>Panel D. Preintervention case file review</i>			
Documentation quality (observations = 436)	6.409 (1.797)	6.133 (1.931)	−0.276 (0.300)
Field investigation quality (observations = 449)	6.537 (1.636)	6.160 (1.823)	−0.378 (0.272)
Whether used scientific methods (observations = 523)	0.179 (0.386)	0.202 (0.402)	0.023 (0.053)

Notes: This table shows the difference between police stations assigned to control versus those in any of the training, weekly day off/duty rotation, transfer freeze, or community observer interventions at the beginning of the program. Statistics in panels A and C reported for baseline household/victim survey, panel B reports data from the first decoy visit, and panel D reports reviews of cases prior to the intervention. Columns 1 and 2 present means with standard deviations in parentheses. Column 3 reports differences with standard errors in parentheses, clustered at police station level for panels A, C, and D. The victim satisfaction outcome in panel C was recorded only conditional on reporting the case to the police.

from survey data, especially for property crimes: survey theft rates are 9.24 times the registered rates and robbery rates are 11.35 times higher than in police records. The differential decreases for violent crimes: rape or molestation surveyed rates are 43 percent rates reported to the police, and assault is only 11 percent higher in the survey (neither difference is statistically significant). While the biggest reason for not reporting a crime was lack of importance perceived by the victim (28 percent), substantial numbers of crime victims did not go to the police because they thought that the police were incapable of helping (20 percent) or unwilling to help (17 percent).⁶

⁶An interesting analysis of the discrepancy between reporting and actual crime is provided in Iyer et al. (2012). They show that reporting of crime against women increases when political positions are reserved for women at the

Police Attitude with Crime Victims.—The poor behavior of police officers with victims is clearly demonstrated by the results of the decoy survey. Overall, in the first decoy visits conducted by our investigators, police were willing to register a First Information Report (know as an FIR, the necessary first step to a criminal investigation) from the decoys only 54 percent of the time. More serious and verifiable cases are registered significantly more frequently: house break-ins led to 92 percent FIR registration while stolen mobile phones had only 40 percent success in registration. Crime victims who attempted to file an FIR had better luck on average (83 percent), likely due to the high threshold before deciding to actually bother to register: they were less likely to attempt to report the sort of minor crime that were reported by our decoys. For example, only 7.4 percent of theft victims who lost less than 1,500 rupees (roughly the value of a mobile phone) even attempted to register their case with the police.

Decoy visits also allow us to assess the politeness of the police officer. They found that 31 percent of the time the police were not polite, and only 11 percent of the time they were “extremely polite.” We took care that decoys would look and act like typical young men in the area, although if they had been identified as decoys, the rate of registration would presumably have been greater, not lower.

Public Perception of Police and Victim Satisfaction.—Few respondents have experience with the police or have ever interacted with them. Only 9 percent of those interviewed have ever had an interaction with the police in their lives, and only 3 percent of women have spoken with a police officer. Even in urban areas, only 20 percent of male respondents ever reported interacting with the police. In the absence of personal experience, 72 percent of those surveyed claimed that word-of-mouth discussions affected their view of police, with only 14 percent and 16 percent stating that they also based their opinions on print and television news sources, respectively. This pattern of information dispersal—opinions formed by hearsay and few real experiences—would suggest that an intervention like the community observer program, which gives a more citizens positive experiences with the police, has potential to improve perceptions of the police.

Unlike the average citizen, crime victims frequently have personal experience with the police. Note, 39 percent of victims were either completely satisfied or satisfied, but 35 percent of victims reported being completely unsatisfied, and 27 percent mostly unsatisfied. Further, 82 percent of the unsatisfied victims stated that they were unhappy because the police “didn’t do anything special” or “didn’t take interest” in their case. In contrast, only 15 percent complained of police corruption, and another 20 percent complained of discourtesy when attempting to report a crime.

These numbers are similar with the only comparable India study (Sarkar et al. 2015), which finds a 44 percent satisfaction rate in the much larger cities of Delhi and Mumbai. They are substantially lower than satisfaction rates found in US studies. In the United States, Ekins (2016) finds 67 percent of Americans satisfied with their recent encounter with police, and Brandl and Horvath (1991) find 69 percent

local level, both in our data and in nationally representative data. Nevertheless, our data shows no increase in the actual incidence of crime.

satisfaction among American crime victims. The average satisfaction with police performance on five common crimes⁷ for the 32 countries taking part in the ICVS (van Dijk, van Kesteren, and Smit 2007; mostly European or other developed countries) was 57 percent, with the most common reason for dissatisfaction also that police “did not do enough.”

Case Review Performance.—The reviews of case files by retired officers revealed substantial heterogeneity in the quality of investigations of different types of cases. The retired officers gave 6.18 out of 10 on average for overall field investigation quality, and 5.91 out of 10 for quality of case documentation. These grades were very dispersed: standard deviations were 1.95 for field investigation and 2.42 for documentation. Only 19 percent of cases investigated during 2006–2007 and reviewed by our reviewers made any use of scientific investigation, and most of these were accident cases.

B. Reform Implementation

In the previous literature, a major obstacle in the implementation of bureaucratic reform at scale has been the resistance to sustained implementation on the part of the middle management, on whom the reform where often imposed. We document similar difficulties in this setting for all the programs that relied on implementation by the station chiefs. In Section IV, we discuss how these challenges, in the context of the organization of the police, shed light on the broader limitations to internal institutional self-reform.

No Transfer.—The extent to which transfers were actually frozen is measured through administrative data from personnel records. We matched names of officers from police station staff lists at the beginning of the program with those from the end of the program, and counted any officers not found as having been transferred between the beginning and the end of the project. Table 4 shows that there was indeed a significant reduction in the number of transfers out of police stations in the ranks that were targeted: 12 percentage points for constables (or a 44 percent reduction from the 29 percent mean in the control group), and 16 percentage points for inspectors (a 23 percent reduction from the 60.6 percent rate of transfer in the control group). Still, transfers during the course of the project remain quite frequent (21 percent on average across all ranks in the treatment group), particularly at higher ranks. These figures are lower bounds, since some personnel may have been transferred into then back out of the sample police stations during this period and thus not included among the transferred staff.

While far from a complete freeze, the ban thus resulted in significantly longer posting duration. Transfers require approval from a senior officer at headquarters, and therefore the ultimate control resides at the top, which insured some amount of adherence to the policy. However, there were pressures to transfer particular

⁷Theft from a car, burglary, robbery, sexual offenses, and assault and threat.

TABLE 4—POLICE STAFF TRANSFERS BY RANK

	Control PS (1)	Treatment PS (no transfer) (2)	Difference (3)
Inspector	60.00%	44.44%	15.56% (17.96)
Sub-inspector	66.67%	46.50%	20.17% (8.47)
Asst. sub-inspector	29.73%	27.04%	2.69% (6.25)
Head constable	34.48%	28.33%	6.16% (7.00)
Constable	29.16%	16.77%	12.39% (3.71)
All ranks	32.09%	21.03%	11.06% (3.21)

Notes: Standard errors in parentheses clustered by police station. Each cell reports the percentage of police staff of a given rank who were posted in a police station at the beginning of the intervention but were not posted there by the end of the intervention. All results calculated from administrative roster data provided before and after the interventions.

staff members both from below and from above (i.e., from politicians), and some instances were more successfully resisted than others.

Training.—The training was the easiest intervention to systematically implement, and there is administrative evidence that it took place largely as planned. Rajasthan Police Academy records show that 88 percent of officers selected for training reported to the academy for training. Although local police chiefs voiced complaints about the temporary reductions in manpower, in this area senior officers have tight control over implementation and could bypass police chiefs to instruct specific constables to report to training. Because existing Rajasthan police trainers had no expertise in soft skills training, the training component was performed by an outside firm.

Weekly Day Off and Duty Rotation.—To measure the strength of implementation of the remaining two interventions (the combination of weekly day off and duty rotation, and community observers), surveyors visited police stations several times during the course of the project.

On paper, the implementation of the weekly day off program was good: 84 percent of treatment police stations reported executing some type of rotating day off, whereas only 34 percent of control police stations reported having any scheduled day off. In order to verify that the police station chief was in fact giving days off, surveyors randomly selected two constables from the staff and interviewed them separately at each of their visits, asking them to recall the last time that they had had a day off for any reason, the weekly day off program or otherwise. The results, presented in Table 5, panel A, show that early on, the weekly day off program did succeed in shortening the time since the last day off, going from an average of 30 days in control stations to about 21 days in treatment stations. This difference shrinks over the course of the program until eventually disappearing by round five, our last visit.

TABLE 5—IMPLEMENTATION OF INTERVENTIONS

	Round 1	Round 2	Round 3	Round 4	Round 5
<i>Panel A. Weekly day off: Days since last day off</i>					
No weekly day off	29.3	33.0	29.2	29.7	30.3
Weekly day off	21.2	21.8	21.4	29.5	32.2
Difference	8.1	11.2	7.8	0.2	-1.9
<i>Panel B. Number of days constable has been doing current duty</i>					
No duty rotation	4.1	4.0	6.1	4.3	2.7
Duty rotation	5.6	4.5	5.4	4.5	2.8
Difference	-1.5	-0.5	0.8	-0.1	-0.1
<i>Panel C. Whether community observer is present</i>					
Community observer station (percent)	20.5	8.5	6.7	8.3	8.3

Notes: Results in panels A and B are based upon data collected in interviews with randomly chosen constables conducted during surprise surveyor visits to the police station. Results in panel C are calculated from surveyors' direct observations during these visits. Surveyors did not record community observer presence in stations not assigned to that intervention.

The implementation of the duty rotation was also close to complete on paper, with 91 percent of station house officers of treatment police stations able to produce a duty roster when asked. Here, 58 percent of program police stations posted the duty roster on the wall as stipulated by the implementation guidelines, as opposed to 39 percent of control police stations, which had independently created duty rosters. Table 5, panel B shows the responses over time to questions about the rapidity of changing duties. It turns out that, even in control stations, the rotation from duty to duty is rapid, if haphazard. Overall there is no significant impact on time spent on specific duties.⁸

Community Observers.—Because police stations were required to publicly post the times that community observers would be present, surveyors could schedule their visits to the police station in order to coincide with the observer's presence and thus monitor the implementation of this branch of the intervention. Table 5, panel C, reveals that even in the first round of visits implementation was far from complete (20.5 percent), but this rapidly decreased so that on average observers were present only 10.29 percent of the time during designated hours—and were about as likely to be at the station at any other time. A villager wishing to visit the police station when a community observer was present would be very hard pressed to decide when the right time would be.

Overall, it appears that the community observer intervention was not properly implemented on a sustained basis. Interestingly, surveyors reported that many police stations independently maintained community observer sign-in books, although these “guest books” were not an official part of the program, and surveyors did not collect consistent quantitative data on their presence. Nevertheless, these books were

⁸One objective of the rotation was to create some predictability in the assignment to limit the jockeying for positions. Constables in police stations with duty rotation did have a significantly higher probability of reporting that they knew what their next duty would be, with an overall average of 50 percent in treatment police stations and 41 percent in controls. Unlike the weekly day off program, this effect appears to have grown stronger over time, with the greatest treatment/control differences in the final periods.

usually filled in completely for previous community observer sessions, including in stations with very poor measured observer attendance. This is consistent with the discrepancy between official records and independent observations of the weekly day off/duty rotation intervention.

Decoy Crime Victims.—Since the decoy crime victims were deployed by the researchers, implementation was straightforward: 27 enumerators conducted 788 police station visits, with all stations except for 3 (due to implementation issues) visited the assigned 5 or 6 times. In online Appendix Table A5 we confirm that the order in which police stations were visited by decoys is uncorrelated with prior characteristics, although it is correlated, purely by chance, with the weekly day off/duty rotation intervention.

A potential concern is that the police were able to detect the decoys, possibly with greater precision over time. To address this issue, we requested the surveyors to report any signs of suspicion: this was recorded to be likely in 3 percent of cases, and possible in 15 percent more. Not surprisingly, these suspicions are highly correlated with case recording and polite treatment, but they are not correlated with the number of past decoy visits (Table 2 and online Appendix Table A6). Later we present specifications controlling for police suspicion and find no significant change in outcomes.

While the average implementation of all interventions was low, data from the first and second rounds show that the project was executed better in the initial months. Anecdotal evidence from the surveyors suggests that police staff may have learned to falsify records to create the impression of full implementation. Handwriting in community observer visitor ledgers was suspiciously consistent, and as mentioned above, almost all police stations were always implementing weekly day off/duty rotation on paper. This is consistent with experience from Banerjee, Duflo, and Glennerster (2008) and elsewhere that agents learn to “play the system.” Alternatively, police chiefs might have learned that there would be little punishment from senior officers for noncompliance. Nevertheless, in this scenario the decoy survey intervention, which was never associated with repercussions from senior officers, might also have lost effectiveness over time. As we show below, this was not the case.

The correlations between station-level characteristics and program implementation unfortunately tell us little about why the interventions were not effective. Execution of the reforms is not correlated with most baseline station characteristics, such as the rank of the police chief, size of the staff, or urban/rural location. We do find that high crime police stations had fewer community observers and find suggestive evidence that districts led by police superintendents selected nationally (rather than promoted from the lower ranks) performed better, in particular on the no transfer intervention. These associations are difficult to interpret, given the potential for correlation with unobservable characteristics. By the end of the program, most of the interventions seem to have been nonfunctional in virtually all police stations, suggesting a systemic rather than idiosyncratic issue with the station-level execution of the reforms. See online Appendix Table A8 (Correlates of Implementation Performance) for full details.

C. Were the Reforms Effective?

We now turn to the impact of the interventions on police behavior. We start from the most direct measure of impact, namely the behavior on the decoy surveys. We then move to crime victim satisfaction measured from our victimization survey, and finally to any possible broader impact on the population.

Using the Decoy Survey Outcomes.—Table 5 presents regressions of the following form for the two outcomes of interest (whether the case was registered and whether the police were very polite):

$$y_{dj} = \beta_0 + \beta_T T_j + \beta_D D_{ij} + \beta_P (\%trained)_j + \beta_X X_{dj} + \epsilon_{dj},$$

where d indexes the decoy visit, j indexes the police station, T_j is a vector of binary variables describing the treatment conditions, D_{ij} is the number of past decoys performed in this police station, $(\%trained)$ is the percentage of the police station staff that were trained, and X_{dj} is a vector of control variables. In particular, all regressions include month fixed effects, crime script used, district fixed effects (districts include multiple police stations but have a single superintendent of police), controls for urban/rural location, and the 2005 crime rate to account for the stratified sampling design. Since order of decoy visits was randomly assigned, conditional on the calendar date, the number of past decoy visits is randomly assigned and indicates impact on police behavior of more past decoy exposure. Standard errors in all regressions are clustered at the police station level to accommodate arbitrary serial correlation across decoy visits within the police station.

Given the combinations of treatment, the treatment variables could have been parameterized in different ways: this, of course, does not affect the results, but it does affect how the tables must be read and interpreted. The list below summarizes our parameterization choices and will be a useful reference for the analysis below:

- *No transfers*: A dummy equal to 1 if transfers were intended to be frozen in the police stations, but the station had no other interventions besides, in some cases, staff training.
- *Weekly day off and duty rotation*: A dummy equal to 1 if transfers were intended to be frozen and the station was assigned the weekly day off/duty rotation intervention.
- *Community observers*: A dummy equal to 1 if transfers were intended to be frozen and the station was assigned to the community observer program.
- *All interventions*: A dummy equal to 1 if the station was assigned no transfers, community observers, and weekly day off/duty rotation.
- *Percentage trained*: Percentage of police staff in the station that were trained.
- *Number decoy visits*: Number of decoy visits that had previously taken place in this police station. Conditional on the calendar date, the number of past decoy visits is randomly assigned and indicates impact of more past decoy exposure on police behavior.

TABLE 6—DECOY SURVEY OUTCOMES

	Case registered			Police were very polite		
	(1)	(2)	(3)	(4)	(5)	(6)
All interventions	-0.0134 (0.049)	-0.0138 (0.048)		-0.0535 (0.038)	-0.0492 (0.039)	
No transfer	-0.0476 (0.045)	-0.0495 (0.044)		-0.0322 (0.036)	-0.0297 (0.036)	
Duty rotation, weekly day off	0.00185 (0.051)	0.00384 (0.051)		-0.00238 (0.038)	0.00306 (0.038)	
Community observer	0.0295 (0.061)	0.0312 (0.061)		-0.0386 (0.038)	-0.0304 (0.038)	
Percentage staff trained	0.101 (0.082)	0.102 (0.083)	0.214 (0.114)	0.102 (0.054)	0.0976 (0.054)	0.184 (0.086)
Number of decoy visits	0.215 (0.068)	0.215 (0.068)	0.219 (0.087)	0.0786 (0.045)	0.0849 (0.045)	0.115 (0.049)
Observations	788	788	788	788	788	788
R ²	0.184	0.187	0.145	0.096	0.123	0.113
District fixed effects	Yes	Yes	No	Yes	Yes	No
Station fixed effects	No	No	Yes	No	No	Yes
Controls for police suspicions	No	Yes	Yes	No	Yes	Yes
Control mean at first decoy visit	0.480	0.480	0.480	0.200	0.200	0.200

Notes: Standard errors in parentheses clustered by police station. All columns report estimates of the linear probability model on the outcome of the surveyors' visits to police stations to attempt to register a case. In columns 1–3, the outcome is an indicator equal to 1 if the police were willing to register a case based on the surveyor's complaint. In columns 4 and 5, the outcome is equal to 1 if the surveyor perceived the police as very polite. The percentage trained variable is equal to the percentage of staff in the police station assigned to training in observations corresponding to decoy visits that occurred after the training, and zero otherwise. All regressions include controls for the month, the number of visits previously performed by the surveyor, and the crime story used. Columns 1, 2, 4, and 5 include controls for the percentage training intervention assignment, whether the police station was urban or rural, and 2005 crime levels. Columns 2, 3, 5, and 6 include controls for whether the surveyor thought police were somewhat or very suspicious that he was a decoy.

- *In study:* A dummy equal to 1 if the police station is not in the pure control group (to identify any possible Hawthorne effect). Since pure control stations were only surveyed in the endline, this effect cannot be identified in regressions with police station fixed effects.

Under this parameterization, the “community observer,” “weekly day off,” and “all interventions” dummies indicate the combined effect of being assigned those additional interventions in addition to being assigned no transfers. So, for example, to calculate the marginal effect of the community observers intervention, we subtract the “no transfer” coefficient from the “community observer” coefficient.

Table 6 presents the results on the decoy visit results, the most direct measure of program effects. The baseline FIR registration rate is low: only 48 percent of crime reports that the decoys tried to get registered in the first visit were in fact filed. Columns 1 and 4 are exactly as above, and column 2 and 5 include in addition an indicator for whether the surveyor thought the police might have suspected that he was a decoy. In columns 3 and 6, we include police station-level fixed effects to control for any random sampling imbalance across police stations.

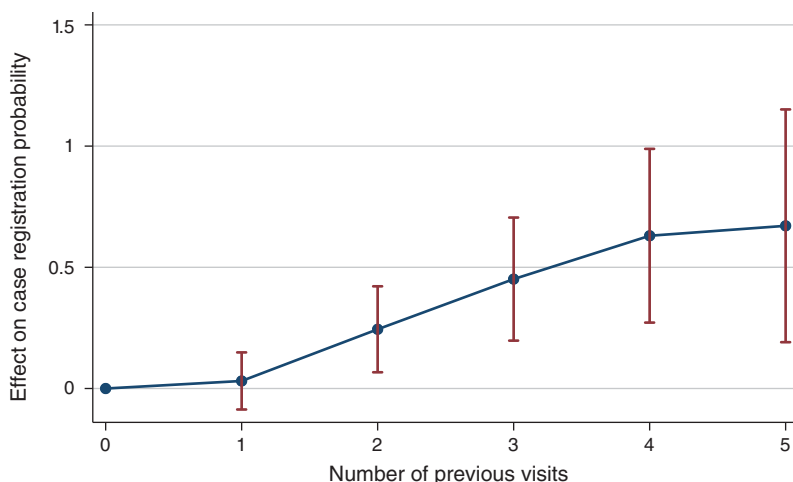


FIGURE 1. EFFECT OF DECOY VISITS ON CASE REGISTRATION

Notes: Graph points are coefficients from a regression of case registered dummies on prior decoy visits, including district fixed effects and controls for case type. Bars show 90 percent confidence interval.

The decoys themselves appear to have a substantial impact on FIR registration. Each previous decoy visit increases the probability of FIR registration at the subsequent visit by 22 percent. Figure 1 shows these effects graphically, plotting the coefficients on dummy variables indicating the number of preceding decoy visits from a regression otherwise the same as column 1 in Table 6. This figure implies that by five previous decoys, the registration rate is close to 75 percent.

We do not believe that this is due to a special treatment of the decoy visitor. Decoys were local people, dressed normally, and reporting very common crimes. Results in columns 2 and 4 show that controlling for cases in which the surveyor suspected that the police were aware he was a decoy does not attenuate the coefficient on the number of previous decoy visits. In addition, the staff training shows a positive effect on registration, becoming much larger and significant when fixed effects are included in column 3.⁹ The other interventions did not have any impact.

Turning to the effect on politeness, both the number of past decoy visits and police officer training had a significant impact on the politeness with which the “victim” was dealt, with $p \approx 0.06$ in columns 4 and 5. Recall that the training for constables (who are in charge of the FIR registration) focused on “soft skills,” and in particular on basic notions of how to interact with victims and regular citizens. This seems to have been effective in changing behavior. Again, the size and significance of the coefficient increases with police station fixed effects in column 6, with the caveat that this source of variations is identified from a slightly different sample.

⁹Introducing police station fixed effects effectively removes the three districts (49 police stations) that had completed staff training prior to the first decoy surveyor visit from the analysis. Thus, the identifying variation for these coefficients comes from only seven of the districts involved in the study.

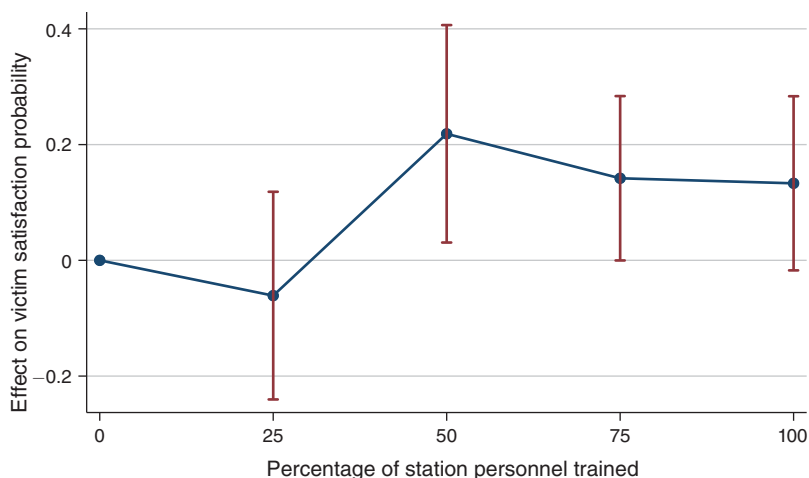


FIGURE 2. EFFECT OF TRAINING ON VICTIM SATISFACTION

Notes: Graph points are coefficients from a regression of victim satisfaction on percentage of staff trained, including police station fixed effects and controls for crime and victim characteristics. Bars show 90 percent confidence interval.

Thus, the decoy survey program had a substantial effect on both FIR registration and behavior of the police, despite the fact that the program was quite explicitly unlinked to any possible sanction: the police officers were warned at the outset that decoy visits could take place, but that the results would not be communicated to their superiors.

Victim Satisfaction with Police.—After the behavior most directly measured by decoys, the second most direct effects of the reform intervention would be found on those citizens who have the most at stake in their interactions with the police: crime victims. We analyze data from the household survey, using the detailed crime reports to infer whether a crime was committed after the intervention and/or training had been conducted in a district, and comparing treatment and control police stations while controlling for district and time effects. In each case in which the victim affirmed that the police had become aware of the crime, we collected data on his or her satisfaction with the police conduct on the case. The interventions had no effect on either incidence of crime (online Appendix Table A2, not surprising, since it was not a target of the operation), or the probability of police becoming aware of the case (online Appendix Table A3). Thus, this selection rule did not induce differential selection in the different treatment police stations.

The main specifications in Table 7 shows the results of the project on the probability that the respondent reports being “satisfied” or “completely satisfied” with the police handling of his or her case. We estimate two specifications. First, a regression similar to the decoy regression above, controlling for survey strata (urban/rural area, past case registration, and district fixed effects) as well as household-level

TABLE 7—PROGRAM EFFECTS ON VICTIM SATISFACTION

	Endline only (1)	Station fixed effects (2)	Station fixed effects (3)
All interventions	−0.00665 (0.0956)	−0.00908 (0.134)	−0.0141 (0.116)
No transfer	0.144 (0.0872)	0.281 (0.123)	0.243 (0.104)
Duty rotation, weekly day off	0.0823 (0.111)	0.204 (0.169)	0.154 (0.164)
Community observer	0.368 (0.117)	0.00272 (0.165)	−0.0576 (0.155)
In study	−0.131 (0.127)		
Percentage staff trained	0.156 (0.0767)	0.193 (0.0843)	0.131 (0.0763)
Registered case			−0.00456 (0.0387)
Asked for bribes			−0.150 (0.0566)
Made arrest			0.166 (0.0618)
Recovered property			0.679 (0.185)
Continued investigation action			0.197 (0.0485)
Recontacted victim			0.0877 (0.0622)
Observations	269	571	571
R^2	0.290	0.222	0.308
District fixed effects	Yes	No	No
Station fixed effects	No	Yes	Yes
Crime and victim controls	Yes	Yes	Yes
Baseline/control mean	0.246	0.299	0.299

Notes: Standard errors in parentheses clustered by police station. All columns contain estimates of the linear probability model of crime victims reporting that they are satisfied or very satisfied with police handling of their case. All specifications include controls for indicators of the type of crime. Victim characteristics: age and gender of the respondent, education and dummies for the occupation of the head of household, indicators for caste or Muslim religion, and indicators for motorcycle ownership. Date of crime controls include dummies for month of crime. All endline satisfaction reports for which victims could not recall the date of the crime were tagged as post-training (24 cases).

characteristics (occupation, religion, sex and age of victim, and motorcycle ownership) and controls for the type of crime:

$$y_{ij} = \beta_0 + \beta_T T_j + \beta_P (\%trained)_j + \beta X_{ij} + \epsilon_{ij}.$$

In this regression, y_{ij} is outcome y reported by citizen i living in the area covered by police station j . In this specification, we use only data collected from crimes that occurred after the training and intervention had both been carried out in police station j . Because the training was rolled out district by district, the relevant control observations for the staff training intervention are those crimes that occurred in the

jurisdiction of a district's control police stations after the staff in that district's treatment police stations had received the training.

Our second approach takes advantage of the police station panel to include fixed effects at the police station level. Here we estimate difference-in-difference type regressions of the form

$$y_{ijt} = \beta_0 + \beta_{T^{FE}}(T_j^{FE} \times POST_{1t}) + \beta_{P^{FE}}(\%trained_j \times POST_{2ijt}) + \beta_X X_{ijt} \\ + \gamma_1 POST_{1t} + \gamma_2 POST_{2ijt} + \eta_j + \epsilon_{ijt},$$

where t now indicates the date at which citizen i was a victim of a crime; $POST_{1t}$ is an indicator for the endline survey, and $POST_{2ijt}$ is an indicator for whether the personnel training had been completed in the district where station j is located by the time crime i was committed. The variation in $POST_{2ijt}$ comes both from the roll out of the training by district as well as from the date at which the crime was committed.

The coefficients of interest (reported in Table 7) are those on the interactions with the treatment vector, $\beta_{T^{FE}}$ and $\beta_{P^{FE}}$. The term X_{ijt} is a vector of the controls listed above, and η_j denotes the police station-level fixed effects. To conserve space we do not explicitly write the full description of the interaction variable into the row labels of the tables when reporting results from regressions with station-level fixed effects. Standard errors are clustered at the police station level.

Consistent with the decoy results on politeness, the training program shows a large impact on the satisfaction of crime victims, both in the fixed effect and the endline-only specifications. The effect of going from 0 percent trained officers to 100 percent is to raise the probability that victims are satisfied with police investigation by between 15 and 19 percentage points, depending on the specification. Since on average less than 30 percent of victims report being entirely or partially satisfied, these changes represent a greater than 50 percent increase in satisfaction.

Unfortunately, the crime victim surveys are not very well suited to estimating the effects of the decoy survey program. Only 149 crime victims were, like the decoy surveyors, males in the age range of 20 to 60 who personally reported the crime to police, and of these only 62 reported crime types that decoys reported (burglary, theft, and sexual harassment). Nonetheless, the analysis performed in this very limited sample is quite suggestive: in the fixed effect specification, each extra decoy visit performed before the crime reportedly happened increased crime satisfaction by 0.18 ($p > 0.08$) in the specification without control for types of crime. This increases to 0.23 ($p > 0.15$) in the specification with control for type of crime—more than doubling baseline satisfaction rates (online Appendix Table A4).

The other interventions do not appear to have been as robustly effective. There is a 14 percent to 28 percent increase in satisfaction attributable to the freezing of transfers (which took place in all the police stations with any other intervention). This average effect is largely or completely undone by combining it with any of the other two interventions (community observer and weekly day off/duty rotation) or with both of them together.

The effect of the training might operate through two different mechanisms: first, it may affect public perception by changing what the police do, for instance

TABLE 8—EFFECT OF TRAINING ON POLICE ACTIONS

	Registered case (1)	Asked for bribes (2)	Made arrest (3)	Recovered property (4)	Continued investigation actions (5)	Recontacted victim (6)	Monthly number of cases registered (7)
Percentage staff trained	0.176 (0.109)	0.183 (0.113)	0.137 (0.0868)	0.0377 (0.0208)	0.160 (0.101)	0.0683 (0.0773)	1.684 (0.976)
Observations	571	571	571	571	571	571	3,852
R^2	0.139	0.107	0.119	0.137	0.087	0.089	0.034
Station fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Crime and victim controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline/control mean	0.590	0.160	0.135	0.0123	0.246	0.135	17.91

Notes: Standard errors in parentheses clustered by police station. Columns 1–6 use data from crime victim surveys. Column 7 uses data from police administrative records. Victim characteristics: age and gender of the respondent, education and dummies for the occupation of the head of household, indicators for caste or Muslim religion, and indicators for motorcycle ownership.

TABLE 9—CASE REVIEW OUTCOMES

	Quality of field investigation (1)	Quality of documentation (2)	Used scientific techniques (3)
Investigating officer trained	0.152 (0.205)	0.203 (0.308)	0.0291 (0.0372)
Fraction officers trained in station	−0.0713 (0.310)	−0.251 (0.396)	−0.0000554 (0.0568)
Observations	937	924	981
R^2	0.286	0.327	0.194
Station fixed effects	Yes	Yes	Yes
Baseline/pure control mean	6.199	6.153	0.195

Notes: Standard errors in parentheses clustered by police station. The outcome variables in columns 1 and 2 are the scores, on a scale of 1–10, given to the case investigations by retired police officers. Column 3 reports a linear probability model of whether scientific methods were used in the case analysis, as determined by the retired officers. All regressions include controls for the type of case and the individual reviewer.

by improving their investigation skills, allowing for faster resolution of the case. Alternatively, the effect may be due to the soft skills training, which would manifest itself through the way in which the police treat victims and not their investigation actions. We examine these two alternative hypotheses in Table 8, Table 9, and column 3 of Table 7.

Table 8 presents the effects of the training on other police activities, such as registering cases, asking for bribes, making arrests, or recontacting the victim. We also add in column 7 variables collected from the official police statistics on the number of cases registered by month. The results suggest that most of the effects of the training operated through the behavior of the police with crime victims rather than their increased investigative success, though the large standard errors imply that we cannot rule out meaningful improvements in police performance. The only significant result relates to property recovery, for which the probability increases

by 4 percent. Consistent with the evidence from the decoy survey, we find suggestive evidence that police officers are more likely to register crime: a 18 percent increase with $p = 0.108$. The estimated impact of the training on the number of cases registered in the administrative records (column 7) is also significant at $p < 10$ percent, and suggests a 9.4 percent increase in case registration for stations with 100 percent of staff trained.

Controlling for the police action and success used as outcomes in Table 8 in the satisfaction regression in column 3 of Table 7 only slightly attenuates the coefficient on training, suggesting that other more subtle changes in attitude affect victim satisfaction. While for the lowest levels in the police hierarchy (the constables), the training was focused on soft skills, it also included a week of refresher training on investigations for the more senior officers. These results support the conclusion that it is the soft skill training that is likely to have had the most impact.

Because crime victims may not be well equipped to judge police competence or effort, we supplement the victim surveys with grades given by retired police officers to the case files of investigating officers. Table 9 reports the effects of training. We find no impact of either an officer's own training or the training of his fellow officers on the quality of case paperwork, the reviewer's evaluation of the officer's field investigation, or whether scientific investigation was used in the case.

These regressions impose the restriction that the effect of the fraction of police officers trained is linear, but our design allows us to relax this restriction, since the fraction of officers trained varied randomly from station to station. Figure 2 displays the coefficients on percentage of personnel trained from a regression like that in column 2 of Table 7, with the fraction of staff trained broken down by each randomization category. The effect of training 25 percent of the station staff is close to zero and insignificant. Only when 50 percent of the staff are trained do we see a 20 percent increase in victim satisfaction, after which the effects are roughly constant or even decreasing. These results suggest that police training, or at least police training intended to change attitudes and treatment of the public, must be given to at least half the staff in order to take root.

Public Perception of Police.—Finally, we investigate whether the interventions had an impact on the more general public perception of the police. Given that very few people actually interact with the police in any capacity, it appears *ex ante* unlikely that we would find results on this dimension, but Table 10 directly investigates this question.

Since the public opinion survey covered a broad range of qualitative indicators, the 14 major questions on police performance are aggregated into four broad categories in order to simplify the interpretation, make sure all questions asked are covered, and guard against false positive results. The categories were: first, responsiveness of police to citizens; second, fear of police; third, corruption; and fourth, adequacy of police resources.¹⁰ The group outcomes are then analyzed following

¹⁰The questions incorporated into the categories were as follows: A. Responsiveness of police to citizens: "How do the police behave with normal citizens?" "Do the police help citizens when required?" "How quick is the police response to distress calls by citizens?" B. Fear of police: "Do you think that citizens like you are afraid of the police?"

TABLE 10—PUBLIC OPINION

	Only endline (1)	Police station fixed effects (2)
<i>Panel A. Police responsiveness to citizens</i>		
All interventions	−0.088 (0.04)	−0.024 (0.07)
No transfer	−0.103 (0.05)	−0.014 (0.08)
Weekly day off	−0.051 (0.05)	0.119 (0.08)
Community observer	−0.083 (0.05)	0.066 (0.09)
Percentage staff trained	−0.03 (0.04)	−0.054 (0.07)
Not pure control	0.112 (0.05)	
<i>Panel B. Reducing police corruption</i>		
All interventions	0.096 (0.05)	0.07 (0.09)
No transfer	0.03 (0.05)	0.02 (0.08)
Weekly day off	−0.028 (0.05)	−0.028 (0.09)
Community observer	0.022 (0.05)	0.063 (0.08)
Percentage staff trained	0.01 (0.05)	0.016 (0.07)
Not pure control	0.078 (0.07)	
<i>Panel C. Lack of fear of police</i>		
All interventions	0.144 (0.07)	0.209 (0.11)
No transfer	0.132 (0.06)	0.242 (0.10)
Weekly day off	−0.043 (0.06)	0.026 (0.10)
Community observer	0.106 (0.06)	0.178 (0.11)
Percentage staff trained	−0.053 (0.06)	−0.003 (0.08)
Not pure control	0.075 (0.08)	
<i>Panel D. Adequacy of police resources</i>		
All interventions	−0.011 (0.06)	−0.017 (0.06)
No transfer	−0.047 (0.06)	−0.034 (0.06)
Weekly day off	−0.063 (0.06)	−0.071 (0.06)
Community observer	−0.077 (0.07)	−0.066 (0.07)
Percentage staff trained	−0.027 (0.05)	−0.032 (0.05)
Not pure control	0.13 (0.08)	
Police station fixed effects	No	Yes
Household characteristics, victim/arrested, opinion source	Yes	Yes

Notes: Standard errors in parentheses are clustered at the police station level. Treatment variables are defined as described in Section VC. Control variables include age and gender of the respondent, household head's education, indicators for occupation, caste, religion, and motorcycle or vehicle ownership. "Victim/arrested" is an indicator for whether the respondent was arrested, knows someone arrested, or is in a household with a crime victim. "Opinion source" is an indicator for sources of information on the police; nonmutually exclusive categories such as TV news, word of mouth, personal experience, etc.

the approach outlined in Kling, Liebman, and Katz (2007). We first normalize the values of each outcome by the variance of the control group and jointly estimate all regressions in the group. The group outcome is, then, the mean of the normalized individual outcome coefficients, with the standard errors of the mean coefficient

"Are law-abiding citizens afraid of the police?" and when asked how the population thinks of the police, replies that they "fear them." C. Corruption: "Would you say that the police in your area are generally honest or generally corrupt?" "Is it necessary to pay the police some money in order to get them to do their job?" and "Do policemen themselves violate the law than the average citizen?" D. Adequacy of police resources: "Do the police have enough personnel to do the work required of them?" "Do the police have enough money and resources to do the work required of them?" "Should the size of the police force be increased, decreased, or stay the same?" and "Do you think that the government should spend more money on the police, even if it means spending less on things like education and roads?"

incorporating the variances and covariances of the individual coefficients. The specification is identical to the specification in Table 7.

Table 10 displays the effect of each reform on each category, with results in column 1 limited to the endline data, and those in column 2 including the baseline data with police station-level fixed effects. The results are expressed in standardized effect size, with standard errors in parentheses. This table indicates that the perceptions of the general public were not particularly affected. Of the four categories, the project has significant positive effect across both specifications only in one, the “Fear of police” category, and for one intervention—the freezing of transfers. The freezing of transfers seems to reduce the fear of police among ordinary citizens. But this effect is undone in stations that also had the weekly day off/duty rotation intervention and, given the limited implementation of this intervention, it seems a little surprising that it has such a robust effect on its own.¹¹ The most likely interpretation, assuming that this is not just a statistical accident, given the lack of familiarity with police that the baseline data demonstrate, is that less frequent transfers allowed the public to become more familiar with the same police staff and hence came to trust them more and fear them less. It is also possible that once the police staff remained in a post for a longer period of time, their behavior changed with respect to the inhabitants of that area and they became less intimidating to the population and more familiar with their beat. This could also help explain why the duty rotation undid the positive effect of the no transfer policy: duty rotation made it impossible to implement the beat system, as it results in duty officers having no attachment to a particular position or neighborhood.¹²

IV. Conclusion

The experimental results in this paper show that it is possible to affect the behavior of the police in a relatively short period of time, using a simple and affordable set of interventions. Decoy visits were effective in getting the police to register more crimes and treat victims more politely. Training police staff in investigation techniques and public relations skills (soft skills) also increased the satisfaction of crime victims. These successes were achieved with no change in existing managers, and virtually no field-level involvement by researchers or consultants. On the other hand, some other *prima facie* plausible interventions were not properly implemented and turned out to be ineffective.

What does this tell us about the nature of organizational change? It is striking that the interventions that failed to work were all carefully selected by the police leadership, partly based on the recommendations of various Police Reform Commissions, who among them combine a huge amount of experience and expertise. These senior police officers also did not lack human capital, being selected through an

¹¹ Online Appendix Table A9 unpacks the effect on fear of police by examining the effects of all interventions on the individual questions that constitute the fear of police outcome separately. While the signs of all coefficients are the same across questions, the effect seems to be driven primarily by the single question that asks respondents whether they believe that “law-abiding citizens” in general fear the police.

¹² It is much less clear why the day off would undo the positive effect of transfer. Given that the implementation of the day off policy quickly diminished over time, the negative interaction is thus more likely to be due to the duty roster.

extraordinarily competitive set of exams. Yet they underappreciated the difficulty of implementing these interventions even with their full backing, suggesting that they did not fully realize the nature of the informal authority enjoyed by the station chiefs.¹³ In contrast, the decoy intervention had never been suggested by Police Reform Commissions or mentioned in the preproject discussions we had with the senior state-level officers. Although it did not specify any explicit punishment, this was clearly a monitoring intervention and raised the possibility, contrary to some of the previous literature, that (well-chosen and well-implemented) incentives can be made to play a positive role even in a government bureaucracy.

While it is not possible to rule out the possibility is that bad luck played a role, we see the relatively poor performance of police force interventions as evidence against the view that the managers know how to translate the general principles of management into solutions that are relevant for their organization. The managers in this case were the elite of the Indian police, with decades of experience and who were motivated enough, among other things, to engage with us to launch this reform program. They had a clear and articulated understanding of the principles that motivated the interventions (“unpredictable work schedules are demoralizing,” “a chance to learn on the job can motivate an employee,” “the customer must be listened to”) and they were recognized as outstanding leaders—both the Director General A.S. Gill and our coauthor Nina Singh were subsequently assigned to important positions in the central (federal) police. And while they clearly shared the view that incentives, as a general principle, are important, they could not see a way to introduce them within the political and administrative constraint. These challenges are surely not limited to the police, or even the government sector—Gibbons and Henderson (2012) document the same difficulties in private corporations with competent, profit-oriented managers.

What perspectives do outsiders with less institutional knowledge and experience provide in these situations? While we cannot directly test this question, we suggest three reasons why these insights were unlikely to have arisen within the police institution. First, the need to introduce individual incentives for subordinate staff was an implicit acknowledgment that the traditional police hierarchy and discipline might be inadequate to obtain satisfactory program implementation. Outsiders with little stake in the existing police power structure were much more willing to consider and advocate interventions that conceded and accommodated the autonomy of middle managers.

Second, the evaluation generated evidence and information not typically available to the police leadership. The police’s own internal monitoring systems, in which subordinates report on their own performance to superiors, typically confirms that implementation is exemplary. Even in the case of the two least successful interventions, official records claimed almost complete compliance. Without external monitoring there would have been no concrete evidence that the basic problem lay not in the design of some of interventions but rather in the incentives for their execution.

¹³ Banerjee, Duflo, and Glennerster (2008) and Dhaliwal and Hanna (2017) are two other examples of how middle-level managers in government bureaucracies undermine programs that limit their discretionary authority. In a private sector context, Atkin et al. (2017) show that a productivity-enhancing innovation was not adopted in firms in Pakistan because it was successfully blocked by cutters and printers who stood to lose from its adoption.

Finally, it seems that the initial underestimation of the importance of incentives and soft skills lay not so much in the lack of information by senior officers, but rather with the view that these programs would not work in that context. (This is similar to the interventions that the management in BEMMR knew about, but did not think would work for them.) In a police system based on the themes of discipline, hierarchy, and duty, the idea of individual incentives, particularly financial ones, has become associated with corruption and the dilution of intrinsic motivation and discipline. Similarly, the Rajasthan Police themselves acknowledged that “soft skills” and positive public relations were not part of their institutional heritage. This explains the decision to completely outsource the soft skills training to consultants rather than do so through police trainers or even a training-of-trainers approach.

This interpretation may help explain the mixed results in Rasul and Rogger (2018). Perhaps the efforts to implement incentives in Nigerian government bureaucracies were sufficiently poorly designed or implemented to backfire; it does not preclude the possibility that consultants would have been able to design them well. The view we end up with falls somewhere in between the “best principle” and “contingencies” view of management. Organizations, both private and public, can be improved by helping managers put into practice some fundamentally sound principles, but the evidence so far suggests that this requires quite careful fine-tuning.

REFERENCES

- Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen.** 2017. “Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan.” *Quarterly Journal of Economics* 132 (3): 1101–64.
- Baicker, Katherine, and Mireille Jacobson.** 2007. “Finders Keepers: Forfeiture Laws, Policing Incentives, and Local Budgets.” *Journal of Public Economics* 91 (11–12): 2113–36.
- Banerjee, Abhijit, Raghavendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh.** 2011. “Rajasthan Police Performance.” Harvard Dataverse. <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/OC35YC> (accessed November 16, 2020).
- Banerjee, Abhijit, Raghavendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh.** 2021. “Replication data for: Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy, and Training.” American Economic Association [publisher], Inter-university Consortium on Political and Social Research [distributor]. <https://doi.org/10.3886/E116807V1>.
- Banerjee, Abhijit, Esther Duflo, Raghavendra Chattopadhyay, Daniel Keniston, and Nina Singh.** 2016. *Police Performance and Public Perception in Rajasthan, India*. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.
- Banerjee, Abhijit V., Esther Duflo, and Rachel Glennerster.** 2008. “Putting a Band Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System.” *Journal of the European Economic Association* 6 (2–3): 487–500.
- Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts.** 2013. “Does Management Matter? Evidence from India.” *Quarterly Journal of Economics* 128 (1): 1–51.
- Bloom, Nicholas, Renata Lemos, Raffaella Sadun, and John Van Reenen.** 2015. “Does Management Matter in Schools?” *Economic Journal* 125 (584): 647–74.
- Bloom, Nicholas, Carol Propper, Stephan Seiler, and John Van Reenen.** 2010. “The Impact of Competition on Management Quality: Evidence from Public Hospitals.” NBER Working Paper 16032.
- Bloom, Nicholas, and John Van Reenen.** 2007. “Measuring and Explaining Management Practices across Firms and Countries.” *Quarterly Journal of Economics* 122 (4): 1351–1408.
- Brandt, Steven G., and Frank Horvath.** 1991. “Crime-Victim Evaluation of Police Investigative Performance.” *Journal of Criminal Justice* 19 (2): 109–21.
- Bruce, David, and Rachel Neild.** 2005. *The Police that We Want: A Handbook for Oversight of Police in South Africa*. Johannesburg: Center for the Study of Violence and Reconciliation.

- Burgess, Simon, Carol Propper, Marisa Ratto, and Emma Tominey.** 2017. "Incentives in the Public Sector: Evidence from a Government Agency." *Economic Journal* 127 (605): F117–41.
- Caparini, Marina, and Otwin Marenin.** 2005. "Crime, Insecurity and Police Reform in Post-Socialist CEE." *Journal of Power Institutions in Post-Soviet Societies* 2: 1–12.
- Centre for Media Studies.** 2005. *India Corruption Study to Improve Governance: Corruption in the Police Department*, Vol. 9. New Delhi: Transparency International. <https://ipc498a.files.wordpress.com/2008/03/ti-india-police-corruption-study-2005.pdf>.
- Dhaliwal, Iqbal, and Rema Hanna.** 2017. "The Devil is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Developmental Economics* 124: 1–21.
- Di Tella, Rafael, and Ernesto Schargrofsky.** 2003. "The Role of Wages and Auditing during a Crack-down on Corruption in the City of Buenos Aires." *Journal of Law and Economics* 46 (1): 269–92.
- Ekins, Emily.** 2016. *Policing in America: Understanding Public Attitudes toward the Police Results from a National Survey*. Washington, DC: Cato Institute.
- Gibbons, Robert, and Rebecca Henderson.** 2012. "Relational Contracts and Organizational Capabilities." *Organization Science* 23 (5): 1350–64.
- Gramckow, Heike P., Jack Greene, Ineke Marshall, and Lisa Barão.** 2016. *Addressing the Enforcement Gap to Counter Crime: Investing in Public Safety, the Rule of Law and Local Development in Poor Neighborhoods, Part 2: Options for World Bank Engagement with Police*. Washington, DC: World Bank.
- Holmstrom, Bengt, and Paul Milgrom.** 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7: 24–52.
- Hope, Kempe Ronald, Sr.** 2015. *Police Corruption and Police Reforms in Developing Societies*. Boca Raton, FL: Taylor and Francis Group.
- Iyer, Lakshmi, Anandi Mani, Prachi Mishra, and Petia Topalova.** 2012. "The Power of Political Voice: Women's Political Representation and Crime in India." *American Economic Journal: Applied Economics* 4 (4): 165–93.
- Kakachia, Kornely, and Liam O'Shea.** 2012. "Why Does Police Reform Appear to Have Been More Successful in Georgia than in Kyrgyzstan or Russia?" *Journal of Power Institutions in Post-Soviet Societies* (13): 1–17.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz.** 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.
- Kolsky, Elizabeth.** 2010. *Colonial Justice in British India: White Violence and the Rule of Law*. Cambridge, UK: Cambridge University Press.
- National Police Commission.** 1980. *National Police Commission Report*. Volumes 1–8. New Delhi: Ministry of Home Affairs, Government of India Press.
- Olken, Benjamin A.** 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–49.
- Opolot, James S.E.** 1992. "The Resilience of the British Colonial Police Legacies in East Africa, Southern Africa, and West Africa." *Police Studies: International Review of Police Development* 15 (90).
- Prendergast, Canice.** 2001. "Selection and Oversight in the Public Sector, with the Los Angeles Police Department as an Example." NBER Working Paper 8664.
- Rasul, Imran, and Daniel Rogger.** 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *Economic Journal* 128 (608): 413–46.
- Sarkar, Abhijit, Dripto Mukhopadhyay, Cheryl Blake, and Devika Prasad.** 2015. *Crime Victimization and Safety Perception: A Public Survey of Delhi and Mumbai*. New Delhi: Commonwealth Human Rights Initiative.
- Sidhu, Harpreet Singh.** 2003. "Management of Reforms in Police—A Study at District Level." Masters dissertation, Guru Nanak Dev University, Amritsar, Punjab.
- Tripsas, Mary, and Giovanni Gavetti.** 2000. "Capabilities, Cognition, and Inertia: Evidence from Digital Imaging." *Strategic Management Journal* 21 (10–11): 1147–61.
- van Dijk, Jan, Jan van Kesteren, and Paul Smit.** 2007. *Criminal Victimization in International Perspective*. Vienna: United Nations Office on Drugs and Crime.
- van Kesteren, Johannes Nicolaas, Pat Mayhew, and Paul Nieuwebeerta.** 2000. *Criminal Victimization in Seventeen Industrialised Countries: Key-Findings from the 2000 International Crime Victims Survey*. The Hague: Ministry of Justice.
- Varona, Glenn.** 2010. "Politics and Policing in the Philippines: Challenges to Police Reform." *Flinders Journal of History and Politics* 26: 101–25.
- Wade, Robert.** 1985. "The Market for Public Office: Why the Indian State is Not Better at Development." *World Development* 13 (4): 467–97.

This article has been cited by:

1. Mitra Akhtari, Diana Moreira, Laura Trucco. 2022. Political Turnover, Bureaucratic Turnover, and the Quality of Public Services. *American Economic Review* **112**:2, 442-493. [[Abstract](#)] [[View PDF article](#)] [[PDF with links](#)]