

Mentoring and Incentives

Arpita Saluja*

December 2021

Latest version available [here](#)

Abstract

Managers play an important role in the human capital development of junior employees through mentoring initiatives and on-the-job training programs. This paper examines the trade-off managers face in allocating their time between two competing tasks: production of own output and mentoring of juniors. Using a stylized model of multitasking and data of the elite civil servants in India, I find that mentors who are in closer proximity to their supervisors and more easily monitored exert relatively less effort in mentoring and more on own output production. These differences in mentoring translates into long-term differences in productivity and effectiveness of juniors bureaucrats and their human capital accumulation. Finally, I show that a country-wide policy reform that was introduced to bring greater transparency to the performance evaluation process and increase the productivity of the seniors distorts the incentives to mentor the junior employees.

*Department of Economics, Columbia University. Email: as4943@columbia.edu
I am grateful to my advisors Michael Best, Jonas Hjort and Suresh Naidu for their input and guidance. I am thankful to Sakshi Gupta, Vinayak Iyer, Tam Mai, Haaris Mateen, Hailey Nguyen and participants of Columbia's Development and Applied Micro Colloquium for helpful comments and feedback. All errors are my own.

1 Introduction

Training and mentoring programs are ubiquitous – medical professionals, police agents, teachers and many public sector employees complete intensive hours of training before they can start their jobs. Similar is the case for bureaucrats¹ who are required undertake training programs that can last from six months to up to two years as a part of their induction (Mackenzie, 1979; Kuhlmann and Wollmann, 2019). Human capital accumulation in the workplace increases productivity, wages and mobility (Becker, 1962). These training programs can especially be important when the newly recruited personnel do not have specific expertise for the job. Indeed, research shows that the largest increases in a worker’s human capital happen at the start of their career (Shaw and Lazear, 2008; Haggag et al., 2017).

A large body of literature on training has been focused on understanding the incentives that drive firms to train its workers (Katz and Ziderman, 1990; Acemoglu and Pischke, 1998; Loewenstein and Spletzer, 1998). However, there is little research that examines the role of mentors or managers in determining the efficacy of training programs. Managers in most organizations perform multi-dimensional roles and they must divide their time between several competing tasks. I examine the effect of this trade-off on training using data from the newly recruited civil servants in the Indian Administrative Services (IAS henceforth). Although specific, it is an ideal setting to study this question as the issues of multi-tasking and perverse incentives are extremely salient in bureaucracies. Furthermore, as governments all over the world continue to allocate increasingly large sums of money to their training budgets and debate on the need to reform their training curriculum, this paper provides the much needed evidence on the returns to such training programs.

This paper studies the training and mentoring program of the Indian Administrative Services officers. These are the top-tier civil servants in India involved in major policy decisions. Every year, junior bureaucrats in the IAS are required to undertake a training in district administration under district heads for a period of one year. These district heads are senior bureaucrats that not only have to mentor junior officers but are also in-charge of the overall development of district. I examine the trade-off that the senior mentors face in allocating their time between district development activities and training of the junior officers. The agency problem that arises gets further exacerbated as the bureaucratic

¹Some countries with long induction training programs for bureaucrats are: Germany, India, Pakistan, Bangladesh, France, Singapore, Taiwan, etc

output is imperfectly observed. I find that district heads choose to allocate greater time on their own output production and less time on mentoring when they are located at a closer proximity to their supervisors and can be easily monitored. These differences in early career mentoring of junior bureaucrats has long-term implications on their productivity. Juniors that receive less mentoring are more likely to be suspended, and less likely to take up training programs in future. Exploiting a country-wide policy reform that created incentives for better monitoring of mentors that were located farther away, I document that after the reform gets enacted, differences in mentoring and output production between farther and closer districts decrease. My findings show that the effectiveness of on-the-job training and mentoring programs depend on the incentives of the seniors who act as mentors. I also find that incentives aimed at improving their productivity has a dynamic effect and is likely to come at the cost of decreased productivity of junior employees.

Entry into the IAS service is very competitive and takes through a rigorous entrance examination annually that is undertaken by over 2 million applicants. All recruits have to undergo a mandatory two year training at the national training academy. This training has two components – classroom instruction and on-the-field district training. I focus on the district-training component of the curriculum, which is a one year on-the-job training wherein new recruits get assigned to district administration. Each district is headed by a *district collector* who is a relatively more experienced bureaucrat² belonging to the IAS. Crucially, the district head acts as a formal mentor to the junior bureaucrat guiding him through the tasks that he is expected to undertake in the next decade of his career.

To illustrate how training efficacy can depend upon the incentives of the managers, I begin with a simple stylized framework based on the multi-tasking model of [Holmstrom and Milgrom \(1991\)](#) as used by [Muralidharan and Sundararaman \(2011\)](#) and [Neal \(2011\)](#). Managers can exert effort in two tasks – mentoring the juniors or producing their own output. Effort exerted in mentoring is unobserved and non-contractible. Managers are evaluated on their performance in output production. There is heterogeneity in how well this output can be observed and monitored by their supervising authority. Comparing between mentors, the model predicts that managers whose output is more accurately observed and measured, spend relatively less time on mentoring and more time on own output production. As a result, high powered incentives targeted towards own output production or changes in technology that allow for a more accurate measurement of managers' output

²On average an officer becomes a district head or collector on completion of 6 years of service. More senior officers occupy senior positions in ministries.

will have an adverse effect on the mentoring of junior employees.

The unique administrative governance setup creates a variation in visibility of the district head’s output that forms the basis of my empirical strategy. District heads report to a senior bureaucrat or group of bureaucrats³ who evaluate the performance of the former. The district where the supervising officer is located is called the headquarter district. When the supervising officer is located in the same district as the district head, output produced by the district head can be more easily observed and measured by the supervisor. Exploiting the “as good as random” assignment of junior bureaucrats to districts and using distance from headquarter as my primary variable of interest, I test if the mentoring effort and consequently human capital formation of the junior bureaucrat varies with distance from the headquarter.

I start my analysis by looking at the performance of junior bureaucrats in their district training. Performance in the district training is assessed by the national academy and is free from bias of the mentor. Consistent with the model’s prediction that mentoring effort increases with distance, I find that junior officers who get assigned to districts that are closer to the district headquarters show a relatively worse performance. Particularly, a 1% decrease in distance from the headquarters moves down the junior bureaucrat by 3 percentiles in the cohort ranking.

The next step is to understand if differences in mentoring also translate into differences in the long term performance of the junior officer. One big challenge of studying civil servants has been the lack of reliable individual performance measures. However, my long panel of work histories allows me to track the performance of the bureaucrats. I use suspension as a measure of bureaucratic performance as used in (Bertrand et al., 2019). Data on suspensions reveals that junior officers that start their career training under mentors assigned closer to headquarters are more likely to get suspended and face disciplinary action. Finally, I explore if the initial mentoring experience of the junior officers has any impact on their future human capital investment decisions. Iyer and Mani (2012) shows that such career investments are perceived as signals of higher competencies and are rewarded by the better future postings. I find that receiving less mentoring decreases the likelihood of taking up training in future. This effect is restricted to only those training that are voluntary and require the officer to sign up for them. No differences

³In certain states district heads report to a single officer called the Divisional Commissioner, where as in states where there are no divisions the district head reports directly to Board of Revenue located at the capital city

are observed in the take-up of mandatory training courses. Taken together, all these four pieces of evidence suggest that the effort exerted by the mentors during the training period can have a lasting impact on the careers of the junior officers.

In the second part of the paper, I rule out other alternative mechanisms that can explain the patterns that I observe. Distance, as interpreted from the lens of the model used, captures the differences in visibility of the mentor's output and hence incentives of the mentor. However, distance of a district from the headquarter is also correlated with other district characteristics – for instance: districts closer to the headquarter are more likely to be urban and populated– which can lead to differences in the performances of junior bureaucrats. In order to show that the effects I see emerge due to differences in mentoring, I proceed in two ways. For the first analysis, I shift focus from the junior officers to the district heads in my sample. While the effort exerted by district heads on mentoring is unobserved, effort exerted in producing own output can be measured. To measure how the effort of district heads on own output production varies with distance I leverage the projects level data from the MPLAD scheme in India. District heads are responsible for implementing key infrastructure projects in the district. Under the Member of Parliament Local Area Development Scheme (MPLADS), each members of parliament gets a fixed sum of money to spend on infrastructure projects in the district. These projects are reviewed and approved by the district heads. Under the legal mandate, the projects have to be sanctioned by the district heads within 75 days. I exploit this to create a measure of bureaucratic performance for the district heads. Restricting attention to the MPLADS projects that were proposed during the training period of the junior employees, I show that the likelihood of project getting sanctioned by the district head on time is higher in districts closer to the headquarters. The result is consistent with the prediction of the model that districts will exert more effort on own output when the output is more accurately observed.

In the second analysis, I exploit a performance evaluation reform that was introduced in 2006-07 with the objective of introducing transparency in the performance evaluation process. In practice, the reform created incentives for a better evaluation of performance of farther away districts. I find no differences in the performance of junior bureaucrats in the training in the post reform period. Furthermore, there are no differences in effort exerted by the district heads on their own output post reform . Both these findings are consistent with the fact that the reform worked as a high-powered incentive especially for districts that were farther away from the headquarter district.

In sum, this paper shows that training and mentoring matters for performance of junior bureaucrats. However, the bureaucrats responsible for mentoring the juniors face agency problems. While designing incentives aimed to improving the productivity and performance of bureaucrats, one must be cognizant of the fact such improvements in productivity may come at the cost of future productivity of bureaucrats.

This paper contributes to several strands of literature. First, it relates to the growing body of work on state effectiveness and capacity. The rigid rules of recruitment, promotion and firing provide a strong job security to public sector employees with no incentives to improve their productivity. A long standing literature has focused on understanding how to improve the performance of civil servants. Existing work has focused on the role of intrinsic motivation, financial incentives (Dal Bó et al., 2013), recruitment and selection (Hanna and Wang, 2017; Moreira and Pérez, 2021) discretionary powers to supervisors Aman-Rana (2017) and autonomy (Bandiera et al., 2021). In this paper, I provide evidence one of the first pieces of evidence on the impact of training and mentoring on the performance of bureaucrats. By highlighting the role of training, the paper also complements the growing literature on “good” management practices in public sector organizations. (Rasul and Rogger, 2017; Banerjee et al., 2021; Sandvik et al., 2021; Fenizia, 2019).

Second, my work contributes to the organizational economics literature on incentive design in hierarchical organizations (Holmstrom and Milgrom, 1991; Dewatripont et al., 1999). One key result of in this seminal work is that when the principal cannot monitor some aspects of the agent’s performance, then high powered incentives cause agents to shift their effort from poorly measured activities to the better-measured and well-compensated activities. In this paper, I show that when senior bureaucrats are also responsible for training or mentoring the junior bureaucrats, high powered incentives can far reaching effect on the inter-generational human capital accumulation. Finally, this paper speaks to the literature on labor economics by examining when why on-the-job training works or not. I highlight the role that managers play in determining the efficacy of training programs. Furthermore, this paper looks at training in public sector while most of literature has been focused on learning within a firm or in sectors such as education (Rodríguez-Planas, 2012), military (Kofoed et al., 2019) or front line bureaucrats such as healthcare workers (Bartel et al., 2014).

The remainder of this paper is organized as follows. Section 2 describes the institutional background of bureaucrats in India, particularly focusing on training. Section 3 details the various data sources and sample used. Section 4 outlines the conceptual framework

that explains why incentives to mentor varies by distance to the headquarter. Section 5 describes the empirical strategy. Main results are presented in Section 6. Section 7 discusses various possible mechanisms. I perform a battery of robustness checks in section 8 and section 9 concludes.

2 Background and Institutional Context

2.1 Indian Administrative Services

Indian Administrative Services (or IAS) along with Indian Foreign Services (IFS) and Indian Police Services (IPS) constitute the most elite officers of the Indian bureaucratic machinery. The bureaucrats belonging to the IAS are responsible for the formulation and implementation of all major government policies. These officers begin their career with a field posting in the district’s revenue department. Specifically, they get appointed as an *assistant collector* as a part of their training followed by the posting of sub-divisional magistrate. As they gain more experience, they go on to occupy key positions in the various government departments such as finance and taxation, health, water resources, power etc. as well lead the charge of large state-owned enterprises. A typical IAS officer’s career spans 30-35 years of service. The senior-most officers have considerable influence and they hold critical posts such as that of principal advisors to the Prime Minister of the country or Chief Ministers of the various states.

As is the case with most bureaucrats, promotions or changes in career level follow a tenure-based progression rule. Officers’ are promoted when they complete 9, 13, 16, 25 years of service. Appendix Figure 6 shows there is not much variation in the salary scales of the officers with most promotions being closely aligned with what is stipulated by the rule. (Bertrand et al. (2019)).

2.2 Training and Mentoring

Since the focus of this paper is on the mentoring received by the junior bureaucrats when they get assigned to the district training, it is important to distinguish the various components of the training. This section explains each component in detail.

After getting recruited, junior officers undergo an extensive training program that is spread over a period of 24 months. Most of the training happens at the Lal Bahadur Shastri National Academy of Administration (LBSNAA) and follows a “sandwich” pattern– which

was introduced in 1969— with institutional training being interspersed by on-the-job field training (Agarwal (2014)). Officers from all the civil services (IAS, IFS and IPS) recruited in a given year start their training together with the *Foundation Course*. This is a 15 week classroom instruction based training focusing on six disciplines- Public Administration, Management and Behavioural Sciences, Indian History & Culture, Economics, Law, and Political Concepts and Constitution. Officers’ performance is evaluated through end-of course examinations.

This is followed by a 26 week training schedule called *Phase I training*. The curriculum focuses on equipping officers with the theoretical knowledge of the responsibilities they are expected to undertake in the first decade of their career. This includes district administration, election duties, disaster management, urban development, and so on. In addition, officers get trained and evaluated on the proficiency in the regional language of their state cadre. As before, the performance of bureaucrats gets evaluated through course-end examinations. The District Administration Training for IAS has its historical roots in tradition of hosting young officers from Britain that were recruited into the ICS, the colonial predecessor of the IAS. In its current form, newly recruited officers are attached to the district head for a period of one year. After the training ends, the officers are required to submit a detailed empirical analysis of a social issue plaguing the district that they were assigned to. The final report is evaluated by the academy.

2.3 District Administration and distance from Headquarters

One of the most important appointments of an IAS officer is that of district head, also known as *District Collector*. District Collector are responsible for the overall development of the district and act as the implementation authority of various programs. Other roles and responsibilities include land revenue administration, disaster management during natural calamities such as floods, famines or epidemics, supervise and coordinate election in the district, among others. The way the hierarchy is structured, district heads reports to a supervising authority who is typically an officer of higher seniority. In most states, this supervising authority is an officer at the post of the *Divisional Commissioner*. This posting was created historically to reduce the span of control from the State headquarters and lead to facilitate greater coordination between districts. For this purpose, a cluster of districts are grouped together to create a *division*. For instance, Figure 3 shows the map of one of largest states in India, Uttar Pradesh, which is divided into 18 divisions for ad-

ministrative purposes. Each division comprises of on average of 4 districts. The divisional commissioner is assigned to one of districts in the division. This district is called the *head-quarter district* (Figure 2). Thus, the district head as well his supervising authority are located in the same location when the district is an headquarter district. This variation in distance of the district head from his supervisor arising out of the decentralized structure of administration is the key variation that I exploit. District heads are more clearly observed and easily monitored lower is the distance of the district from the headquarter. The basis for this assumption is that geographical proximity facilitates greater monitoring and access to information. In a handful of states⁴, district heads report directly to the state headquarters and there is no intermediate post of divisional commissioner. For the purpose of my analysis I don't differentiate between the two.

2.4 Performance Appraisal

An officer's performance is assessed and recorded annually following a multi-tiered review process. Performance is assessed primarily by the immediate supervisor of the appraised officer called the *reporting authority*. Once the reporting officer records his evaluation, the report is passed on to the *reviewing* and *accepting* authority for their validation. Until 1985, the evaluation process was simple and recommended and IAS officer to be "fit", "unfit" or "not yet fit" for promotion. Thereafter, quantifiable targets were introduced and officers were required to submit a self-assessment of their performance. Furthermore, the scale to evaluate the performance was changed to a 5-point system ranging from "Outstanding" to "Below Average". More reforms were introduced in 2007 with the objective of introducing greater transparency in the performance appraisal⁵. One big departure from the old process was the move to disclose the performance grade to the assessed officer. Prior to 2007, the performance report was kept confidential and only "adverse" remarks were communicated to the officer. Following the landmark judicial case of "Dev Dutt Vs Union Of India", the final report is now required to be revealed to the appraised officer who is empowered to appeal against any incorrect factual observations made by the reporting and reviewing authority. In case of disagreement, the officers gets an opportunity to get the grade amended or adverse remarks expunged by submitting to a representation to a referral board.

⁴Tamil Nadu, Andhra Pradesh, Kerala and Gujarat

⁵The guidelines governing the process are The All India Services (Performance Appraisal Report) Rules 2007

3 Data

My analysis combines data from several sources. In this section I describe these sources of data used in the final sample.

1. **Civil Lists:** Published annually, civil list is a roster of bureaucrats serving each state at the beginning of year. These lists record a snapshot of all the details of the employees including their current posting, date of joining that post, salary scales, etc. I identify the first assignment of the junior bureaucrat using these rosters. In years where this information is missing I use the executive record sheets of the IAS officers, details of which are given below. To create this data, I undertook a massive exercise of digitizing the annual civil lists from years 1996 to 2021.
2. **Executive Record Sheets:** The Department of Personnel & Training publishes the executive record sheets of all the officers that have served the Indian Administrative Service. These records are essentially resumes of the officers that are publicly available. I use these records to create a panel of various training courses attended by the officer and obtain the rich demographics details such as gender, date of birth, education qualification, etc.
3. **Inter-se seniority lists:** In order to measure the performance of the junior officers in their district training, I obtain the inter-se seniority lists for all the officers admitted into the civil services. This data contains raw scores of officers in all the training courses undertaken at the Lal Bahadur Academy. This data is publicly available for the officers admitted post 2004. For earlier years, I obtain this information through multiple government agencies under the Right to Information Act (RTI).
4. **MPLADS data:** To measure bureaucratic performance I leverage the project level data from the MPLADS. Started in 1993, the scheme provides each Member of Parliament with a fixed sum of money to be spent on public works in their constituency. Crucially, the projects are implemented at the level of district and are required to be sanctioned by the district head. As this data is no longer available for years before 2016, I use the replication files provided by [Bohlken \(2018\)](#) and [Rivera \(2020\)](#) to create this data. The universe of data comprises of around 35000 projects that were proposed between the period 1999 and 2014.

5. **Other administrative data sources:** I create an annual panel data of officers serving as district collectors or district heads for each district. These collectors are the assigned mentors of the junior officers. Information about the district characteristics come from the latest census available at the time of district assignment.

4 Conceptual Framework

This section illustrates the simple framework that formalizes the mechanism guiding the empirical analysis. The model is based on the multi-tasking model of [Holmstrom and Milgrom \(1991\)](#) as used by [Muralidharan and Sundararaman \(2011\)](#) and [Neal \(2011\)](#).

Suppose that the district head can allocate time on two types of tasks T_1 and T_2 . The first task T_1 represents the district development activities undertaken by the district head and T_2 captures the activities undertaken by the district head to mentor the junior new recruit assigned to him. Let t_1 and t_2 be the time spent on each of these two tasks. Let $f(t_1)$ be the payoff that the district head gets from tasks related to public service delivery. Even though the pay structure of bureaucrats are fixed, a good performance can translate into non pecuniary benefits like a transfer to a better location or posting to a more lucrative department. Additionally, the model assumes that the output produced by the district head is not perfectly observed by the supervisor. Let $1 - \alpha$ capture how imperfectly the district output is observed. $g(t_2)$ is the payoff that the district head gets from mentoring the juniors. This represents the warm glow utility from mentoring or the present discounted value of return from investing in a relationship with juniors. This results in the district head or mentor's maximization problem as follows:

$$\max_{t_1, t_2} \alpha f(t_1) + g(t_2) \quad \text{st } t_1 + t_2 = T$$

Optimization gives:

$$\alpha f'(t_1) = g'(t_2)$$

Comparing between district heads we find that an increase in α should result in decline in t_2 . This is the standard prediction of model of [Holmstrom and Milgrom \(1991\)](#). It is pretty standard to assume that α varies with distance as geographical proximity facilitates greater monitoring and access to information ([Degryse and Ongena \(2005\)](#); [Kang and Kim \(2008\)](#); [Giroud \(2013\)](#)). In our setup what this means is that districts that are located farther away from the headquarter would have district heads allocating more time on

mentoring and less time on district development tasks. Further, changes in visibility of performance due to better technology, etc will induce district heads to reduce mentoring and shift effort towards district development. Figure 4 graphically shows the prediction of the model.

5 Empirical Strategy

5.1 Estimating Equation

The theoretical model predicts that the district head’s incentives to mentor the junior bureaucrat varies with the distance of the district from the headquarter. To estimate the effect of distance on the outcomes of the junior bureaucrats, I run the following baseline specification:

$$(1) \quad Y_{b(si)mdt} = \beta \log(1 + Dis_{dt}) + \gamma X_{b(si)} + \eta Z_{mdt} + \delta W_{dt} + \alpha_{si} + \epsilon_{b(si)mdt}$$

The outcome variable $Y_{b(si)mdt}$ is some measure of performance of the junior bureaucrat b who is recruited into the IAS in year i and allocated to state cadre s . d denotes the district where the officer gets assigned for the one year district training at time period t ⁶. As described in Section 2.3, assignment to a district automatically entails assignment to a mentor, m denotes the district head who is the mentor of the junior bureaucrat. $\log(1 + Dis_{dt})$ is the logarithm of distance of district d from its headquarter district at time period t , where distance Dis_{dt} is measured in kilometers. The term $X_{b(si)}$ contains the set of all observable characteristics of the junior bureaucrat, including gender, age at entry into civil service, age squared, bureaucrat’s performance in the entrance exam, whether the bureaucrat has advanced educational degree. As performance in the training may depend on the social identity of the bureaucrat, I control for whether the bureaucrat belongs to SC, ST, OBC caste. W_{dt} controls for differences in district characteristics such as the proportion of rural households, proportion of households belonging to SC/ST population, logarithm of total population of district from the last available census, etc. I also control for the characteristics of the district head who mentors the junior bureaucrat. Specifically, I include the gender dummy, social caste of the mentor, experience of the mentor at time period t . As the goal is to compare the performance of junior bureaucrats that enter the civil service in the same year and are allocated to the same state cadre, all

⁶As the training calendar is fixed, t is always $i + 1$

specifications include state cadre X intake year fixed effects, α_{si} . Standard errors are also clustered at the state cadre - intake year level. β is the coefficient of interest measuring the effect of distance on performance of the junior bureaucrat.

5.2 Identification Strategy

The identification strategy relies on the “as good as random” assignment of junior bureaucrats⁷ to districts as used in Bhavnani and Lee (2021) and Bhavnani and Lee (2018). They document that junior officers’ assignment to districts are arbitrary with the precise allocation mechanism varying by states and years. Using the posting orders for four large state cadres (Andhra Pradesh, Karnataka, Rajasthan and Uttar Pradesh) the two papers show that state cadres seem to use heuristics such as alphabetical order and serial number – which are orthogonal to district and bureaucrats’ characteristics – while assigning junior officers to the districts. For instance, in 2013 Andhra Pradesh state cadre assigned officers by the alphabetical order of their names to districts that were ordered by their geographical position proceeding clock wise.

Use of such heuristics for initial assignments is unsurprising. While bureaucratic appointments are often politicized (Coviello and Gagliarducci (2017); Iyer and Mani (2012)) and officers get transferred to less desirable posts and locations as rewards or punishment for their performance (Khan et al. (2019)), such endogenous appointments are more common for experienced bureaucrats. Initial postings of bureaucrats have been found to be apolitical and plausibly random in many countries. (Rasul and Rogger (2017), Aman-Rana (2017)). Qualitative fieldwork done by Bhavnani and Lee (2021) corroborates similar findings for junior officers in India.

In this paper, I focus on the *first posting* of the bureaucrat where the “as good as” random assignment is even more credible. As described in Section I, officers are first randomly allocated to a state cadre using a rotating roster. They, then, get assigned to a training district and mentor while still training at the academy *before* joining their state cadre. The randomized allocation to the state cadre along with the timing of the posting makes it highly unlikely for the junior officer to be known to the politicians or senior bureaucrats⁸ responsible for bureaucratic appointments. Further, the fact that these bureaucrats are still on probation⁹ when they get assigned, alleviates remaining concerns

⁷Bhavnani and Lee (2021) define junior bureaucrats as those with tenure less than five years

⁸District postings are usually decided by very senior bureaucrats with more than 25 years of service

⁹Probationary officers are not confirmed into the services until they complete their training and

related to them wielding any power over their posting.

I formally test the random assignment assumption through a number of tests. First, I check that the officers that get assigned to farther away districts are not systematically different from the ones that are assigned to closer districts. Table 1 looks at individual characteristics of junior bureaucrats and quite reassuringly none of these can predict the distance of the assigned district from the headquarter. This is true for both pre- and post-reform period. The F statistic value for both the samples is quite small. Table 2 further verifies that none of the officers' characteristics can predict the characteristics of the districts where he gets assigned for her training for the pre-reform period. The results are similar for post reform sample. Finally, assignment to a district also means assignment to mentor. In Table 3, I evaluate how the characteristics of the junior bureaucrat vary with the characteristics of the mentors they get assigned to. As before, most of the coefficients are insignificant and the F statistic for joint significance is also insignificant. To sum, all randomization tests strongly support the key identifying assumption.

6 Main Results

In this section, I first present the results on the outcomes of the junior bureaucrats followed by outcomes of the district heads or mentors.

6.1 Immediate impact on performance

First, I look at the immediate impact of mentoring. As discussed in Section 2.2, junior bureaucrats are mentored by the district heads for a period of one year as part of their district training. Once the training period ends, they go back to the academy and submit an empirical study which is then evaluated. Thus, if mentoring effort varies with distance from the headquarter then the junior officers performance in the district training exam will reflect the same. Table 4 shows how junior officers performance in the district training varies with distance of the district from its respective headquarter using equation (1). The sample covered only includes officers that got recruited before 2007. Column (1) controls for only officers' characteristics, while column (2) and (3) successively add controls for district characteristics and mentor's characteristics. The coefficient is pretty stable across all specifications. Moving closer to the district headquarter by 1% pushes down the junior

successfully pass all the exams

officer in the cohort’s ranking by 3 percentiles. Figure 1 presents the graphical evidence of the relationship between logarithm of distance and percentiles of the officer in the district training scores. Consistent with the results from Table 4, I find that getting assigned to a farther away district is associated with better performance in the field training exam.

6.2 Long term outcomes

To explore the long term impact of mentoring, I leverage the work histories of the junior bureaucrats to create a balanced panel of long term career outcomes. The sample comprises of bureaucrats hired between 1996 and 2006 which implies that I can only examine outcomes up to 14 years in their careers¹⁰. Specifically, for each bureaucrat b of state cadre s , recruited in year i and trained in period t , I test:

$$(2) \quad Y_{b(si)mdtj} = \beta \log(1 + Dis_{dt}) + \gamma X_{b(si)} + \eta Z_{mdt} + \delta W_{dt} + \alpha_{si} + \eta_{sj} + \epsilon_{b(si)mdt}$$

$Y_{b(si)mdtj}$ is the outcome variable for the bureaucrat at experience level j . All notations are same as equation (1), except that now I also include state specific tenure fixed effects η_{sj} where j is the tenure or experience of the bureaucrat. Standard errors are clustered at state cadre X intake year level as well as at individual officer level. For all long term outcomes, results are reported for all officers recruited between 1996 and 2006, as well as for the sample restricted to years with information on training performance available.

6.2.1 Suspension

First, I use suspensions as direct measure of performance and investigate if the likelihood of getting suspended in future varies with the distance of the training district from the headquarter. Officers that face suspension have been found to be perceived as less effective, more corrupt, and less pro-poor by Bertrand et al. (2019) in their survey of subjective evaluations of IAS officers. Table 5 columns 1-3 report the estimates for all officers recruited between 1996 and 2006 and column 4 includes officers from restricted sample years. The negative coefficient shows that officers are more likely to face suspension and disciplinary action in any given year if they had their field training in districts closer to the headquarter. This is consistent with the hypothesis that these bureaucrats receive less

¹⁰Attrition is very small, less than 3% of officers, and there is no relationship between initial training location and attrition

mentoring from the district heads at the start of their career and hence are less effective. An important caveat to consider is that suspensions are extremely rare events – only 2% of all officers ever get suspended. Furthermore, suspension decisions are often contentious and politically motivated, hence one should interpret these results with caution.

6.2.2 Future Training Participation

Next, I examine the impact of initial mentoring on future human capital acquisition decisions. Such career investments are perceived as signals of higher competencies and are rewarded by the better future postings (Iyer and Mani, 2012). Every year a schedule for various training courses available to the officers is released by the Department of Personnel and Training. The schedule includes training courses that are mandatory for promotion (for instance, Mid Career Training), as well as many “in-service” training courses that interested officers can sign up for. I investigate if the lack of mentoring received during the district training deters bureaucrats from making such career investment in future. In terms of outcome variables, I look at both the extensive and intensive margin: the likelihood of undertaking training and total number of days spent in training in a given year.

Table 6 presents the estimates. Panel A Newly recruited officers who get trained and mentored in districts closer to the headquarter district are less likely to take up training in future on their own (Panel A). I also examine if this effect exists for training courses that are mandatory as it is a natural placebo outcome variable. Reassuringly, Panel B shows that there is no difference in the take-up of training that are mandatory for the officers to do. The results suggest that training experiences in early career can have long lasting effect on officers’ general attitude towards training.

7 Mechanisms

The results presented so far suggest that junior officers who are mentored in “farther” away districts show better performance. Next, I explore the various mechanisms that can potentially drive these results. Looking from the lens of the model, distance from head-quarter captures differences in information and monitoring of performance of the district heads. However, distance is likely to be correlated with other district characteristics that can independently generate the differences in performances of junior bureaucrats. Table

7 lists the correlation between $\log(1 + \text{distance})$ and various district characteristics. For both the pre- and post-2007 sample, I find that junior bureaucrats assigned to districts farther away from the headquarter are working in districts that are relatively more rural and less populated. It is not straightforward to predict the direction in which these associations may affect a junior bureaucrat’s performance. For instance, juniors working in a more rural and less populated location could potentially have less activities for leisure, and hence spend more time working on the empirical project that gets evaluated in district examination. This can explain why these officers do better in the training exam. Another possibility is that junior bureaucrats working in the headquarter district are likely show better performance as they can get more help and support as more bureaucrats work in the headquarter district. To rule out such competing explanations, I provide additional pieces of evidence below.

7.1 Performance of District Heads

I move the focus of the analysis from junior bureaucrats towards the district heads. While the effort exerted by a district head on mentoring is not directly observed, the output of the district head on district development activities can be measured. District heads are responsible for the overall development of the district, however, there are certain development programs that are solely controlled by the district head. One important example are the projects sanctioned under the Members of Parliament Local Area Development Scheme (MPLADS). Launched in 1993, the scheme provides each Member of Parliament with a fixed sum of money to be spent on public works in their constituency. Crucially, the projects are sanctioned by the district head. The time taken by the district head to sanction the project from the date it is proposed is a direct measure of the performance or output of the district head (Nath et al., 2016; Rivera, 2020; Bohlken, 2018; Bhavnani and Lee, 2021). In addition, projects are legally mandated to be sanctioned within 75 days of proposal. Therefore, I use two measures of output of district heads— time taken to approve the application and likelihood that the project is approved within 75 days. Since the objective is to compare the output of district heads while they are entrusted with the additional responsibility of mentoring the newly recruited bureaucrats, I restrict attention to the projects that are proposed in the time frame that coincides with the training period. I am able to map around 11000 projects to the district heads. I run the following baseline

regression:

$$(3) \quad Y_{pd(s)mt} = \beta \log(1 + Dis_{dt}) + \eta Z_{mdt} + \delta W_{dt} + \gamma_{st} + \epsilon_{pd(s)t}$$

For project p that is proposed within time frame t to be implemented district d located state cadre s and headed by district head m , $Y_{pd(s)mt}$ is the project's outcome variable. As before Dis_{dt} is the district's distance from the headquarter district and W_{dt} controls for district specific characteristics. As the outcome of the project is likely to be affected by the district head in-charge, Z_{mdt} controls for district head's experience level, age at entry into the IAS, gender and whether the officer was recruited through the competitive exam. γ_{st} fixed effects for In alternative specification, I also control for characteristics of the junior bureaucrat that the district head has to mentor as this may affect the amount of time the district head can allocate to district development activities. There is a lot of variation in time taken by the district head to approve the projects. Table ?? presents the results. Column (1) controls for district characteristics.

7.2 Performance Appraisal Reform

To disentangle the effect of district characteristics from the role of monitoring and information, I leverage the policy change in the performance appraisal process that happened in 2007. As described in Section 2.4, under the new system performance evaluation has become more transparent and timely. The policy effectively imposed a reputation cost on the supervisors of the district head creating incentives for them to evaluate the performance of their subordinates more accurately and on time. In my model, I interpret this as a change in α . Specifically, one expects the reform to have more effect on districts that were farther away from the headquarters. As predicted by the model, one should expect differences in mentoring to decrease post the reform. To test this empirically, I run the following specification similar to (Bertrand et al., 2019):

$$(4) \quad Y_{b(si)mdt} = \beta \log(1 + Dis_{dt}) + \theta \log(1 + Dis_{dt}) \times \text{Post Reform}_i + \\ \gamma X_{b(si)} + \eta Z_{mdt} + \delta W_{dt} + \alpha_{si} + \epsilon_{b(si)mdt}$$

Post Reform_i is a binary variable that takes value 1 if the intake year i of the bureaucrat is after the year 2007. All other variables are defined same as in equation (1). As before α_{si} are the state cadre X intake year fixed effects and standard errors are clustered at

the level of state cadre X intake year. Given how recent the reform is I can only look at the immediate impact of reform. Longer term outcomes such as performance of the junior bureaucrats as district heads is not possible as the average experience required to get such an appointment is 8 years. Nevertheless, I can explore how the reform impacted the performance of juniors on the district training scores and the district output of the mentors.

Table 8 reports the estimates from equation (4). Column (1) and (2) looks at the pre- and post-reform cohorts separately. Consistent with the model that the performance reform created incentives for the mentors to focus more on own output, the results show that after the reform there is no difference in performance of junior bureaucrats in their district training exam. Column (3) uses the entire sample from 1997 to 2012 and Column (4) interacts all the variables with the reform dummy. The results remain stable and robust across all the specifications. Graphical evidence in Figure 5 is also consistent with these estimates. Post 2007, the relationship between distance and performance becomes flat. Next, I investigate if the output of the district head changes post the reform. Following the analysis in Section 7.1, I re-estimate equation 3 by including a binary variable Post Reform_p which takes value 1 for projects that are proposed after 2007. Table 9 reports the coefficients. Column (1) includes only projects that were proposed in the pre-period, Column (2) uses projects that were proposed in the post-reform period and column (3) uses all projects together. The results suggest that after the reform, officers who were district heads of districts farther away from the headquarter get an incentive to allocate greater effort to own output production as their output is likely to be measured more precisely.

8 Robustness Checks

To ensure that I am capturing the effect of the reform, I perform a battery of robustness checks. First, I check that the reform did not have impact on outcomes that are not expected to be affected by the reform. I use the performance of the junior bureaucrats in training exams other than the district training to create placebo outcomes. Table 10 presents the results. The junior bureaucrats undertake four other exams as a part of their training at the academy. Coefficients for none of the exams are significant.

One threat to the interpretation of the results in Section 7.2 would be if the assignment process of junior officers to districts changed post 2007. Similarly, district assignments need

to be stable with respect to mentors' characteristics across the two periods. I test these assumptions directly. Table 11 examines if the junior officers assigned to farther away districts are characteristically different *after* the reform. Reassuringly, all except for one coefficient are insignificant. Table 12 repeats the analysis for the mentor's characteristics and finds similar results.

9 Conclusion

Mentoring of junior employees is one of the key ways through which human capital is augmented. In this paper, I look at the on-the-job training and mentoring of junior employees for bureaucrats in India. The unique institutional context allows me to exploit the exogenous variation in hiring and appointment of junior recruits to mentors which is very difficult to find in firms using observational data. To answer the question, I undertake a large scale digitization exercise to create a new source of data of locations where the new recruits undergo training for a period of one year. My main findings suggest the effectiveness of the mentoring programs depend primarily on the incentives of the managers or seniors who are responsible for several goals. Mentors choose to devote more time to their projects over mentoring when their output is more clearly observed and measures. This is an important finding as it highlights the need to look at the dynamic effect of an incentive policy. Many papers have looked at the role of perverse incentives that arise in multi-tasking settings. However, this paper shows that there could be other types of trade-off that have not been taken into account. Policy aimed at increasing the performance of economic agents must balance the tension between current productivity and future productivity. Finally, though the focus of this paper is on bureaucrats the results are quite general and relevant to several other sectors such as academic, medical professionals, military and police academy that have similar training programs.

References

- Acemoglu, D. and Pischke, J.-S. (1998). Why do firms train? theory and evidence. *The Quarterly journal of economics*, 113(1):79–119.
- Agarwal, K. (2014). Report of the committee to review the content and duration of induction training of ias officers. *DOPT Report*.
- Aman-Rana, S. (2017). Meritocracy in a Bureaucracy. *Working Paper*.
- Bandiera, O., Best, M. C., Khan, A. Q., and Prat, A. (2021). The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats*. *The Quarterly Journal of Economics*, 136(4):2195–2242.
- Banerjee, A., Chattopadhyay, R., Duflo, E., Keniston, D., and Singh, N. (2021). Improving police performance in rajasthan, india: Experimental evidence on incentives, managerial autonomy, and training. *American Economic Journal: Economic Policy*, 13(1):36–66.
- Bartel, A. P., Beaulieu, N. D., Phibbs, C. S., and Stone, P. W. (2014). Human capital and productivity in a team environment: Evidence from the healthcare sector. *American Economic Journal: Applied Economics*, 6(2):231–59.
- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49.
- Bertrand, M., Burgess, R., Chawla, A., and Xu, G. (2019). The Glittering Prizes: Career Incentives and Bureaucrat Performance. *The Review of Economic Studies*, 87(2):626–655.
- Bhavnani, R. R. and Lee, A. (2018). Local embeddedness and bureaucratic performance: Evidence from india. *The Journal of Politics*, 80(1):71–87.
- Bhavnani, R. R. and Lee, A. (2021). Does affirmative action worsen bureaucratic performance? evidence from the indian administrative service. *American Journal of Political Science*, 65(1):5–20.
- Bohlken, A. T. (2018). Targeting ordinary voters or political elites? why pork is distributed along partisan lines in india. *American Journal of Political Science*, 62(4):796–812.

- Coviello, D. and Gagliarducci, S. (2017). Tenure in office and public procurement. *American Economic Journal: Economic Policy*, 9(3):59–105.
- Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*. *The Quarterly Journal of Economics*, 128(3):1169–1218.
- Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *The Journal of Finance*, 60(1):231–266.
- Dewatripont, M., Jewitt, I., and Tirole, J. (1999). The economics of career concerns, part ii: Application to missions and accountability of government agencies. *The Review of Economic Studies*, 66(1):199–217.
- Fenizia, A. (2019). Managers and productivity in the public sector. Technical report, Technical report, Mimeo, Available at: <https://sites.google.com/site/...>
- Giroud, X. (2013). Proximity and investment: Evidence from plant-level data. *The Quarterly Journal of Economics*, 128(2):861–915.
- Haggag, K., McManus, B., and Paci, G. (2017). Learning by driving: Productivity improvements by new york city taxi drivers. *American Economic Journal: Applied Economics*, 9(1):70–95.
- Hanna, R. and Wang, S.-Y. (2017). Dishonesty and selection into public service: Evidence from india. *American Economic Journal: Economic Policy*, 9(3):262–90.
- Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *JL Econ. & Org.*, 7:24.
- Iyer, L. and Mani, A. (2012). Traveling agents: political change and bureaucratic turnover in india. *Review of Economics and Statistics*, 94(3):723–739.
- Kang, J.-K. and Kim, J.-M. (2008). The geography of block acquisitions. *The Journal of Finance*, 63(6):2817–2858.
- Katz, E. and Ziderman, A. (1990). Investment in general training: the role of information and labour mobility. *The economic journal*, 100(403):1147–1158.

- Khan, A. Q., Khwaja, A. I., and Olken, B. A. (2019). Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings. *American Economic Review*, 109(1):237–70.
- Kofoed, M. S. et al. (2019). The effect of same-gender or same-race role models on occupation choice evidence from randomly assigned mentors at west point. *Journal of Human Resources*, 54(2):430–467.
- Kuhlmann, S. and Wollmann, H. (2019). *Introduction to comparative public administration: Administrative systems and reforms in Europe*. Edward Elgar Publishing.
- Loewenstein, M. A. and Spletzer, J. R. (1998). Dividing the costs and returns to general training. *Journal of labor Economics*, 16(1):142–171.
- Mackenzie, C. (1979). The école nationale d’administration and the civil service college. *Comparative Education*, 15(1):11–16.
- Moreira, D. and Pérez, S. (2021). Civil service reform and organizational practices: Evidence from the pendleton act. Technical report, National Bureau of Economic Research.
- Muralidharan, K. and Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from india. *Journal of political Economy*, 119(1):39–77.
- Nath, A. et al. (2016). Bureaucrats and politicians: Electoral competition and dynamic incentives. *Unpublished working paper. Boston University*.
- Neal, D. (2011). The design of performance pay in education. In *Handbook of the Economics of Education*, volume 4, pages 495–550. Elsevier.
- Rasul, I. and Rogger, D. (2017). Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *The Economic Journal*, 128(608):413–446.
- Rivera, C. V. (2020). Loyalty or incentives? how party alignment affects bureaucratic performance. *The Journal of Politics*, 82(4):1287–1304.
- Rodríguez-Planas, N. (2012). Longer-term impacts of mentoring, educational services, and learning incentives: Evidence from a randomized trial in the united states. *American Economic Journal: Applied Economics*, 4(4):121–39.

Sandvik, J., Saouma, R., Seegert, N., and Stanton, C. T. (2021). Treatment and selection effects of formal workplace mentorship programs. Technical report, National Bureau of Economic Research.

Shaw, K. and Lazear, E. P. (2008). Tenure and output. *Labour Economics*, 15(4):704–723.

10 Figures and Tables

Table 1: Randomization check: Individual characteristics

	Log Distance	
	Pre 2007	Post 2007
Mean of dep. var	3.073	3.296
caste = OBC	-0.573 (0.392)	0.403 (0.245)
caste = SC	-0.360 (0.592)	0.373 (0.363)
caste = ST	-0.353 (0.712)	0.605 (0.394)
female	-0.382 (0.327)	0.070 (0.256)
age	-0.044 (0.074)	-0.017 (0.042)
Entrance exam percentile	-0.003 (0.009)	0.002 (0.005)
Education degree	-0.093 (0.300)	-0.200 (0.214)
F stat	0.75	1.06
State X Intake year FEs	Y	Y
Observations	258	450

This table relates shows that junior officer's characteristics cannot predict the distance of their training district from its headquarter. Standard errors are in parentheses and clustered at State cadre intake year level.

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 2: Randomization check: District characteristics

	% rural (1)	% literate (2)	% Low SES (3)	Log size (4)
caste = OBC	-0.040 (0.025)	0.001 (0.016)	-0.018 (0.020)	0.022 (0.089)
caste = SC	-0.034 (0.028)	0.002 (0.023)	-0.036 (0.027)	-0.189 (0.135)
caste = ST	-0.064 (0.045)	-0.012 (0.036)	-0.016 (0.040)	0.019 (0.154)
female	-0.030 (0.025)	0.012 (0.012)	-0.015 (0.027)	0.157 (0.080)
age	0.004 (0.004)	-0.002 (0.002)	0.006 (0.004)	0.003 (0.012)
Entrance exam percentile	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.002)
Education degree	0.024 (0.018)	0.004 (0.011)	-0.006 (0.024)	-0.069 (0.061)
F stat	1.54	0.52	1.63	1.65
Mean of dep. var	0.704	0.578	0.266	14.59
State X Intake year FEs	Y	Y	Y	Y
Observations	258	258	258	258

This table relates shows that junior officer's individual characteristics cannot predict the characteristics of the districts where they get assigned for the district training . Unit of observation is IAS officer inducted before 2007. SStandard errors are in parentheses and clustered at State cadre intake year level.

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 3: Randomization check: Mentor's characteristics

	Recruited by exam (1)	Female (2)	Age (3)	Experience (4)
caste = OBC	0.035 (0.078)	-0.049 (0.094)	-0.236 (1.340)	0.458 (0.526)
caste = SC	0.119 (0.089)	0.093 (0.137)	-0.911 (1.702)	-0.620 (0.661)
caste = ST	0.264 (0.140)	-0.119 (0.151)	-5.117* (2.097)	0.685 (0.924)
female	-0.062 (0.070)	-0.039 (0.059)	1.218 (1.212)	0.159 (0.436)
age	-0.013 (0.013)	-0.142 (0.175)	0.287 (0.201)	-0.011 (0.073)
Entrance exam percentile	-0.000 (0.001)	0.001 (0.002)	-0.014 (0.020)	0.010 (0.010)
Education degree	0.038 (0.047)	0.011 (0.048)	-0.179 (0.848)	0.286 (0.308)
F stat	1.24	1.43	1.61	1.25
Mean of dep. var	0.806	0.136	28.31	9.984
State X Intake year FEs	Y	Y	Y	Y
Observations	258	258	258	258

This table shows that junior officer's characteristics cannot predict the characteristics of the mentor that they get assigned to. Standard errors are in parentheses and clustered at State cadre intake year level.

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 4: Performance in district training

	(1)	(2)	(3)
	Percentile District Training		
Log Distance	3.740*** (0.755)	3.432*** (0.790)	3.235*** (0.725)
Individual Controls	Y	Y	Y
District Controls	N	Y	Y
Sup Controls	N	N	Y
State X Intake year FEs	Y	Y	Y
Observations	258	258	258

This table relates performance of junior officer to the distance of the district from its headquarter. Unit of observation is IAS officer inducted before 2007. Dependent variable is percentile of the junior officer in the district training exam. Standard errors are in parentheses clustered at state cadre intake year level:

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 5: Suspension

	Suspension X 100			
	(1)	(2)	(3)	(4)
Mean of dep var	0.125	0.125	0.125	0.129
Log (1 + Dis)	-0.034*	-0.043*	-0.043*	-0.097**
	(0.020)	(0.025)	(0.023)	(0.047)
Individual Controls	Y	Y	Y	Y
District Controls	N	Y	Y	Y
Mentor Controls	N	N	Y	Y
State X Intake year FEs	Y	Y	Y	Y
State X Tenure FEs	Y	Y	Y	Y
Sample	All Officers			Restricted
Observations	5582	5582	5582	3109

Unit of observation is the IAS officer in a given year. Dependent variable is a binary variable taking value 1 if the officer was under suspension in that year. Dependent variable has been scaled by a factor of 100 to improve readability of the coefficients. Log (1 + Dis) is the natural log of the distance of the training district from the headquarter in kilometers. Individual, mentor and district controls are same as before. State X tenure FE are dummies for State specific experience level. All officers sample includes all IAS officers recruited in the year 1996-2006. Restricted sample is a sub-sample of years where district training scores for officers are available. Standard errors are in parentheses and clustered at State cadre intake year level and officer level.

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 6: Take-up of trainings in future

	(1)	(2)
	Taking a training	Number of training days
<i>Panel A: Elective training</i>		
Mean of dep. Var	0.0887	0.693
Log Distance	0.007*	0.051**
	(0.004)	(0.025)
<i>Panel B: Mandatory training</i>		
Mean of dep. Var	0.0868	4.003
Log Distance	-0.001	-0.006
	(0.001)	(0.034)
Individual Controls	Y	Y
State X Intake year FEs	Y	Y
State X tenure FEs	Y	Y
Observations	3190	3190

Unit of observation is the IAS officer in a given year. Column (1) 's dependent variable is a binary variable taking value 1 if the officer took any training in that year. Column (2) 's dependent variable is a number of days spent by officer on training in that year. Dependent variable has been scaled by a factor of 100 to improve readability of the coefficients. $\text{Log}(1 + \text{Dis})$ is the natural log of the distance of the training district from the headquarter in kilometers. Individual, mentor and district controls are same as before. State X tenure FE are dummies for State specific experience level. Standard errors are in parentheses and clustered at State cadre intake year level and officer level. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 7: Relationship between $\log(1+\text{Distance})$ and district characteristics

	(1)	(2)	(3)	(4)
	Proportion Rural	Proportion Literate	Proportion Low SES	Log Population
<i>Panel A: Pre-reform sample</i>				
Mean of dep. Var	0.706	0.576	0.266	14.60
Log (1+ Dis)	0.012* (0.006)	0.002 (0.004)	-0.011*** (0.003)	-0.034** (0.017)
Observations	258	258	258	258
<i>Panel B: Post-reform sample</i>				
Mean of dep. Var	0.732	0.580	0.266	14.56
Log (1+ Dis)	0.015*** (0.005)	-0.000 (0.002)	-0.003 (0.004)	-0.053*** (0.010)
Observations	450	450	450	450

This table relates shows the correlation between district characteristics and distance. Log (1+ Dis) is the logarithm of distance between district and the district headquarter. Pre-reform sample includes only officers that were recruited before year 2007. Proportion Rural is the percentage of households that are classified as rural by the census, proportion low SES is the proportion of population in the district that belongs to the marginalized social caste of SC and ST, proportion literate is proportion of population who can read or write, population size is the logarithm of (1+population) of the district in the last available census. Standard errors are in parentheses:

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 8: Impact of the reform: District Training Exam Performance

	(1)	(2)	(3)	(4)
	Percentile District Training			
Mean of dep. var	48.36	50.83	49.93	49.93
Log (1 + Dis)	3.235*** (0.725)	0.0255 (0.658)	3.366*** (0.708)	3.235*** (0.714)
Log (1 + Dis) X Post reform			-3.436*** (0.917)	-3.210** (0.973)
State X Intake year FEs	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y
District Controls	Y	Y	Y	Y
Mentor Controls	Y	Y	Y	Y
Controls \times Post reform cohort				Y
Sample	Pre	Post	Pre & post-reform cohorts	
Observations	258	450	708	708

Notes: Unit of observation is an individual IAS officer. Dependent variable is the percentile the officer belongs to in the district training exam. Log (1 + Dis) is the natural log of the distance of the training district from the headquarter in kilometers. Standard errors are in parentheses and clustered at State cadre intake year level.

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 9: Impact of Reform: District Head's Output

	Probability project sanctioned on time		
	(1)	(2)	(3)
Mean of dep. var	0.747	0.690	0.708
Log Distance	-0.0602*	0.0215	-0.0418*
	(0.033)	(0.015)	(0.025)
Log Distance X Post reform			0.0619**
			(0.028)
District Controls	Y	Y	Y
Mentor Controls	Y	Y	Y
State X Intake year FEs	Y	Y	Y
Sample	Pre	Post	Pre & post-reform
Observations	10894	22547	33441

Notes: Unit of observation is an individual project sanctioned by the district head or the mentor. Dependent variable is the likelihood that a MPLADS project is sanctioned within the legal mandate of 75 days. $\text{Log}(1 + \text{Dis})$ is the natural log of the distance of the training district from the headquarter in kilometers. Only includes projects that were proposed during the time frame when the district head mentors a junior officer. Standard errors are in parentheses and clustered at State cadre intake year level.

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 10: Impact of Reform: Other Exam scores

	(1)	(2)	(3)	(4)
	Foundation Course		Phase 1	
	Written	Other	Written	Other
Mean of dep. var	49.37	49.78	49.59	50.32
Log Distance	-0.311 (0.544)	0.574 (0.592)	-0.457 (0.590)	0.605 (0.942)
Log Distance X Post Reform	-0.577 (0.813)	0.271 (0.830)	0.346 (0.698)	-0.351 (1.075)
State X Intake year FEs	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y
District Controls	Y	Y	Y	Y
Mentor Controls	Y	Y	Y	Y
Observations	708	708	708	708

Notes: Unit of observation is an individual IAS officer. Dependent variable is the percentile the officer belongs to in the specific training exam. Dependent variable is the percentile the officer belongs to in the district training exam. Log (1 + Dis) is the natural log of the distance of the training district from the headquarter in kilometers. Standard errors are in parentheses and clustered at State cadre intake year level.

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Table 11: Composition of Junior officers' across periods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	age	female	exam score	edu degree	OBC	SC	ST
Mean of dep. var	27.44	0.225	50.43	0.537	0.326	0.149	0.0730
Log Distance	-0.102 (0.115)	-0.0153 (0.016)	0.711 (1.392)	0.00914 (0.021)	-0.0281 (0.021)	-0.00121 (0.017)	-0.00135 (0.010)
Log Distance	0.0797	0.0176	-1.172	0.00199	0.0422*	0.00580	0.00750
X Post reform	(0.137)	(0.020)	(1.589)	(0.024)	(0.024)	(0.020)	(0.011)
State X Intake year FE	Y	Y	Y	Y	Y	Y	Y
Observations	750	750	750	750	750	750	750

Notes: Unit of observation is an individual IAS officer. Dependent variable is the percentile the officer belongs to in the district training exam. Log (1 + Dis) is the natural log of the distance of the training district from the headquarter in kilometers. Post reform is a binary variable taking value 1 if the officer has been recruited post 2006. Standard errors are in parentheses and clustered at State cadre intake year level.

*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

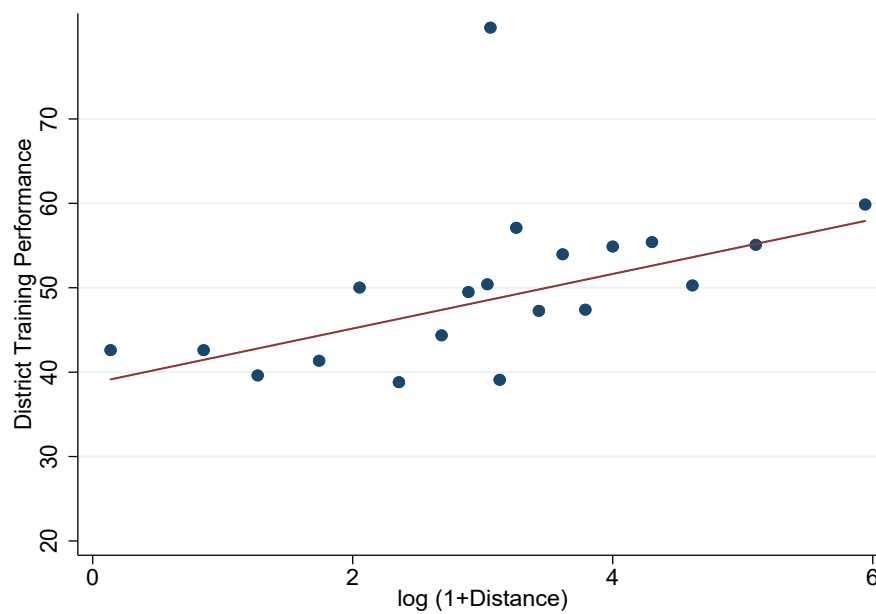
Table 12: Composition of Mentors' characteristics across periods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	age	female	recruit type	tenure	OBC	SC	ST
Mean of dep. var	29.47	0.108	0.767	9.886	0.285	0.0971	0.0485
Log Distance	0.00764 (0.181)	-0.0139 (0.010)	-0.00261 (0.010)	-0.210** (0.081)	0.00323 (0.011)	0.00657 (0.008)	-0.00739 (0.007)
Log Distance X Post reform	0.415* (0.244)	-0.0140 (0.014)	-0.0181 (0.013)	0.147 (0.109)	-0.00532 (0.017)	-0.00882 (0.012)	0.0168* (0.009)
State X Intake year FE	Y	Y	Y	Y	Y	Y	Y
Observations	750	750	750	750	750	750	750

Notes: Unit of observation is an individual IAS officer who is the district head during the training period. Sample only includes districts in years when junior officers gets assigned for training there. Dependent variable is the percentile the officer belongs to in the district training exam. Log (1 + Dis) is the natural log of the distance of the training district from the headquarter in kilometers. Post reform is a binary variable taking value 1 if the officer has been recruited post 2006. Standard errors are in parentheses and clustered at State cadre intake year level.

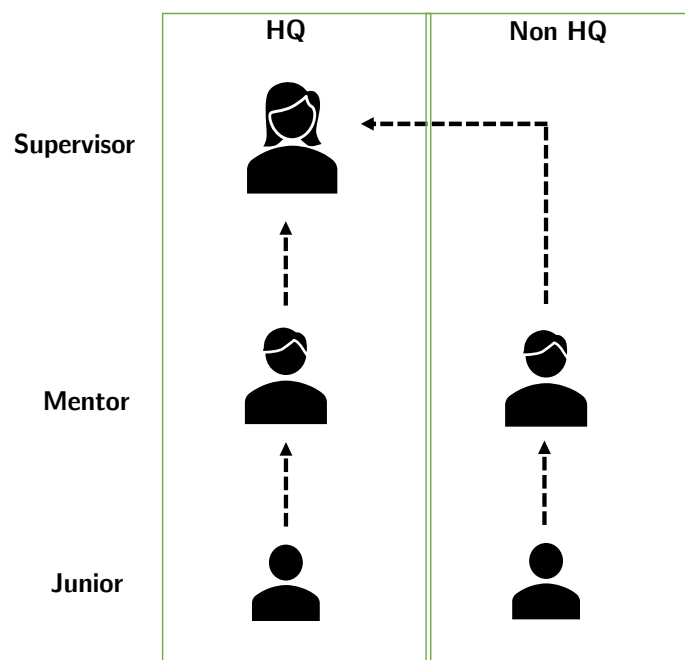
*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$

Figure 1: Distance from HQ and performance in training



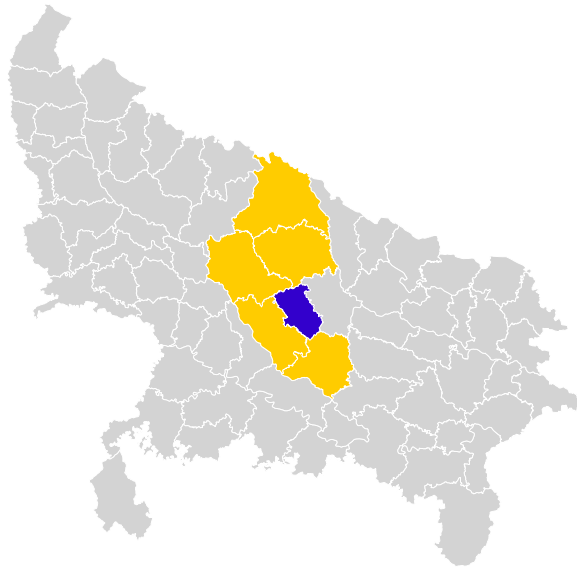
These binned scatterplots show the relationship between distance of training district from the headquarter and performance in district training. Controls included are performance in entrance exams and other components of training, individual characteristics and district and mentors characteristics.

Figure 2: Distance of district from HQ and Performance Evaluation: Illustration



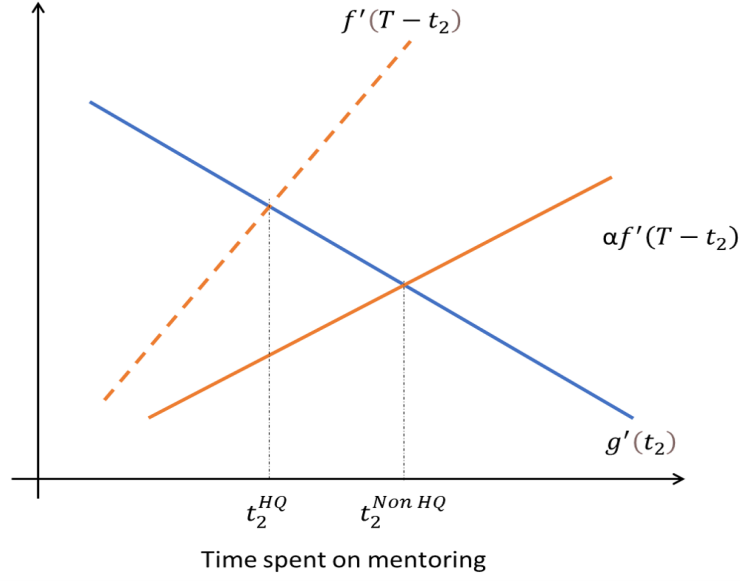
Notes: This figure shows a simplified illustration of variation that I exploit. Juniors get assigned to districts that are headed by the mentor. Districts are headquarter districts (HQ) when the supervisor of the district head is also physically located in that district.

Figure 3: Example of headquarter district: Uttar Pradesh



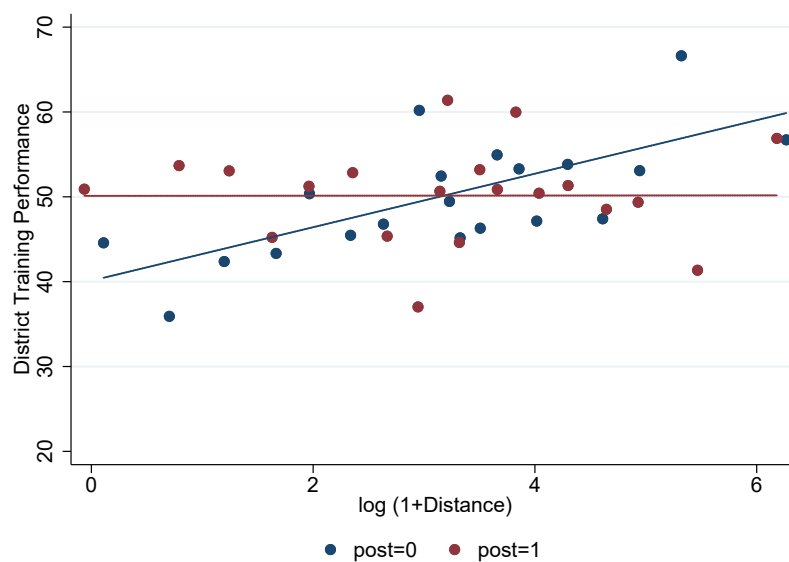
Notes: Blue coloured district is the headquarter district while the yellow coloured districts are non headquarter districts. All district heads in the non-grey districts report to one supervisor called the divisional commissioner.

Figure 4: Conceptual Framework: Illustrative Example



This is a simplified intuition of the conceptual framework. t_2 is the time spent on training the juniors and t_1 is the time spent on own output. FOC requires that $\alpha f'(t_1) = g'(t_2)$. The solid blue line plots the RHS of the FOC. RHS is decreasing function of t_2 and LHS is an increasing function of t_2 . In Headquarter district (HQ) the degree of visibility of the mentor by his supervisor is higher, (here normalized to 1 while visibility of output in districts away from headquarter (Non HQ) α is less than 1.

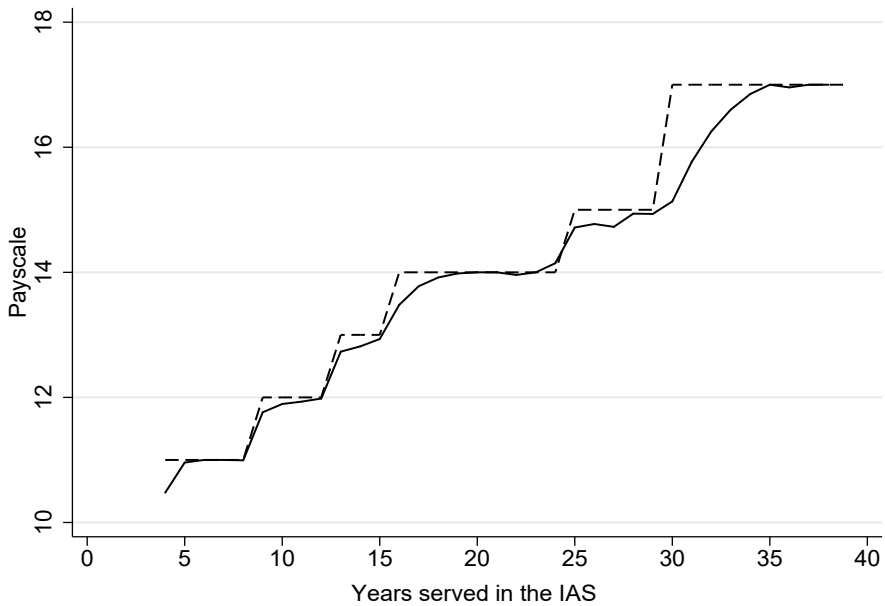
Figure 5: Impact of Reform: District Training Performance



These binned scatter plots show the relationship between distance of training district from the headquarter and performance in district training. Controls included are performance in entrance exams and other components of training, individual characteristics and district and mentors characteristics.

Appendix

Figure 6: Career Level



Tenure or experience based progression rule for the average pay of the IAS. Solid line represents the observed average pay, dotted line represents the rule based average pay