# Breaking Out of Low-Effort Traps: Bureaucratic Leadership by Persuasion*

Martin Haus†

August 14, 2025

**Abstract**

This paper tests if bureaucratic leadership impacts the quality of service provision of a difficult-to-monitor task: learning in public schools. Using the empirical case of rural India, it exploits an administrative setup mimicking a natural experiment with two types of bureaucrats for the same spatial unit, the district, that either have more authority or more ability to engage in time-intensive persuasion. Utilising blocked randomisation inference and bias-corrected variance decomposition on bureaucratic postings linked to independent learning data from household surveys across ten years, it shows that only those bureaucrats with less authority but more ability to engage in persuasion impact learning. Drawing on novel interview data, it illustrates how bureaucratic leaders can increase effort levels of subordinates through persuasion to overcome collective action problems rather than relying on orders and monitoring as principal-agent frameworks would suggest. The findings illustrate that for difficult-to-monitor tasks managerial intensity and persistence trumps formal authority.

***Keywords:*** *state effectiveness; bureaucracy; service delivery; leadership; India; public schools*

***Word count: approx 9,800***

## 1 Introduction

Bureaucracy matters for development in general (Besley et al., 2022) and public service provision in particular (Pepinsky et al., 2017). The quality of core public services like health and education has direct consequences for future life chances or even bare survival. This paper considers the empirical case of public schools in India that fail children on a massive scale, with millions of them leaving the school system every year without being able to read a short story or do division (ASER, 2022; Pritchett, 2013). While studies of developing country bureaucracies mostly focused either on the top or frontline level, more recent research has highlighted that the mid-level of bureaucracy is often neglected (Asim et al., 2023) but matters (Mangla, 2022). If the mid-level of bureaucracy impacts implementation, this implies variation in the quality of service provision

*This work was supported by the Economic and Social Research Council. Fieldwork was supported by the Sir Patrick Gillam Scholarship Fund. This research underwent an ethics review at LSE (ref #42572). I am thankful for feedback on earlier versions of this paper by Daniel Berliner, Sarah Brierley, Joachim Wehner, participants at the King's College Business School PhD colloqium, and the Politics of the Global South workshop at LSE. All remaining errors are mine. I thank Aravind PK for his excellent research assistance.

†Department of Government, London School of Economics and Political Science. m.haus@lse.ac.uk

across subnational units under different leadership. Keeping macro-structures such as formal rules and resources constant, this paper uses independent learning data linked to posting data of different types of bureaucrats to test whether bureaucratic leadership matters for the quality of service provision, whether this impact differs across types of bureaucrats who have available to them different modes of governance, and explores how bureaucratic leaders are able to affect change.

Recent contributions in the field of bureaucracy and service provision have highlighted, e.g., the role of discretion in ensuring effective bureaucracies (Rasul and Rogger, 2018; Rasul et al., 2021; Bandiera et al., 2021; Honig, 2018), how 'overloading' an under-resourced bureaucracy hampers development and what drives uneven bureaucratic resources (Dasgupta and Kapur, 2020), how bureaucrats' transfer preferences can lead to uneven development (Agnihotri, 2022), or how bureaucratic norms shape state effectiveness (Mangla, 2022). Closely related, an emerging literature highlights that state effectiveness is also influenced by performance heterogeneity among individual bureaucrats even if formal rules, political oversight, and resources are kept constant. Fenizia (2022) finds that managers explain significant variation in the output of an Italian public sector agency and Best et al. (2023) show that individual bureaucrats matter for public procurement in Russia.

However, none of these studies on the impact of individual managers focuses on the quality of service provision in a developing country context, likely due to a lack of accessible and reliable data, a factor frequently hampering public administration research in these contexts (Bertelli et al., 2020). Such evidence, however, is particularly relevant as in many low- and middle-income countries, low performance equilibria exist, and weaker organisations possibly amplify the impact individual bureaucratic leaders can have on aggregate performance.

The contribution of this paper is two-fold: first, it extends the literature on leader effects in the public sector to tasks that are "thicker" and less observable, and it tests between the effects of different types of bureaucrats: those that have more authority or those that have less authority

but can invest more time in interaction-intensive persuasion.

The different avenues for bureaucratic leadership are determined by the institutional setup of the empirical case: like other countries whose bureaucratic setup was shaped by the British Empire, the Indian state's core implementation unit, the district, is headed by nationally recruited elite bureaucrats who overlook dozens of departments and provincially recruited officers who head the implementation units of the line departments at the same spatial level of the district. This setup constrains national bureaucrats in deploying a more labour-intensive mode of governance that I term "governance by persuasion" while simultaneously putting them at a position of the highest authority in the district. In contrast, the under-researched provincial bureaucrats heading the line departments can deploy this mode of governance by persuasion but have less authority and power. This presents a setup mimicking a natural experiment which allows to test whether bureaucrats with either more formal power or more possibilities to persuade have an impact on service provision. While not formally a natural experiment, it is nonetheless rare to identify a situation where bureaucrats with the same geographical scope of service responsibility but different forms of authority can be so directly compared.

Taking the case of learning in public schools across two states in India (Bihar and Karnataka with a combined population of around 200 million), I show that provincial bureaucrats have consistent leader effects across their postings over time in different districts with no such effects for their nationally recruited superiors who are equally observed across their different district postings. In the realm of school education, this paper is the first to systematically assess the impact of mid-level bureaucrats on learning in a low- and middle-income country (for a review of the literature, see Asim et al. (2023)).

Similar to other developing country contexts, data availability and reliability is a challenge. To overcome this, I rely on learning data from independent household surveys widely considered the most reliable learning data for the Indian subcontinent. I further assemble novel posting data of hundreds of bureaucrats to then link individual bureaucrats to learning.

To make this argument, I deploy a mixed methods approach that first uses blocked randomisation inference (Berry and Fowler, 2021) and bias-corrected variance decomposition techniques (Bonhomme et al., 2023) to test which bureaucrats impact learning. These methods are recent advancements to those used in earlier pioneering work by Bertrand and Schoar (2003) who considered CEOs and their impact on firm policies and performance or Besley et al. (2011) who tested if more educated leaders lead to more economic growth. For this, I link novel posting data for different types of bureaucrats with independent learning data for each district over time to then test, separately for each bureaucratic type, if a given type impacts learning. Intuitively, these techniques exploit turnover among bureaucratic leaders and, in the case of variance decomposition, the movement of the same bureaucrat across different districts to disentangle the impact of individuals from other factors. The null hypothesis here is interchangeable leaders, i.e., it does not matter for learning which bureaucrat heads a district. In a second step, I shed light on the mechanisms by analysing novel interview data from around 40 interviews across both states. Based on this interview data, I show in the second part of the paper how exceptional provincial bureaucrats deploy "governance by persuasion" to break out of low-effort traps.

More broadly, this paper makes several contributions to the wider literature. First, I show that the mechanism of overcoming low effort equilibria in service provision is less a principal-agent problem but one of collective action (see Ostrom (1998)). Based on interviews, I illustrate how persuasion rather than monitoring and incentives increases effort among subordinates, highlighting the role of intrinsic motivation and trust between public managers and frontline officers. I thereby build on recent work illustrating the importance of management practices beyond surveillance (Khemani et al., 2020; Rasul and Rogger, 2018; Rasul et al., 2021; Bandiera et al., 2021; Honig, 2018; Esteve and Schuster, 2019; Honig, 2021) and contrast this "governance by persuasion" with "governance by order", the predominant mode among national bureaucrats. Second, I add to recent work on state effectiveness and public administration in developing countries (Besley et al., 2022; Bertelli et al., 2020) but move the focus beyond national bureaucrats who have been subject to several studies (see, e.g., Bertrand et al. (2020); Xu et al. (2023); Bhavnani and Lee (2021, 2018); Thakur (2021); Iyer and Mani (2012)) to a small but emerging literature on

provincial bureaucrats (e.g., Dasgupta and Kapur (2020); Agnihotri (2022); Mangla (2022)).

# 2 Modes of governance

## Effort, leadership, and governance in bureaucracies

In a bureaucratic hierarchy, the bottom is populated by bureaucratic agents tasked with policy implementation. At the frontline, these might be teachers or health workers. These agents choose an effort level that can be expressed, following Akerlof (1982, 547), as

$$e_n \geq e_{min} \tag{1}$$

This implies that each agent n will do at least the bare minimum to avoid being fired or, in the case of public bureaucracies, face disciplinary action. $e_{min}$ might be also called $e_{compliance}$, i.e., the level of effort required to comply with administrative rules. Put differently, the problem a principal faces is that of an incomplete contract. An efficient bureaucracy, then, is a bureaucracy where agents go beyond compliance. From the perspective of bureaucratic managers tasked with implementation rather than policy-making, $e_{compliance}$ can be considered as exogenous: administrative rules mandate the legal minimum effort and this minimum effort is often easier monitored from above as it relates to, e.g. the timely "production of paper", i.e., written documents. Improving policy implementation requires more than $e_{compliance}$. Yet, any $e_n > e_{min}$ is more than contractually mandated and hence unenforceable. How, then, can public managers affect a *voluntary increase in effort* by subordinates?

One way of achieving this, following the conventional principal-agent literature, is through monitoring and incentives (Dixit, 2002). Knowing the effort levels of subordinates is a precondition for the provision of monetary or non-monetary incentives that can ensure that agents perform well.

Another way is persuasion. The role of leadership, as Levi (2006) noted, while being about shaping preferences and institutions, mainly "provides the human agency that coordinates the efforts of others" (Levi, 2006, 10). Individuals have utility functions that are not confined to personal benefits but include "concerns about fairness and reciprocity" (Levi, 2006, 13), and leaders play an important role in making sure these concerns are satisfied. This implies that, if leaders succeed in instilling mutually higher expectations, an equilibrium of higher efforts can be achieved. This is a marked departure from simplistic rational choice approaches based on what Ostrom (2000) called "rational egoists" towards a more complex understanding of human (and therefore bureaucratic) behaviour.

I next present two distinct modes of governance based on these contrasting conceptions of leadership.[1]

Governance by order is built on the principal-agent framework and follows the conventional roles with the leader as principal ordering agents what to do and monitoring their activities. This is similar to the notion of agenda-setters in Shepsle (2010). This role of governance is confined to formal structures and the presence of the leader is at decision points but hardly in-between. In contrast, governance by persuasion is based on a collective action framework with leaders as what Shepsle (2010) called entrepreneurs. Interactions between leaders and subordinates are much more frequent and not confined to formal structures, and leaders are present throughout the implementation process. The sources of power of these distinct modes of governance also differ: governance by order builds on the hierarchy of the bureaucratic setup (a formal institutional feature). In contrast, governance by persuasion uses trust as the source of power to affect actions of subordinates. Governance by persuasion goes beyond desks and meeting rooms and uses all interactions with subordinates to affect them. The role of subordinates also differs: while under governance by order, recipients of orders are passive agents tasked with implementing the principal's wishes, governance by persuasion takes the autonomy of agents seriously and con-

[1]Here, I focus on active understandings of the leadership role. In theory and practice, another mode in which bureaucratic managers can act is what Shepsle (2010) called conformity. This flips the principal-agent framework on the head and considers leaders as followers of the wishes of agents. They are mainly concerned about getting to know the preferences of subordinates and gain their power through being submissive.

siders them as active peers and less as the bottom of a hierarchy. At the same time, it does not deny the importance of hierarchy, but leaders differ mainly from subordinates in their special role as entrepreneurs to enable collective actions. Table 1 summarises these features.

| | **Governance by order** | **Governance by persuasion** |
|---|---|---|
| **Theoretical framework** | Principal-agent | Collective action |
| **Role of leadership** | Agenda setter | Entrepreneur |
| **Role of subordinates** | Passive agents | Active peers |
| **Tools of governance** | Ordering & monitoring | Frequent interaction & motivation, symbolic acts |
| **Places of governance** | Desk & meeting rooms (formal structures) | Everywhere (office, corridor, in the "field", virtual spaces) |
| **Frequency of governance** | Low: at the beginning and the end | High: permanent |
| **Source of power** | Hierarchy | Trust |

**Table 1:** Modes of governance

The effectiveness of different modes of governance depends on the kind of task on which they are deployed. I next apply them to the context of this paper: difficult-to-monitor public services.

## The effects of leadership on difficult-to-monitor tasks

Governance by order assumes a benevolent manager who wishes to increase effort levels for better service provision. However, in the case of difficult-to-monitor tasks, managers are unable to enforce compliance by ordering higher effort levels as they cannot verify which individuals shirk and agents, being aware of this impediment, do not transform their believes about the effort levels of others. For governance by order to be effective, principals need to overcome this information asymmetry. Yet, the scale and hierarchical layers make this practically impossible. E.g., a public manager cannot monitor thousands of classrooms or health centres to measure effort levels of a substantial number of frontline bureaucrats. They therefore fail to trigger a cycle of mutually higher expectations that could push the administration out of a low effort equilibrium.

Governance by persuasion, in contrast, has the potential of triggering such a cycle of mutually higher expectation but hinges on a leader's communication ability to generate and foster productive norms based on reciprocity, reputation, and trust (Ostrom, 1998). The main challenge for the public manager who governs by persuasion is coordinating the hierarchy to elicit higher effort levels by creating mutually higher expectations (Miller, 1992, 199). Achieving this flips the information asymmetry problem of governance by order on its head: the public manager needs to find ways for subordinates to know the manager's "intentions and degree of commitment to

cooperation" (Miller, 1992, 196) as superiors have an exceptional role to play (Miller, 1992, 186). In the mechanisms section, I will illustrate creative ways in which public managers minimise uncertainty among subordinates. Public managers can, through acts of cooperation and equality, and by building up a reputation, significantly reduce information asymmetry about their own efforts. Trustworthiness here is not a simplified version of a "nice" manager but whether subordinates can be confident that others are likely to cooperate, i.e., provide high effort (Miller, 1992, 186). This is only possible when moving beyond "rational egoists" (Ostrom, 2000) and allowing for "concerns about fairness and reciprocity" (Levi, 2006, 13) to play a role.

In sum, governance by order poses a major challenge: overcoming information asymmetry about effort levels of thousands of subordinates. In contrast, governance by persuasion requires minimizing uncertainty about effort levels and trustworthiness of a single individual: the public manager.

I next explain why the bureaucratic structure of the empirical case shapes a bureaucrat's propensity to deploy either mode of governance and what predictions this leads to about which bureaucrats can make a difference.

# 3 Empirical case

As Kapur (2020) notes, the Indian state does rather well when implementing episodic events or programs (such as elections or "missions" with a clear end-point) or logistical tasks (such as road-building). This paper instead focuses on a major weakness of the Indian state: the provisioning of a permanent public service that is interaction-intensive and difficult-to-monitor, with social hierarchies leading to segregation and "exit" (Hirschman, 1970) by better-off families from the public system. Others have termed this kind of tasks "wicked hard" (Andrews et al., 2017). Close observers of the Indian state have for long been puzzled by its inability to provide high quality services at scale while simultaneously having one of the most prestigious and competitive civil services in the world (Pritchett, 2009). By putting the spotlight on the underbelly of the Indian state, this paper highlights some aspects that can explain this puzzle.

This paper draws on evidence from two contrasting states in India: Bihar in the North, often seen as the poorest state in India and severely lagging in human development, and Karnataka in the South which is one of the wealthier states in India.

## Shadow of the empire: two types of bureaucrats who implement

The structure of the Indian state, similar to its neighbours of Pakistan or Bangladesh, is inherited from the British Empire. This has led to a distinct bureaucratic setup where nationally recruited bureaucrats who form the bureaucratic elite of the country, in the early stages of their careers, are overlooking the core implementation unit of the Indian state: the district. I term these bureaucrats national bureaucrats as they are largely recruited through the Union Public Service Commission (UPSC) of the national government and become part of the Indian Administrative Service (IAS). Later, these bureaucrats will move on to policy-shaping roles and, as generalists, head the bureaucracy across ministries at the provincial and national level. These national bureaucrats have high social prestige and a distinct standing within the bureaucracy. During their posting in districts, they are the point of contact for ministers and overlook dozens of departments (from education to health, revenue, or law and order). Equipped with government cars featuring red number plates, they are rather insulated from common people, not least as they are guarded by gun-carrying security personnel. Also at the district-level are bureaucrats who are recruited through the public service commissions of the provincial (state) governments. I term these provincial bureaucrats. At district-level, they head district offices of line departments. While they, too, enjoy high social prestige, especially in rural India (Mangal, 2024), their standing within the bureaucracy is different: national bureaucrats as well as politicians frequently look down on them while they are often the most important bureaucratic leaders for frontline bureaucrats. They never move to a policy-shaping role, frequently interact with frontline bureaucrats, and are more accessible for common people or special interests. Provincial bureaucrats differ from frontline bureaucrats as they have a managerial role, often overlooking thousands of frontline personnel. This paper focuses on national and provincial bureaucrats at the district-level. It is important to stress here that although national and provincial bureaucrats differ in their recruitment channels and career prospectives, they both fill leadership roles at the same spatial

unit, the district. This provides a unique opportunity to directly compare two types of bureaucrats with the same scope of service responsibility but different forms of authority, mimicking a natural experiment. Table 2 summarises the features of both types of bureaucrats.

| | National | Provincial |
|---|---|---|
| **Predominant channel of recruitment** | National government: competitive examination | Provincial government: competitive examination or promotion |
| **Social prestige** | Very high | High |
| **Intra-bureaucratic standing** | Very high | From elite bureaucrats: low; from local bureaucrats: high |
| **Policy making role** | Yes (at later career stages) | No |
| **Frequency of interaction with frontline bureaucrats** | Low | High |
| **Frequency of interaction with common people** | Low | Medium |
| **Insulation from special interests** | High | Medium |

**Table 2:** Types of managerial bureaucrats at district-level

This administrative design leads to different propensities for deploying either governance by order or persuasion: as national bureaucrats at the district-level have dozens of departments under them, they are constrained to govern by orders and monitoring meetings. In contrast, provincial bureaucrats can become more involved in the actual provision of services. This allows them to opt for governance by persuasion. I therefore argue that the bureaucratic setup enables provincial bureaucrats to consistently deploy governance by persuasion if they choose to do so. This is not a forced choice though, and they might also choose to govern by order (i.e., there is heterogeneity in leadership strategies). For national bureaucrats, in contrast, choices are more restricted: they can largely only govern by order as they cannot deploy governance by persuasion, with its need for high presence and frequent interactions, across dozens of departments.

Table 3 summarises these constraints in the case of difficult-to-monitor, thick tasks where the ability to verify effort-levels of subordinates ex-post is severely limited.

| | National bureaucrats | Provincial bureaucrats |
|---|---|---|
| **Governance by order** | Feasible, but ineffective for difficult-to-monitor & thick tasks | Feasible, but ineffective for difficult-to-monitor & thick tasks |
| **Governance by persuasion** | Infeasible due to the administrative setup (multiple departments, low frequency of interaction with frontline) | Feasible and effective for difficult-to-monitor & thick tasks |

**Table 3:** Modes of governance by bureaucrat type

## Schools in India

Schools in India are segregated. While those parents who can afford it oftentimes send their children to private schools, public schools mainly cater to the marginalised sections. In 2022, around

three quarters of rural children between 6 and 14 years went to public schools while about one quarter were enrolled in private schools (ASER, 2022). These shares differ across states, with Bihar still having around 80% in public schools while in Karnataka, this share dropped below 70% in 2018 (ASER, 2022). There is also segregation within public schools as the Central government runs its own schools to which, e.g., many members of the all-India services, such as the elite Indian Administrative Service (IAS), send their children. India's public schools for the majority of the population remain heavily under-resourced with large shares violating legally binding minimum infrastructure norms (Rai, 2014). This neglect of education for the masses (Dréze and Sen, 2013; Weiner, 1991) has profound impacts on human capital formation with a huge share of children leaving school in India without being able to read and write (Pritchett, 2013). The neglect of public schools perpetuates social inequality, undermining the ability of historically marginalised sections (such as Schedules Castes and Scheduled Tribes) to gain quality education. This is not a coincidence as upper caste groups have hindered the expansion of primary education by channelling scarce resources instead to secondary schools where their own children went (Chaudhary, 2009). Historically, too, upper castes have undermined state capacity by crippling the state's ability to generate revenues (Suryanarayan, 2021). Even in today's urban India, public schools "are systematically allocated away from neighborhoods where many Muslims and members of Scheduled Castes live" (Asher et al., 2023, 1).

## Administration at district level

The core implementation unit of the Indian state is the district. Districts have vastly different sizes, with the smallest district being home to around 8,000 and the largest one to more than 11 million people (Government of India, 2013). For the two states considered in this paper (Bihar and Karnataka), these numbers are mostly between half a million and 5 million people (Government of India, 2013).[2] Districts have no legislative power thus keeping education policy and the administrative setup constant within states. The following paragraphs explain which posts I treat as national and provincial bureaucrats as well as placebos across both states and why. Their postings will later be linked with learning data to test, separately for each bureaucratic

[2]Educational districts usually coincide with revenue districts except for a few cases in which two education districts fall within a revenue district. As ASER data is following revenue districts, I exclude those cases in Karnataka where educational districts do not coincide.

type, whether they impact learning.

At district level, the general administration is headed by a member of the Indian Administrative Service, recruited by the national-level public service commission, called District Magistrate (DM) or District Collector. The DM has the overall responsibility for the district including law-and-order, revenue generation, but also public service provision including education (Government of India, 2009; Sinha, 2007). Below the DM, early-career IAS bureaucrats or senior bureaucrats of the state administrative services [3] (in Bihar this post is called Deputy District Collector – DDC; in Karnataka CEO of the Zilla Parishad, the district council) are tasked with developmental work. Based on my fieldwork which included interviews and background discussions with DMs, DDCs, and CEOs, it became clear that in Bihar, the DM plays the larger role for the provision of education whereas in Karnataka, the CEO is more involved with education provision, likely a manifestation of more advanced decentralisation efforts in the state. I could observe direct interactions between a CEO in Karnataka and a DM in Bihar with officers of the line department during my fieldwork as well. For these reasons, I have selected DMs for Bihar and CEOs for Karnataka as those national bureaucrats at district level that are most likely to have an influence on education provision.

Turning to the education line department at district level and provincial bureaucrats, the education office is headed by the District Education Officer (DEO) which is called Deputy Director of Public Instruction (Administration) (DDPI (Administration)) in Karnataka. They have either been promoted from the ranks of teachers or entered the service through competitive examinations. Overall, these officers are the nodal officers in charge of school education in a district. For instance, the Bihar Education Code, 1961, notes in article 65 that the DEO has broadly three functions: inspection, supervision, and administration. The code further describes the crucial position of the DEO in assisting the DM (article 72 and 75) which indicates the hierarchy within the

[3] While this implies that not all bureaucrats treated in this paper as "national bureaucrats" are recruited through the national-level public service commission, this further boosts the persuasive power of the empirical strategy: as some of the "national" bureaucrats are from state-services, heterogeneity in training and competitiveness for recruitment is heightened, making it more likely that if the post of the DM or CEO ZP allows to impact learning, this should show up in the empirical test for heterogeneity in leader quality as it increases the heterogeneity among this group of bureaucrats.

bureaucracy, and lays down the powers of DEOs (article 79) that include making appointments of certain teachers and lower subordinate services; promotions, punishments, and suspensions; sanctioning of smaller infrastructure and school investments; or awarding scholarships. This central role of both DEOs and DDPIs (Administration) was also confirmed during my interviews. I will therefore contrast the impacts of DEOs and DDPIs (Administration) as provincial bureaucrats with DMs and CEOs as national bureaucrats across districts.

Parallel to the DEO and DDPI (Administration), there is another post of provincial bureaucrats who head the district teacher training institute. This post is of equal rank and called District Institute of Education and Training principal (DIET principal) in Bihar and Deputy Director of Public Instruction (Development) (DDPI (Development)) in Karnataka. The DDPI (Administration) is more powerful as an executive post and is the main managerial post overlooking service provision. This is also supported by the rule that in case of a departmental enquiry, a prosecution in a criminal case, or an officer is undergoing a punishment or penalty, they are barred from such executive posts [4] but not from being DDPI (Development). This has been echoed during my fieldwork where the DDPI (Development) post was described as a "place to rest". DIET principals and DDPIs (Development) are therefore unlikely to impact learning, and I will refer to them as placebos throughout this paper.

# 4 Data

## Outcome data

As India is known for unreliable learning data even for high-stakes board exams due to mass cheating (see, e.g., BBC News (2015)), I use data gathered by a non-government entity instead. The Annual Status of Education Reports (ASER) provide the highest quality of learning data available in India. This household survey collects annual learning data (usually from September in each year) by testing children in their homes for basic literacy and numeracy with the highest of five levels corresponding to being able to read a short story or doing division, a level that

[4] Rule 5(6) and (7) of the Karnataka State Civil Services (Regulation of Transfers of Block Education Officers/Deputy Directors and other officers in the Department of Public Instruction) Rules, 2015.

should be reached by the end of class two. In addition to learning levels, they also gather the type of school a child goes to (private vs. public). The data is cross-sectional and hence does not track individual children over the years. Data coverage includes (almost) all rural districts in India and is representative at the district level (ASER, nd). I accumulate average learning levels for classes 3 to 8 for each district-year as class 3 is the first grade in which all children should be at level five (the highest level) of the learning test and class 8 is the grade at which the free and compulsory schooling period ends. I use the survey data for the years 2007-2018 in my analyses with no survey undertaken in 2015 and 2017, leading to 10 years of survey rounds with each round covering several thousand children per year in both, Bihar and Karnataka.

## Posting data

I have compiled six novel posting databases of the three types of bureaucrats explained above for each of the two states: national (DMs and CEOs), provincial (DEOs and DDPI (Administration)), and placebos (DIET principals and DDPI (Development)). Online Appendix Section A details the data sources for each bureaucrat type and provides information on posting durations and turnover dynamics which are approximately similar for national and provincial bureaucrats.

Essential for the analyses that follow is how bureaucrats get assigned to districts. While early-career postings of national bureaucrats have been shown to be quasi random (Bhavnani and Lee, 2018), there is no similar mechanism for the later career postings of national bureaucrats or any of the career stages of the provincial bureaucrats under consideration here. This endogeneity in matching bureaucrats to districts poses some threats to valid inference which will be addressed later.

## Linking learning and posting data

To connect bureaucrats with learning outcomes, I determine which bureaucrat oversaw a district for the most days in the 365 days prior to 1 September of each year when ASER usually starts gathering data. This avoids bureaucrats who might have temporarily been in charge for a month or two from being held responsible for the performance in a given year. Online Appendix Section

B provides further details on the learning data.

# 5 Which bureaucrats matter?

I next test which bureaucrats matter, or, to be more precise, whether there is variation across individual bureaucrats of a given type (i.e., whether they have heterogeneous quality) that can explain (part of) the variation in learning. To test this, I will first deploy blocked randomisation inference before applying an alternative method of bias-corrected variance decomposition on one state for which the data allows this (reasons for this are explained in that section).

In both approaches, the aim is not to detect which individual leaders make a difference but whether a given type of bureaucrat as a whole has an impact on learning. For the null-hypothesis of interchangeable leaders to be rejected, two things must hold: first, leaders must have an impact on the outcome of interest; and second, leaders must differ in their quality, i.e., have a heterogeneous impact on the outcome. For my empirical strategy to work, it also requires leaders to have a rather instant effect. Before detailing both empirical approaches, I therefore address this concern of temporality.

## Temporality

While one might worry that changing norms and building trust takes time, based on interview data detailed in the next section, respondents repeatedly noted that such changes usually set in within weeks after a new officer comes in as subordinates quickly try to assess what kind of leadership they can expect. Subordinates hence have a strategic interest to overcome uncertainty about a manager's type. E.g., one interviewee stated that provincial bureaucrats transferred to a new district "will be familiar within 15 days" (KD01003, 2023).

For leader effects to become visible, however, not only the speed of change must be rather instant but it also must not persist after a bureaucrat leaves the post. This, too, was confirmed during fieldwork. E.g., a subordinate in Karnataka made clear that a new equilibrium of higher

effort depended on the bureaucrat being present. Once the officer got transferred, the high-effort cycle was disrupted, and effort levels collapsed (KD01003, 2023). Such an immediate effect was also noted during another interview where an incoming provincial bureaucrat was described as taking many initiatives, including a special support program for lagging students or night classes, leading to changes in results from the first year which kept improving while he was there. Once a new officer took charge, results got weaker (KB01007, 2023). This unravelling of the high-effort convention was also visible during my fieldwork when a provincial bureaucrat was transferred. Once the transfer order became known in the district office, one could immediately sense a difference in the behaviour of subordinates. People were aware that the incoming officer was less "strict", and subordinates started to visit house-warming parties during work hours. There was fewer staff present in the office, and they arrived at irregular times. It hints at the issue of a lack of institutionalisation which might be particularly prominent in low- and middle-income countries: whereas in highly institutionalised contexts, changes to management strategies might be achieved through changes to procedures and rules, in less institutionalised contexts, this change process is much more labour-intensive for managers and contingent on persons rather than institutions rendering governance by order ineffective. At the same time, it is this lack of institutionalisation that enables the empirical strategy in this paper.

## Blocked randomisation inference

Applying the blocked randomisation inference approach by Berry and Fowler (2021), I compare the measure of fit ($r^2$) of the real data with permutations of the spells of officers within districts, keeping the durations of spells unchanged. Take, for example, an officer heading district A for the years 2005 to 2007 and assume we have data for 2005 to 2010. Permutations of this stint would then include simulations where the officer heads district A instead from 2006 to 2008 or 2008 to 2010. Keeping durations constant has the advantage that concerns around serial correlation driving results are mitigated: on average, these correlations in the permutations should be at a similar magnitude. Equally, the number of parameters to estimate (the fixed effects for officers and districts) does not change, hence any incidental parameter bias should equally affect permutations. The approach then records the fit statistic for each permutation, allowing to compare the $r^2$ for the real data with the average of permuted data. This enables a one-sided

test whether the $r^2$ for the real data is exceptionally high in which case the type of bureaucrat is considered to matter. For this approach, I can only reliably use those years where there is no gap in coverage of the outcome data, i.e., 2006-2014 with 2007 being the first year for which I can calculate the change from the previous year. This results in eight year-to-year changes.

Table 4 presents the results of the approach, showing that only provincial bureaucrats matter for learning. The last column indicates the effect size for each officer type. This is based on Monte Carlo simulations which allow to link the differences in model fit back to proportions of variation that can be attributed to officers. Online Appendix Section C illustrates how to link effect sizes back to $\Delta r^2$. I also document that the statistical power across officers showing that for both states, national and provincial bureaucrats are very similar in their statistical power, providing further support to the claim that results are not driven by differences in the ability to detect effects. Although for Bihar only the effect on reading for provincial bureaucrats is statistically significant at the 5%-level, math goes in a similar direction and the effect size would be of a similar magnitude.

The results for Karnataka are more difficult to interpret. First, the statistical power for Karnataka is lower, making it more difficult to detect statistically significant effects. Nonetheless, for provincial bureaucrats in Karnataka math is very close to being significant at the 10%-level. For Karnataka, the "placebo" bureaucrats are also show a significant result even at the 1%-level. This is due to the fact that the cadres of provincial and "placebo" bureaucrats are not separate, unlike in Bihar, and bureaucrats can move between posts of what I call here provincial and placebo bureaucrats. This leads to high correlations in the timing of leadership transitions (a transition in one post predicts the transition in the other for a given year, see Online Appendix Section A.3). The leadership transitions of Karnataka's provincial bureaucrats are highly predictive of transitions in placebo posts ($p = 0.015$).[5] As such, the 'placebo' test in the case of Karnataka is not suitable to serve as a placebo, and the rejection of the null hypothesis can be considered indicative of the confirmation that provincial bureaucrats matter.

[5]In contrast, this is not the case for Bihar where cadres are separate.

Reassuringly, all placebo tests with private school outcomes remain insignificant for all types of officers in both states. These findings indicate that provincial bureaucrats explain more than 30% of the variation in year-to-year changes.

**Threats to inference**

While this approach is relatively robust to serial correlation, endogenous turnover of officers can bias the results, with the direction of bias unknown ex ante. I therefore follow Berry and Fowler (2021) in estimating endogenous turnover using lagged outcomes to use this estimate in Monte Carlo simulations for assessing the magnitude and direction of bias that could result from officers being replaced based on lagged learning (i.e., their past performance). Online Appendix Section C.3 shows the results of this exercise, implying that the amount of bias is relatively small and unlikely to drive results. Providing further reassurance in the case of provincial bureaucrats in Bihar, there are two events that led to turnovers in almost each district at the same time (see Online Appendix Section A.3). First, a restructuring of the Bihar Education Service[6] led to a major reshuffle. Second, a related judgement of the Supreme Court[7] later forced the government to another reshuffle. This makes the timing of turnovers in these years likely exogenous, limiting some of the concerns of bureaucrat performance driving transfers.

## Bias-corrected variance decomposition

An alternative approach to estimate officer effects is based on Abowd et al. (1999) (henceforth AKM) who pioneered a method to decompose earning contributions of workers and firms in payment-linked matched employer-employee networks. The structure of these data is similar to the learning-linked matched officer-district networks in this paper. Following the AKM approach, an additive model with two-sided heterogeneity can be considered:

$$Y_{it} = \psi_i + \tau_t + \alpha_{o(i,t)} + X_{it} + v_{it} \tag{2}$$

where $\psi_i$ is the district effect of district $i$, $\tau_t$ is the year effect, $\alpha_{o(i,t)}$ is the matched officer $o$ overseeing district $i$ in year $t$, $X_{it}$ are time-varying controls, and $v_{it}$ idiosyncratic error terms. For

[6]Office resolution 788 of 2011, Department of Human Resource Development, Government of Bihar.
[7]Judgement, Contempt Petition (Civil) No. 88-89 of 2013, Supreme Court of India (07 May 2014).

| State | Type of bureaucrat | Type of school | Outcome | Real $r^2$ | Average $r^2$ | $\Delta r^2$ | $p$-value | Effect |
|---|---|---|---|---|---|---|---|---|
| Bihar | national | Government | math | 0.324 | 0.308 | 0.016 | 0.3535 | 0 |
| Bihar | national | Government | reading | 0.313 | 0.304 | 0.008 | 0.404 | 0 |
| Bihar | provincial | Government | math | 0.307 | 0.271 | 0.036 | 0.1655 | 0.254 |
| Bihar | provincial | Government | reading | 0.364 | 0.292 | 0.071 | **0.0485*** | 0.33 |
| Bihar | placebo | Government | math | 0.249 | 0.322 | -0.073 | 0.9515 | 0 |
| Bihar | placebo | Government | reading | 0.242 | 0.343 | -0.101 | 0.9775 | 0 |
| Karnataka | national | Government | math | 0.549 | 0.556 | -0.007 | 0.574 | 0 |
| Karnataka | national | Government | reading | 0.576 | 0.582 | -0.005 | 0.568 | 0 |
| Karnataka | provincial | Government | math | 0.419 | 0.342 | 0.077 | 0.101 | 0.385 |
| Karnataka | provincial | Government | reading | 0.288 | 0.313 | -0.025 | 0.573 | 0 |
| Karnataka | placebo$^{\diamondsuit}$ | Government | math | 0.364 | 0.295 | 0.069 | 0.0935 | 0.42 |
| Karnataka | placebo$^{\diamondsuit}$ | Government | reading | 0.432 | 0.283 | 0.149 | **0.005**** | 0.5 |
| Bihar | national | Private | math | 0.323 | 0.355 | -0.032 | 0.7175 | 0 |
| Bihar | national | Private | reading | 0.282 | 0.332 | -0.050 | 0.8185 | 0 |
| Bihar | provincial | Private | math | 0.239 | 0.240 | -0.001 | 0.489 | 0 |
| Bihar | provincial | Private | reading | 0.256 | 0.238 | 0.018 | 0.3015 | 0 |
| Bihar | placebo | Private | math | 0.298 | 0.320 | -0.022 | 0.696 | 0 |
| Bihar | placebo | Private | reading | 0.298 | 0.289 | 0.009 | 0.4075 | 0 |
| Karnataka | national | Private | math | 0.543 | 0.562 | -0.019 | 0.637 | 0 |
| Karnataka | national | Private | reading | 0.571 | 0.548 | 0.023 | 0.3825 | 0 |
| Karnataka | provincial | Private | math | 0.338 | 0.379 | -0.041 | 0.702 | 0 |
| Karnataka | provincial | Private | reading | 0.264 | 0.331 | -0.068 | 0.865 | 0 |
| Karnataka | placebo | Private | math | 0.327 | 0.306 | 0.021 | 0.2965 | 0 |
| Karnataka | placebo | Private | reading | 0.308 | 0.321 | -0.013 | 0.5925 | 0 |

**Table 4:** Results using the blocked randomisation inference. Outcomes are period-to-period changes. The real $r^2$ indicates the model fit with the actual officer spells and permuted $r^2$ provides the mean of model fits using permuted spells. A positive value of $\Delta r^2$ implies that the real spells provide a better fit than most permutations. The $p$-value indicates whether this better model fit is exceptional. The effect column links back the $\Delta r^2$ to the effect size (based on Monte Carlo simulations with details for the effect size provided in Online Appendix Section C.1). Significant results are highlighted in bold. Other relevant results discussed in the main text are underlined. Maximum imposed effect size is 0.5. Effect sizes for results with $p < 0.2$ are shown as 0. Results are based on 2000 permutations for each line.
$^{\diamondsuit}$ See the main text for why the placebo in Karnataka is not an actual placebo.
$^{*}p < 0.05, ^{**}p < 0.01$

the present purpose, there are no time-varying controls and the learning outcomes have been detrended, allowing to simplify to:

$$Y_{it}^{*} = \psi_i + \alpha_{o(i,t)} + \epsilon_{it} \tag{3}$$

This equation is only identified within a connected set. A connected set is built up by movers, i.e., officers who are transferred over time from one district to another. All districts connected by at least one mover form such a set and estimated officer or district effects are only comparable within connected sets, but not across them. A fixed effects estimation determines the share of variation explained by officers and districts within a connected set, following AKM:

$$\underbrace{Var(Y_{it}^{*})}_{\text{Variation in outcome}} = \underbrace{Var(\psi_i)}_{\text{District effects}} + \underbrace{Var(\alpha_{o(i,t)})}_{\text{Officer effects}} + \underbrace{2Cov(\alpha_{o(i,t)}, \psi_i)}_{\text{Sorting}} + \underbrace{Var(\epsilon_{it})}_{\text{Error term}} \tag{4}$$

This requires the estimation of many parameters, frequently leading to a type of incidental parameter bias known as "limited mobility bias" (Andrews et al., 2008). Intuitively, this occurs when networks are identified (i.e., the connected set requirement is fulfilled), but the connectivity is weak, i.e., the number of movers is relatively small. Jochmans and Weidner (2019) provide formal measures to estimate this bias[8] which depends on both, the graph structure (i.e., how densely connected the network is through movers) and the sample size. The bias does not necessarily decrease with an increasing sample size in cases where connectivity decreases. I will now apply this approach to two bureaucrat types in Bihar (national and provincial bureaucrats) as both Bihar placebo officers and all Karnataka graphs are disconnected, creating too small subgraphs to be useful for further analysis. For the Bihar national and provincial graphs, all districts are connected with each other, i.e., all are within a single connected set (see details in Online Appendix Section D.1).

In table 5 I provide estimates of the share of variation explained by national and provincial bureaucrats, as well as districts and the sorting effect for math scores. As both graphs are rather

[8] see also Bonhomme (2020, 91-92)

well connected (see summary statistics in Online Appendix Section D.2),[9] it is unlikely that differing results are driven by statistical artefacts. I present two bias-corrected estimates, using the approach for homoscedastic bias correction by Andrews et al. (2008) and the heteroskedastic version of Kline et al. (2020) both implemented by Bonhomme et al. (2023).[10] This bias-correction is necessary as the plug-in estimator vastly overestimates the variances for the given data (see the Monte Carlo simulations in Online Appendix Section D.3 for an illustration). In both cases, I use not the largest connected set but the leave-out-spell set, an approach that increases connectivity by omitting those districts which are only weakly connected. I further collapse measures at the spell level, lowering the impact of idiosyncratic shocks during an officer's tenure. This decreases bias despite a lower sample size.

| | Estimator | National bureaucrats | Provincial bureaucrats |
|---|---|---|---|
| **Officer** | Homescedastic | -5% | 14% |
| | Heteroscedastic | 4% | 45% |
| **District** | Homescedastic | 70% | 49% |
| | Heteroscedastic | 72% | 45% |
| **Sorting** | Homescedastic | -35% | -45% |
| | Heteroscedastic | -39% | -70% |

**Table 5:** Estimates of the variance decomposition for math. The shares are given in percentage by dividing estimates through the total variance of the outcome. The homoscedastic correction is by Andrews et al. (2008) and the heteroskedastic bias correction by Kline et al. (2020). All estimates are based on the leave-out-spell connected set. I use the implementation by Bonhomme et al. (2023)

The findings indicate that provincial bureaucrats matter for learning, and that their contribution is substantial, explaining between 14% and 45% of the overall variation in detrended math (the number for reading are similar: 25% and 55%, see online Online Appendix Section D.5.1 for details).The difference in estimates depending on the correction techniques is likely due to the imprecision caused by the small sample size which also explains negative estimates for national officer effects.[11] The estimates for national bureaucrats are negative or close to zero (for both math and reading), as are estimates for both officers in private schools (see Online Appendix Section D.5). In practice, negative estimates can be interpreted as 0%. Due to this imprecision for the bias-corrected estimators for the kind of data used in this paper, the numeric values above require caution and should merely be interpreted to confirm that provincial bureaucrats matter

[9] This is particularly the case when compared to widely used teacher value-added applications where graphs are often much less well connected (see Jochmans and Weidner (2019, 1556).

[10] I use their python package PyTwoWay throughout this section.

[11] I confirm the effectiveness of the bias reduction in Monte Carlo simulations documented Online Appendix Section D.3 which also illustrate this imprecision.

whereas no similar effects can be found for national bureaucrats.

## Threats to inference

This approach assumes exogeneity of the network formation. More intuitively, as Fenizia (2022) notes, it requires that mobility is as-if random, conditional on officer and time fixed effects. This does allow for sorting based on time-invariant effects (e.g., if the government sends better quality officers into time-invariantly low-performing districts) but excludes other sources of mobility. Following Fenizia (2022) based on Card et al. (2013), three sources of endogenous mobility appear particularly relevant: a match-component between officers and districts, district-specific trends in learning, and transitory shocks. I address these concerns now.

First, if match-components were driving results, this should show up in patterns of mean residuals (this approach was pioneered by Card et al. (2013) and is a standard way to check for matching effects; see also Abowd et al. (2019, 411), or Fenizia (2022, 1073-74)). In Online Appendix Section D.4, I show that mean residuals are low (for the non-collapsed connected set), and no evident pattern emerges, suggesting that additive separability holds. Second, district-specific trends could bias results if, for instance, officers were able to strategically move into districts that are already improving. Testing for such pre-trends is often done using event-study designs pioneered by Card et al. (2013). However, the infrequent measurements of learning and short tenures of officers in my sample do not allow me to follow this approach. Based on my qualitative fieldwork, I consider it highly unlikely that political or bureaucratic principals can recognize such drift-components, even if they existed, not least because there has been an overall decline in learning and a change in composition of which children go to government schools making disentangling cohort from officer effects difficult. On the other end, as officers are not rewarded for improved learning, there is little reason for them to bear extensive search costs to discover districts that might be improving. Similarly, for transitory shocks to become a threat to inference, political or bureaucratic principals would need to transfer officers due to a "bad $\epsilon_{it}$ draw" (Fenizia, 2022, 1073). This seems unlikely as these draws are specific to the outcome considered in this paper (ASER data), and results get only published in January with district-specific results not widely discussed. Although ASER reports regularly make national headlines, these focus on

national and state-level findings. Exogenous shocks causing both learning and easier-to-monitor process outputs to decline could be one potential source of endogenous network formation but are likely dominated by other causes of transfers. For these reasons it appears justified to maintain the assumption that endogenous network formation is unlikely to drive results. As noted above, a strategic sorting of "better" officers to low-performing districts is no threat to inference.

# 6 Mechanisms: Governance by Order vs Governance by Persuasion

This section largely builds on novel interview data. These interviews were conducted in person together with a research assistant. Around 40 semi-structured interviews took place in Bihar and Karnataka over a period of around two months in 2023, at times spending multiple days in a district interviewing national, provincial, and placebo bureaucrats at district-level as well as their subordinates (at district or block-level, the level below the district). Prior to conducting interviews, non-objection certificates were obtained from the state-level departments and copies were handed over to respondents together with information sheets to provide reassurance. Respondents were asked about their leadership styles or the styles of their superiors (in the case of subordinates), and about any exceptional leaders in the past. Specific examples were probed wherever possible, and respondents were encouraged to elaborate at length leading to interviews at times taking an hour or more. Interviews were conducted in Hindi, English, Kannada, and Tamil (or a mixture).[12] I use the male form irrespective of the actual gender of the bureaucrat to avoid any inference of their identity given the low number of female bureaucrats at the managerial level.

As noted earlier, for the null-hypothesis of interchangeable leaders to be rejected, two things must hold: leaders must differ in their quality and leaders must impact the outcome of interest. I will first address how leaders differ in their quality before zooming in on how leaders are able to bring about positive change for learning in public schools.

[12]This was only possible due to the exceptional work by my research assisant, Aravind PK.

## Leadership heterogeneity

The literature on national bureaucrats has documented heterogeneity among them, based both on perceptions (Xu et al., 2023; Bertrand et al., 2020) as well as actual performance (Thakur, 2021; Bhavnani and Lee, 2018). Such literature does not exist for the case of provincial bureaucrats, but my interviews provide ample illustrations about how provincial bureaucrats, too, differ vastly in their approaches to leadership. E.g., one subordinate noted that provincial bureaucrats can have three types: some are "authoritarian" and give orders; some are friendly and ask for suggestions; and some are in-between. Pressed which one the subordinate would prefer, he chose the third one, the "balanced one" (BI01001, 2023). This heterogeneity was confirmed by another interviewee who described four working styles of provincial bureaucrats as 'one-man show' (governance by order), 'soft', 'motivating people' (governance by persuasion), and 'rule-based'. The interviewee noted that strictly following rules makes work impossible, taking the case of a provincial bureaucrat who opted for this mode but left after a few months as people refused to cooperate. In contrast, another provincial bureaucrat opted for the motivation mode, i.e., governance by persuasion, leaving a lasting positive impression on subordinates (KB01003, 2023).

## Leading change

How can leaders impact the outcome of interest, i.e., learning in public schools, a "thick" task that is difficult to monitor, at the scale of a district with several hundred schools?

#### Feasible but ineffective: Governance by order

Interviewees provided further evidence for why governance by order remains ineffective for this type of task and the scale of a district. As noted earlier, governance by order is based on the need to be able to verify compliance with these orders. However, as one officer noted, a managerial bureaucrat ordering frontline bureaucrats what to do "will never visit the school in next one year because we have 1,200 schools [in the district]", stressing how this way of management is rendered ineffective (KD01003, 2023). The same interviewee, probed which kind of bureaucrat can make a larger impact on education: "100% DDPI [provincial bureaucrat]. Other IAS officers [national bureaucrats] might support but he is not concerned with education, he is concerned

with many departments. [...] But the whole program that should be planned, monitored, and executed by DDPI. The successfulness depends on his involvement, how he supports the other staff, how he is open to all." (KD01003, 2023). It is therefore due to the administrative setup that national bureaucrats are mainly confined to governance by order, a mode of governance rendered ineffective due to the scale and nature of the task at hand.

This was further confirmed in interviews with national bureaucrats themselves. E.g., one national bureaucrat described their role as regularly conducting meetings with the education department where provincial bureaucrats present data based on which solutions are sought (BG01001, 2023). This illustrates the desk-based nature of governance by order where orders are given, and their implementation reviewed at discrete points in time within formal settings. Interestingly, national bureaucrats were often dismissive of the possibility for provincial bureaucrats making a difference: when asked whether provincial bureaucrats can make a difference, a national bureaucrat noted that there is little chance to make a difference as government rules are there (BF01001, 2023), implicitly limiting governance to formal rules without considering the labour-intensive persuasive task that many provincial bureaucrats undertake.

**Governance by persuasion**

The difficulty of providing a clear definition of governance by persuasion is that its deployment is much less structured than governance by order (see also (Shepsle, 2010) on this). What follows are descriptions of some of the ingredients of this management bundle as described by interviewees.

A district-level subordinate in Karnataka noted:

> The nature of work inspires others also. I'm always very much inspired by his [provincial bureaucrat] work. So, I will be present at the same time or when he is coming I will be present. That is the attachment between he and me in office works. When others are not impressive, I am also not interested to come in the same time. Ok I will come. If you are impressing me, I will follow. [...] Another one more point is he is asking colleagues to follow in school visits. He is asking: do you have any

> work? No? Then come with me. We are jointly visited. Some DDPIs [provincial bureaucrats] don't accompany colleagues with them. They are visiting separately; I'm visiting separately. This is the major difference I observed from last three DDPIs. Not only me he asks subject inspectors also. [...] If the same atmosphere is present, I will continue. If it disturbs, I can't move. I am ready to conduct HM´s [headmaster's] meeting regularly. But if the support from the higher authority is not supportive...

(KD01003, 2023)

The provincial bureaucrat "inspired" him to increase effort, he was "impressed" by him (reputation). He worked hard himself and went far beyond working by rule, making subordinates believe that change is possible. After he proved his trustworthiness by working hard and by making sure that others believed in the possibility for positive change, subordinates responded with higher effort levels (reciprocity). Because both the provincial bureaucrat and subordinates gained utility from being part of something normatively desirable, a new equilibrium was reached where both started to provide higher voluntary effort as they had high mutual expectations and believed that their counterpart will equally respond with higher effort (trust). As a result, there was no uncertainty about the manager's own effort and ability of persuasion.

*Frequent interactions, motivation, and symbolic acts*

Exceptional bureaucrats at times also took on the role of training subordinates. As one subordinate described, a provincial bureaucrat provided them with specific examples of how to do their work while others would simply give orders about what to do. This bureaucrat, instead, would train them (KD01005, 2023). While this builds trust with subordinates and increases one's reputation, it also illustrates why national bureaucrats, being largely absent during implementation by design, miss out on many opportunities to inspire subordinates or to signal equality. One interviewee noted that a provincial officer would distribute work in the morning and review and "clear" (i.e., sign files) in the evening and noted that he "enjoyed" this work style a lot and appreciated the "full freedom" in conducting specific tasks (KD01003, 2023), a sign of trust as the bureaucratic leader granted autonomy to subordinates. Interestingly, the same interviewee

noted that this perceived increase in autonomy was accompanied by less ambiguity: the strong presence of the bureaucratic leader included clear instructions and tasks but at the same time left details of implementation open to professional judgement. In the words of the interviewee, this led to "more systematic" work while under other officers, it "is often not very clear what is my task today [...] many people are feeling like we don't really know what our task is" (KD01003, 2023).

Many exceptional officials also used symbolic acts to signal to subordinates that they do not consider themselves above them. Such actions, noted in interviews by other officers, included asking subordinates to join them on school visits in their cars or cleaning their own toilets when cleaning staff was unavailable, a very unusual act for a high-ranking official, especially in India. Some of these acts were done years ago but still vivid in the memories of staff.

*Sources of power: trust and reputation*

Another interviewee described the events during Independence Day celebrations: the provincial officer would arrive before time and make sure that people are punctual. If one came late, the gate would be closed but there was no punishment. Nonetheless, this led to everyone being on time (KB01003, 2023). This leading by example and inspiration was mirrored in many interviews with subordinates. Protection against undue influence has also repeatedly been noted by interviewees as an aspect of exceptional bureaucratic leadership by provincial bureaucrats and is a core ingredient for building trust and reputation. E.g., one provincial bureaucrat provided responsibility and autonomy to subordinates but when there was an attempt of undue influence by organised groups during an event, the provincial bureaucrat took the responsibility to solve it and protect subordinates (KB01003, 2023). Another interviewee recalled an exceptional provincial bureaucrat who, if complaints about subordinates were brought to him, called subordinates directly listening to both sides to come to a fair conclusion while others act under political or other pressure (KB01005, 2023). This builds both, reputation, and trust.

*Places of governance*

One subordinate recalled a provincial bureaucrat during his own probation period who, when visiting schools, would sit for 40 minutes, listen to the class, and teach himself to give teachers specific examples (KB01004, 2023). This "hands-on" approach was mentioned repeatedly when subordinates were asked about exceptional bureaucratic leaders. One provincial bureaucrat himself provided a detailed illustration of governance by persuasion as much more "hands-on" and less desk-based. He elaborated how chat groups are used to motivate teachers, and teachers are encouraged to contribute with their inputs. Exceptional teachers received awards, and local media highlighted their achievements. When clean-up drives in schools take place, the district office participates by cleaning their own premises and visiting schools to support them (BE01001, 2023).

Importantly, such illustrations of different ingredients to governance by persuasion were confined to accounts of the management strategies by provincial bureaucrats, not national bureaucrats, throughout the around 40 interviews.

# 7 Discussion

This paper so far has shown which type of managerial bureaucrat matters for a "thick", difficult-to-monitor task, and provided evidence on how provincial, but not national, bureaucrats are able to make a difference for learning. That provincial bureaucrats can make a difference by increasing effort levels of subordinates also implies that the binding constraint to learning is not a technical limitation (such as teachers' ability) but a low-effort equilibrium. Overcoming a low-effort equilibrium in public organisations mirrors challenges of overcoming suboptimal equilibria in society more broadly. Persson et al. (2013) document how mutual expectations of others acting corruptly have undermined reforms to overcome high corruption equilibria. Such attempts are frequently based on a principal-agent framework and fail not least because principals face few incentives to enforce anti-corruption efforts. While this stream of the literature focuses on "unprincipled principals" (see also Brierley (2020)) lacking incentives for enforcement, I argue that even "principled principals" deploying governance by order lack tools to punish widespread non-compliance with orders of higher effort in contexts of public organisations for difficult-to-

monitor tasks as compliance with these orders is impossible to monitor.

My argument is that overcoming low-effort equilibria in service provision is foremost about getting traction, not about adjusting the speed or steering the wheel. Many reform efforts are changing policy content but like a car with tyres in the air, this does not get it going. Applying this on school education illustrates that which pedagogy teachers apply is a secondary problem. They first need to teach. Reforming pedagogy is, in contexts where few teachers actively teach, a misplaced focus on a counterfactual (it does not matter which pedagogy teachers do not teach) and pilot reforms that simultaneously come with their own cars (whether through special administrative units or non-governmental organisations) provide little insights in how to overcome this first-order challenge. The first-order challenge of getting out of a low-equilibrium trap is to get traction, i.e., to start teaching (however poor the pedagogy might be). This paper has provided evidence that an overlooked actor – provincial bureaucrats – plays a key role in overcoming this first-order challenge.

# 8 Conclusion

Does bureaucratic leadership matter for the quality of service provision and for tasks that are difficult-to-monitor? This paper has shown that even with fixed formal rules and a given stock of inputs, mid-level managerial bureaucrats can and do make a difference. The empirical case of district bureaucrats in India allowed to exploit the administrative setup to test which type of bureaucrat can move effort levels of subordinates beyond compliance. Governance by persuasion needs a "hands-on" approach by public managers who use frequent interactions to lead by example, building trust and reputation. This is in sharp contrast to the desk-based administrator who orders and reviews but does not take part in implementation labour in-between this start- and endpoint. As this paper has illustrated, national bureaucrats who have seen much attention in the literature (e.g., Xu et al. (2023); Bertrand et al. (2020); Thakur (2021); Bhavnani and Lee (2018)) are confined to rely on governance by order which is ineffective to increase voluntary effort levels beyond compliance. Instead, it is the often-overlooked provincial bureaucrats who are able to impact learning. For difficult-to-monitor tasks, intensity and persistence of interac-

tions trumps formal authority. If IAS officers are the steel frame of the Indian state, provincial bureaucrats are the bricks and glass.

# References


Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Abowd, J. M., K. L. McKinney, and I. M. Schmutte (2019). Modeling endogenous mobility in earnings determination. *Journal of Business & Economic Statistics 37*(3), 405–418.

Agnihotri, A. (2022). Transfer preferences of bureaucrats and spatial disparities in local state presence. *World Development 159*, 1–18.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics 97*(4), 543–569.

Andrews, M., L. Pritchett, and M. Woolcock (2017). *Building State Capability: Evidence, Analysis, Action*. Oxford University Press.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008). High wage workers and low wage firms: Negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society, Series A (Statistics in Society) 171*(3), 673–697.

ASER (2022). Enrolment and learning report card.

ASER (nd). Sample design of rural ASER.

Asher, S., K. Jha, P. Novosad, A. Adukia, and B. Tan (2023). Residential segregation and unequal access to local public services in India: Evidence from 1.5m neighborhoods. Unpublished manuscript.

Asim, M., K. Mundy, C. Manion, and I. Tahir (2023). The "missing middle" of education service delivery in low- and middle-income countries. *Comparative Education Review 67*(3), 353–378.

Bandiera, O., M. C. Best, A. Q. Khan, and A. Prat (2021). The allocation of authority in or-

ganizations: A field experiment with bureaucrats. *The Quarterly Journal of Economics 136*(4), 2195–2242.

BBC News (2015, March). India students caught 'cheating' in exams in Bihar. `https://www.bbc.co.uk/news/world-asia-india-31960557`. Accessed: August 4, 2025.

BE01001 (2023, March). Interview with provincial in Bihar.

Berry, C. and A. Fowler (2021). Leadership or luck? randomization inference for leader effects in politics, business, and sports. *Science Advances 7*(4), 1–16.

Bertelli, A. M., M. Hassan, D. Honig, D. Rogger, and M. J. Williams (2020). An agenda for the study of public administration in developing countries. *Governance 33*(4), 735–748.

Bertrand, M., R. Burgess, A. Chawla, and G. Xu (2020). The glittering prizes: Career incentives and bureaucrat performance. *The Review of Economic Studies 87*(2), 626–655.

Bertrand, M. and A. Schoar (2003). Managing with style: the effect of managers on firm policies. *Quarterly Journal of Economics 118*(4), 1169–1208.

Besley, T., R. Burgess, A. Khan, and G. Xu (2022). Bureaucracy and development. *Annual Review of Economics 14*, 397–424.

Besley, T., J. G. Montalvo, and M. Reynal-Querol (2011). Do educated leaders matter? *The Economic Journal 121*(554), F205–F227.

Best, M., J. Hjort, and D. Szakonyi (2023). Individuals and organisations as sources of state effectiveness. *American Economic Review 113*(8), 2121–2167.

BF01001 (2023, March). Interview with national bureaucrat in Bihar.

BG01001 (2023, March). Interview with national bureaucrat in Bihar.

Bhavnani, R. R. and A. Lee (2018). Local embeddedness and bureaucratic performance: Evidence from India. *The Journal of Politics 80*(1), 71–87.

Bhavnani, R. R. and A. Lee (2021). Does affirmative action worsen bureaucratic performance? evidence from the Indian Administrative Service. *American Journal of Political Science 65*(1), 5–20.

BI01001 (2023, March). Interview with subordinate in Bihar.

Bonhomme, S. (2020). Econometric analysis of bipartite networks. In B. Graham and de Paula (Eds.), *The Econometric Analysis of Network Data*, pp. 83–121. London: Academic Press.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2023). How much should we trust estimates of firm effects and worker sorting? *Journal of Labor Economics 41*(2), 291–322.

Brierley, S. (2020). Unprincipled principals: Co-opted bureaucrats and corruption in Ghana. *American Journal of Political Science 64*(2).

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Chaudhary, L. (2009). Determinants of primary schooling in British India. *The Journal of Economic History 69*(1), 269–302.

Dasgupta, A. and D. Kapur (2020). The political economy of bureaucratic overload: Evidence from rural development officials in India. *American Political Science Review 114*(4), 1316–1334.

Dixit, A. (2002). Incentives and organizations in the public sector: An interpretative review. *The Journal of Human Resources 37*(4), 696–727.

Dréze, J. and A. Sen (2013). *An Uncertain Glory: India and its Contradictions*. Princeton: Princeton University Press.

Esteve, M. and C. Schuster (2019). *Motivating Public Employees*. Elements in Public and Nonprofit Administration. Cambridge: Cambridge University Press.

Fenizia, A. (2022). Managers and productivity in the public sector. *Econometrica 90*(3), 1063–1084.

Government of India (2009). Second administrative reforms commission: Fifteenth report – state and district administration.

Government of India (2013). Census of India 2011: Primary census abstract, total population.

Hirschman, A. O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.

Honig, D. (2018). *Navigation by Judgment: Why and When Top Down Management of Foreign Aid Doesn't Work*. New York: Oxford University Press.

Honig, D. (2021). Supportive management practice and intrinsic motivation go together in the public service. *Proceedings of the National Academy of Sciences 118*(13), e2015124118.

Iyer, L. and A. Mani (2012). Traveling agents: Political change and bureaucratic turnover in India. *The Review of Economics and Statistics 94*(3), 723–739.

Jochmans, K. and M. Weidner (2019). Fixed-effect regressions on network data. *Econometrica 87*(5), 1543–1560.

Kapur, D. (2020). Why does the Indian state both fail and succeed? *Journal of Economic Perspectives 34*(1), 31–54.

KB01003 (2023, February). Interview with subordinate in Karnataka.

KB01004 (2023, February). Interview with subordinate in Karnataka.

KB01005 (2023, March). Interview with subordinate in Karnataka.

KB01007 (2023, February). Interview with provincial bureaucrat in Karnataka.

KD01003 (2023, February). Interview with subordinate in Karnataka.

KD01005 (2023, February). Interview with subordinate in Karnataka.

Khemani, S., S. Chaudhary, and T. Scot (2020). Strengthening public health systems: Policy ideas from a governance perspective. Policy research working paper 9220, World Bank Group, Washington, DC.

Kline, P., R. Saggio, and M. Sølvsten (2020). Leave-out estimation of variance components. *Econometrica 88*(5), 1859–1898.

Levi, M. (2006). Why we need a new theory of government. *Perspectives on Politics 4*(1), 5–19.

Mangal, K. (2024). How much are government jobs in developing countries worth? *The World Bank Economic Review*, lhad047.

Mangla, A. (2022). *Making Bureaucracy Work: Norms, Education and Public Service Delivery in Rural India*. Cambridge: Cambridge University Press.

Metcalfe, R. D., A. B. Sollaci, and C. Syverson (2023). Managers and productivity in retail. Working paper 31192, National Bureau of Economic Research, Cambridge.

Miller, G. J. (1992). *Managerial Dilemmas: The Political Economy of Hierarchy*. Cambridge: Cambridge University Press.

Ostrom, E. (1998). A behavioral approach to the rational choice theory of collective action: Presidential address, american political science association, 1997. *American Political Science Review 92*(1), 1–22.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives 14*(3), 137–158.

Pepinsky, T. B., J. H. Pierskalla, and A. Sacks (2017). Bureaucracy and service delivery. *Annual Review of Political Science 29*, 249–268.

Persson, A., B. Rothstein, and J. Teorell (2013). Why anticorruption reforms fail – systematic corruption as a collective action problem. *Governance 26*(3), 449–471.

Pritchett, L. (2009). Is India a flailing state?: Detours on the four lane highway to modernization. Hks faculty research working paper series rwp09-013, John F. Kennedy School of Government, Harvard University.

Pritchett, L. (2013). *The Rebirth of Education: Schooling Ain't Learning*. Washington: Center for Global Development.

Rai, A. (2014). Implementation of the RTE Act: RTE Forum's stocktaking report. *Social Change 44*(3), 439–449.

Rasul, I. and D. Rogger (2018). Management of bureaucrats and public service delivery: Evidence from the Nigerian civil service. *The Economic Journal 128*(608), 413–446.

Rasul, I., D. Rogger, and M. J. Williams (2021). Management, organizational performance, and task clarity: Evidence from Ghana's civil service. *Journal of Public Administration Research and Theory 31*(2), 259–277.

Shepsle, K. A. (2010). *Analyzing Politics: Rationality, Behavior, and Institutions* (2nd ed.). New Institutionalism in American Politics. New York: W. W. Norton & Company.

Sinha, C. (2007). *Public Sector Reforms in India: New Role of the District Officer*. SAGE Publications.

Suryanarayan, P. (2021). Status politics hollows out the state: Evidence from colonial India. Unpublished manuscript.

Thakur, A. (2021). Matching in the civil service: A market design approach to public administration and development. Econtribute discussion paper, no. 087, University of Bonn and University of Cologne, Reinhard Selten Institute (RSI), Bonn and Cologne.

Weiner, M. (1991). *The Child and the State in India: Child Labor and Education Policy in Comparative Perspective*. Princeton: Princeton University Press.

Xu, G., M. Bertrand, and R. Burgess (2023). Organization of the state: Home assignment and bureaucratic performance. *The Journal of Law, Economics, and Organization 39*(2), 371–419.

# Appendix for Breaking Out of Low-Effort Traps: Bureaucratic Leadership by Persuasion

## A Bureaucratic posting dynamics and data sources

### A.1 Data sources

#### A.1.1 Bihar

The posting data for provincial and placebo bureaucrats was provided by the Education Department of Bihar. Where gaps or inconsistencies remained, these were filled by additional research at times obtaining photographs of incumbency boards that are in each office containing tenures and names of officers. DM posting data (national bureaucrats) was assembled from the annual civil list published by the Department of Personnel and Training of the Central government. Where gaps remained or inconsistencies were found, manual searches in newspapers and the executive record sheets of IAS officers on the SUPREMO platform[13] were consulted. As the latter is self-reported data with numerous gaps and errors, the annual civil list was given preference.

#### A.1.2 Karnataka

Posting data for provincial and placebo bureaucrats were provided by the Department of Public Instruction of Karnataka. Where gaps or inconsistencies remained, manual searches in newspapers etc. were undertaken and for some districts, pictures of the incumbency boards were obtained. CEO Zilla Parishad data (national bureaucrats) is based on photos of incumbency boards from all districts which were obtained with the help of the Department of Rural Development and Panchayat Raj of Karnataka.

*I am grateful to the many bureaucrats who supported the gathering of the data, making it possible to compile these novel databases.*

[13] https://supremo.nic.in/KnowYourOfficerIAS.aspx

## A.2 Posting durations

This provides an overview of the different posting durations by bureaucrat type for each state.

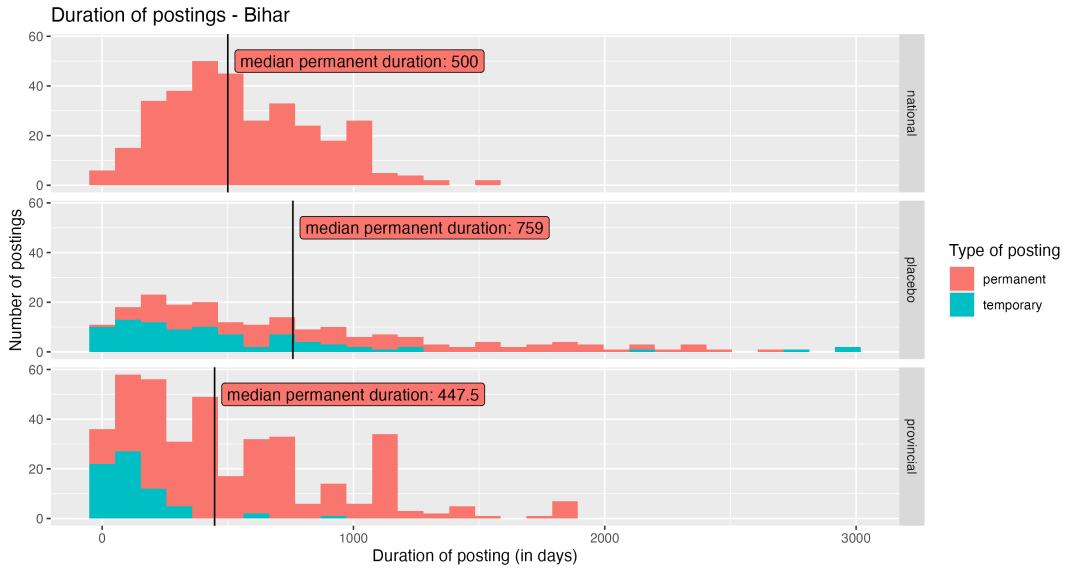


**Figure 1:** Posting durations by officer type for Bihar. The x-axis indicates the number of days a posting lasted. Colours indicate the posting type (permanent or temporary).

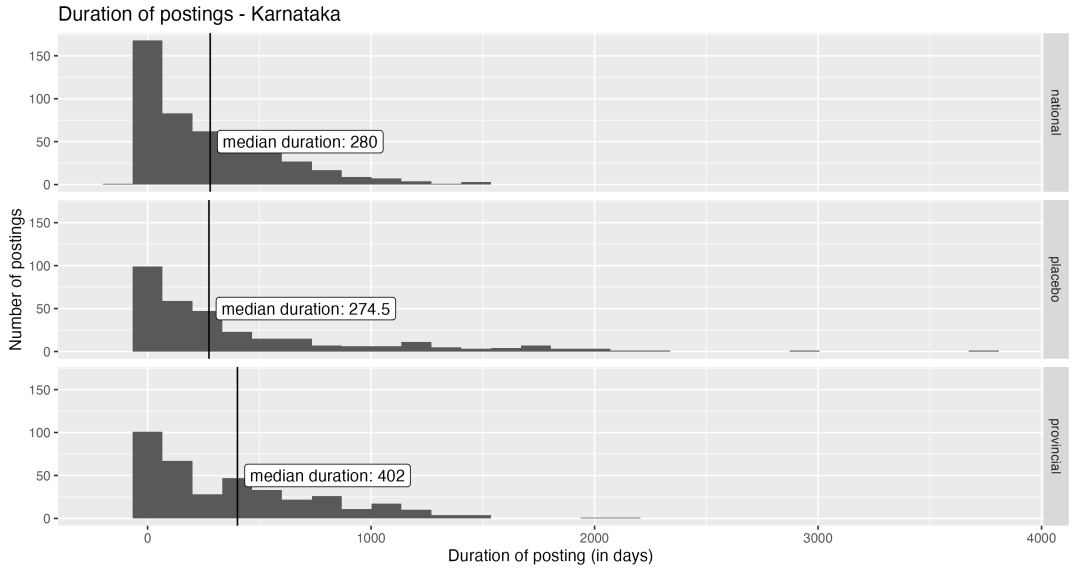


**Figure 2:** Posting durations by officer type for Karnataka. The x-axis indicates the number of days a posting lasted. As this data does not contain information on whether a posting is permanent or not, I cannot differentiate between the two.

## A.3 Turnover dynamics

This provides a quarter-wise overview of transfers by bureaucrat type for each state. The Bihar figure also highlights two instances where almost all districts saw a change in leadership for the

provincial bureaucrats due to changes in the service rules and a reversal of these changes.

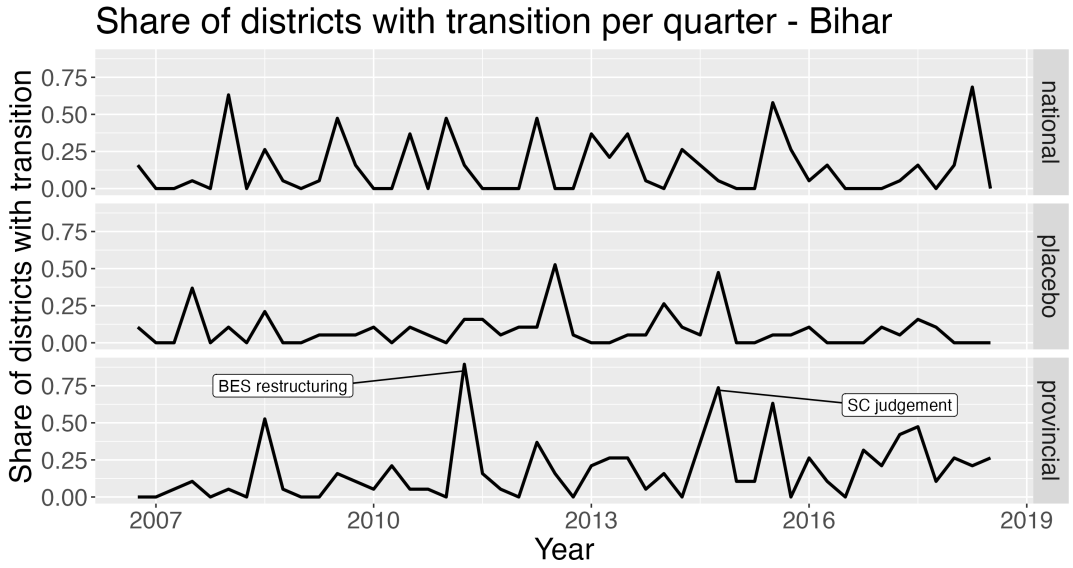


**Figure 3:** Bihar quarter-wise transfers by officer type. The y-axis indicates the share of districts which experienced a turnover in a given quarter. For the highlighted maxima of turnover for the provincial bureaucrats, see the main text.

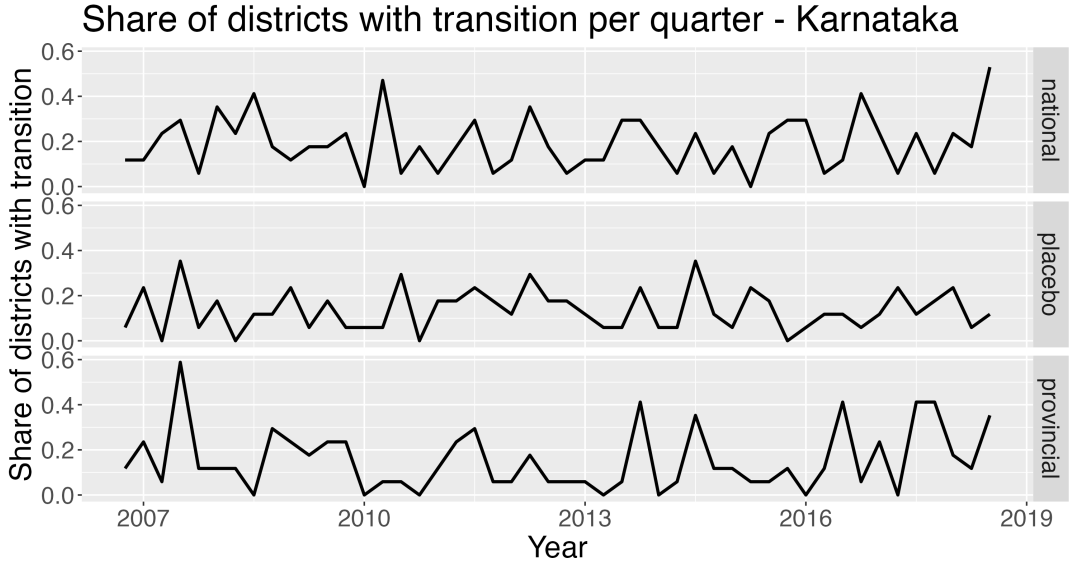


**Figure 4:** Karnataka quarter-wise transfers by officer type. The y-axis indicates the share of districts which experienced a turnover in a given quarter. The close correlation between provincial and placebo officers is likely due to the two cadres not being separated.

# B Descriptive statistics for learning data

This section provides summary statistics of the ASER data used. These are aggregate statistics for the district-year split by state and school type.

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| **average reading score** | 372 | 3.894 | 0.317 | 3.122 | 4.683 |
| **average reading score** | 372 | 3.806 | 0.334 | 2.921 | 4.687 |
| **number of children in reading sample** | 372 | 568.694 | 140.615 | 278 | 1149 |
| **number of children in math sample** | 372 | 566.142 | 139.298 | 280 | 1135 |

**Table 6:** Summary statistics for learning in government schools in Bihar

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| **average reading score** | 225 | 4.042 | 0.311 | 3.213 | 4.733 |
| **average reading score** | 225 | 3.642 | 0.281 | 2.875 | 4.435 |
| **number of children in reading sample** | 225 | 348.133 | 120.363 | 115 | 678 |
| **number of children in math sample** | 225 | 347.756 | 120.159 | 115 | 677 |

**Table 7:** Summary statistics for learning in government schools in Karnataka

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| **average reading score** | 372 | 4.454 | 0.346 | 2.875 | 5 |
| **average reading score** | 372 | 4.326 | 0.365 | 2.762 | 5 |
| **number of children in reading sample** | 372 | 39.742 | 28.535 | 2 | 197 |
| **number of children in math sample** | 372 | 39.629 | 28.524 | 2 | 196 |

**Table 8:** Summary statistics for learning in private schools in Bihar

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| **average reading score** | 225 | 4.202 | 0.288 | 3.494 | 4.841 |
| **average reading score** | 225 | 3.893 | 0.229 | 3.256 | 4.505 |
| **number of children in reading sample** | 225 | 84.836 | 39.734 | 7 | 219 |
| **number of children in math sample** | 225 | 84.76 | 39.718 | 7 | 219 |

**Table 9:** Summary statistics for learning in private schools in Karnataka

# C Additional statistics for blocked randomisation inference

This section provides additional statistics for the blocked randomisation inference approach. It details effect sizes, statistical power, and bias for each bureacrat type in each state.

## C.1 Effect sizes

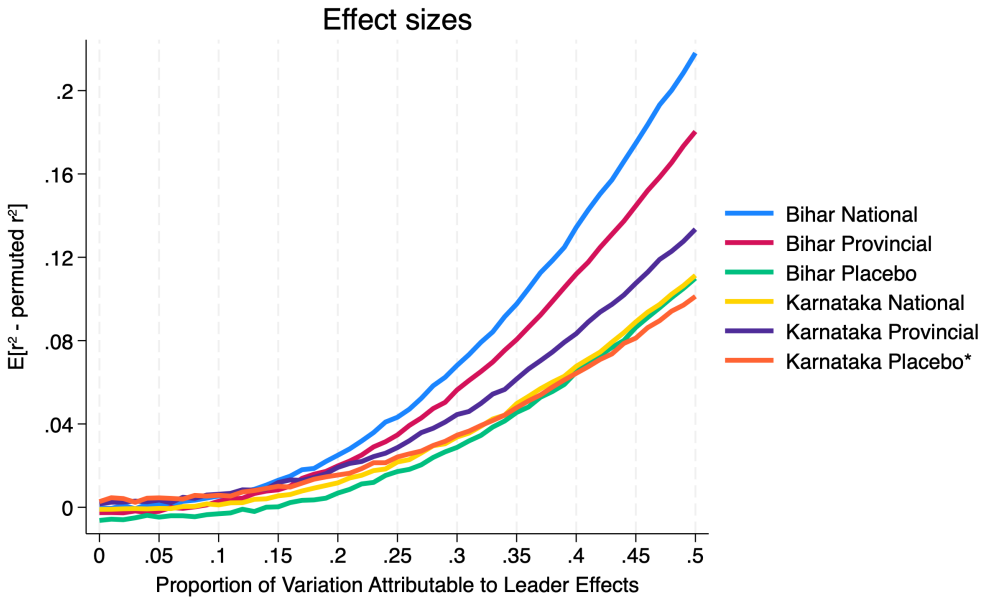


**Figure 5:** Effect sizes for different officer types linking the difference in model fit between the real data and permutations with the proportion of variation explained by officer effects, following figure 3 in Berry and Fowler (2021). The graph is based on 200 iterations of the Monte Carlo simulation.
*See the main article for why the placebo bureaucrats for Karnataka are not suitable as actual placebos.

## C.2 Statistical power

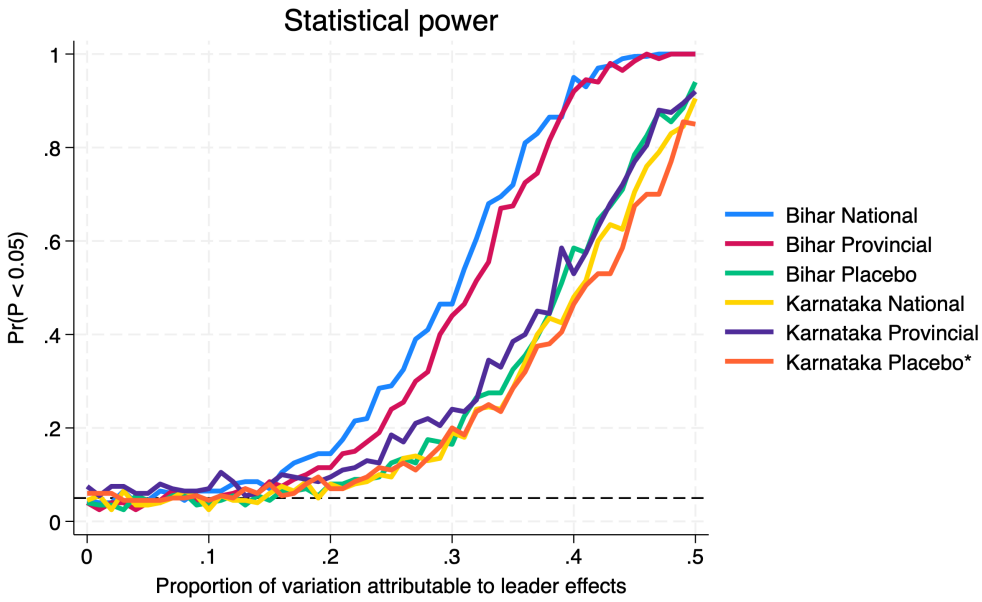


**Figure 6:** Statistical power for different officer types showing the probability of rejecting the null-hypothesis (y-axis) depending on the proportion of variation explained by officer effects, following figure 2 in Berry and Fowler (2021). This illustrates how large the effect sizes need to be in order to be detected. The graph is based on 200 iterations of the Monte Carlo simulations.
*See the main article for why the placebo bureaucrats for Karnataka are not suitable as actual placebos.

## C.3 False rejection rate

In this section, I follow Berry and Fowler (2021) in estimating endogenous turnover using lagged outcomes to utilise this estimate in Monte Carlo simulations for assessing the magnitude and direction of bias that could result from officers being replaced based lagged learning. Table 10 shows the results of this exercise (following table 1 in Berry and Fowler (2021)). The False Rejection Rate (FRR) indicates the share of iterations for which the null hypothesis that leaders do no matter is wrongly rejected (i.e., in these simulations, all leader effects are zero) using the extent of endogenous turnover as observed in the real data.

| state | type of bureaucrat | type of school | outcome | average turnover | slope | $\Delta r^2$ | serial correlation | FRR |
|---|---|---|---|---|---|---|---|---|
| Bihar | national | Government | reading | 0.606 | -0.023 | 0.002 | -0.444 | 0.035 |
| Bihar | national | Government | math | 0.606 | -0.023 | 0.002 | -0.446 | 0.045 |
| Bihar | provincial | Government | reading | 0.468 | -0.037 | 0.005 | -0.444 | 0.055 |
| Bihar | provincial | Government | math | 0.468 | -0.048 | 0.010 | -0.446 | 0.05 |
| Bihar | placebo | Government | reading | 0.409 | -0.015 | 0.001 | -0.461 | 0.08 |
| Bihar | placebo | Government | math | 0.409 | -0.010 | 0.000 | -0.440 | 0.05 |
| Karnataka | national | Government | reading | 0.642 | -0.002 | 0.000 | -0.404 | 0.025 |
| Karnataka | national | Government | math | 0.642 | -0.018 | 0.001 | -0.436 | 0.06 |
| Karnataka | provincial | Government | reading | 0.469 | -0.041 | 0.007 | -0.412 | 0.03 |
| Karnataka | provincial | Government | math | 0.469 | -0.034 | 0.005 | -0.462 | 0.045 |
| Karnataka | placebo | Government | reading | 0.371 | -0.056 | 0.014 | -0.347 | 0.04 |
| Karnataka | placebo | Government | math | 0.371 | -0.014 | 0.001 | -0.406 | 0.045 |
| Bihar | national | Private | reading | 0.606 | -0.044 | 0.009 | -0.515 | 0.045 |
| Bihar | national | Private | math | 0.606 | -0.033 | 0.005 | -0.403 | 0.065 |
| Bihar | provincial | Private | reading | 0.468 | -0.009 | 0.000 | -0.515 | 0.045 |
| Bihar | provincial | Private | math | 0.468 | -0.014 | 0.001 | -0.403 | 0.035 |
| Bihar | placebo | Private | reading | 0.409 | -0.003 | 0.000 | -0.503 | 0.065 |
| Bihar | placebo | Private | math | 0.409 | 0.006 | 0.000 | -0.425 | 0.06 |
| Karnataka | national | Private | reading | 0.642 | -0.027 | 0.003 | -0.404 | 0.035 |
| Karnataka | national | Private | math | 0.642 | -0.057 | 0.015 | -0.367 | 0.06 |
| Karnataka | provincial | Private | reading | 0.469 | 0.011 | 0.000 | -0.423 | 0.055 |
| Karnataka | provincial | Private | math | 0.469 | -0.032 | 0.004 | -0.402 | 0.06 |
| Karnataka | placebo | Private | reading | 0.371 | -0.024 | 0.002 | -0.240 | 0.04 |
| Karnataka | placebo | Private | math | 0.371 | -0.035 | 0.005 | -0.332 | 0.035 |

**Table 10:** False rejection rate (FRR) based on simulations using turnover rate, slope, and serial correlation as in the real data. $\Delta r^2$ indicates the improvement in fit when adding a lagged depended variable. The outcome is a detrended period-to-period change measure. An FRR close to 0.05 indicates low bias. All FRRs are close to 0.05. Each line is based on 200 iterations.

# D Additional statistics and results for variance decomposition

This section provides further statistics and results for the variance decompoistion approach.

## D.1 Connected sets for different bureaucrats

This illustrates why only national and provincial bureaucrats in Bihar can be used for the variance-decomposition approach. All other bureaucrat types have disconnected graphs leading to much smaller connected sets.

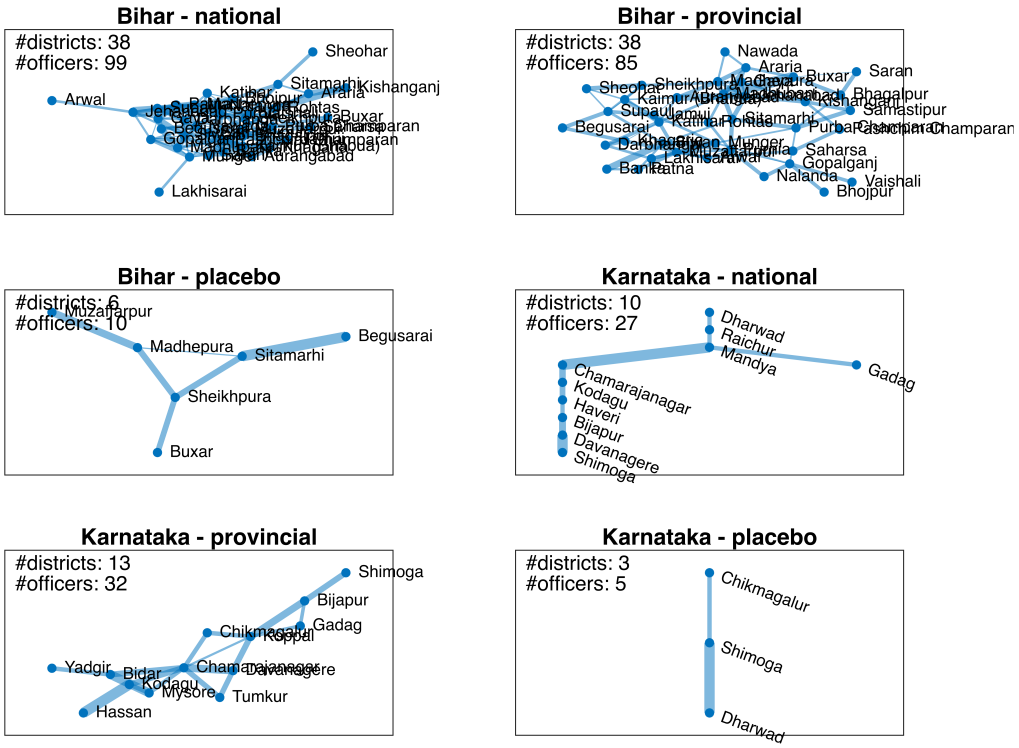


**Figure 7:** Largest connected sets for different types of bureaucrats. Note that for analysis, I use leave-out-spell connected sets that exclude some of the weakly connected districts to further reduce bias.

## D.2 Summary statistics for the connected set used for analysis

This section provides summary statistics for the leave-out-spell connected set used for analysis. Note that the number of officers and districts is lower than for the connected set shown above as for analysis, I use the leave-out-spell connected set which further reduces bias by excluding weakly connected districts.

| Officer type | Number of observations | Number of districts | Number of officers | Number of movers |
|---|---|---|---|---|
| **national** | 179 | 35 | 90 | 54 |
| **provincial** | 125 | 31 | 72 | 37 |

**Table 11:** Summary statistics for the variance decomposition. Observations are collapsed at the spell level and the number of districts, officers, and movers (i.e., officers heading different districts within the sample) relate to the leave-out-spell connected set.

## D.3 Monte Carlo simulations for different estimators

This section details the Monte Carlo simulations using different estimators showing how bias correction matters. I first detail the simulation procedure before providing a plot showing the

performance of different estimators for similar data as the Bihar provincial bureaucrats.

I adapt the python package BipartitePandas[14] to simulate outcomes-linked bipartite networks of a similar size with known true variances, creating networks that unlike the canonical case of employer-employee networks ensure that each district is always and only led by a single officer each year. These simulated networks are hand-calibrated with the real data for provincial bureaucrats in Bihar to approximate a wide range of network statistics of the real data. First, the simulated networks include moderate sorting effects, leading to better performing officers being sorted into lower performing districts. Additional serial correlation is also included. A recent addition to potential biases has been noted by Metcalfe et al. (2023): selection-based negative assortative matching. This occurs if the bureaucratic or political principal is indeed unable to recover the latent type of officers or districts but has some sense of the joint performance. Then, if this joint performance falls below an arbitrary threshold, an officer is moved either to another district or out of the dataset. This appears to be a more realistic assumption than other biases that assume knowledge about the disentangled officer fixed effects or latent type as it involves instead a rather crude rule-of-thumb on the joint (i.e., observable) performance. As their paper shows, this implies a differing stability of matches: particularly poor performing matches (irrespective of the underlying reasons, therefore including bad luck) are more likely to break up whereas the inverse might also be true: if an officer does particularly well in a district, it might lead to an extension of her tenure. This implies a difference in stability of matches, leading to a negatively biased sorting estimate because better districts "will "tolerate" worse managers on average in a stable match" and as "the estimation approach relies exclusively on managers who move (i.e., that were part of an unstable match), this conditionality results in a negatively biased estimate of the correlation between managers and store fixed effects" (Metcalfe et al., 2023, 17). In the given scenario, the nature of the bias might differ as unlike retail stores, the focus of their paper, officers are by default moved (either across districts or to other posts) and staying in-place for many years is on purpose avoided. Nonetheless, their rule-of-thumb approach to differing match stability depending on the observed performance appears a possible scenario. In the simulations, those officers who performed more than one standard deviation above the long-term

[14]https://tlamadon.github.io/bipar5tepandas/index.html

average of a district are more likely to remain in the district while those who performed more than one standard deviation below are more likely to be replaced. This also partially incorporates transitory shocks discussed in the main article. I also include a retirement or promotion "boost" to ensure that officers who have been in the dataset for seven years and more are likely to leave the dataset. Movers receive a "mover boost" to allow them to cover frequently more than two distinct districts and a "newbie boost" increases the probability of a turnover event if a new officer held the district in the previous year, resembling temporary covers that never reappear in the dataset (so-called "single stayers"). Additionally, I incorporate two "shake-up boosts" where turnover probabilities are significantly increased in two years, similar to the real provincial bureaucrat data for Bihar where the education service was restructured. I simulate data for 12 years, not ten, and then drop two years in- between like the real data as there was no learning data collected for 2015 and 2017. I then verify the resulting data by comparing it with the real data using core statistics from Jochmans and Weidner (2019) including $\lambda_2$ as a measure for global connectivity and the estimated bias of the plug-in estimator, as well as the number of movers and distinct officers, the number of edges between districts, the amount of serial correlation and endogenous turnover. A graphical inspection of the resulting graphs confirms that they are indeed similar to the real graph. The simulated data closely resembles the real data on all these statistics simultaneously. I base my justification for the use of my preferred estimator on these Monte Carlo simulations. Figure 8 presents the results of applying different estimators to these simulated data.

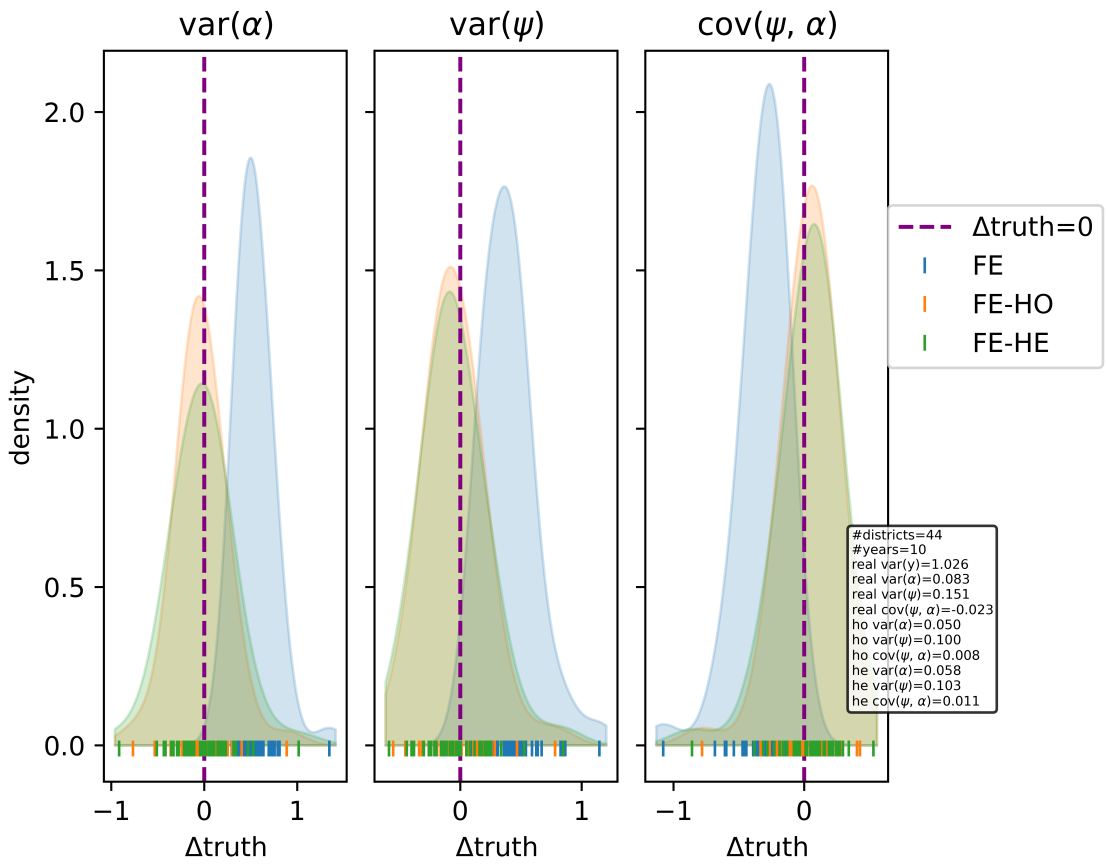


**Figure 8:** Monte Carlo simulations of the variance decomposition using different estimators where $\alpha$ refers to officer effects and $\psi$ to district effects. Each of the three subgraphs indicates the deviation from the true value (at 0) for each estimator. The true values as well as the mean values of the estimators are noted in the textbox on the right. FE refers to the biased plug-in estimator, FE-HO to the homoscedastic correction by Andrews et al. (2008), and FE-HE to the heteroskedastic bias correction by Kline et al. (2020). All estimates are based on the leave-out-spell connected set. The graph is based on 48 Monte Carlo simulations. The graph illustrates the large limited mobility bias of the plug-in estimator (blue) and the the imprecision of the corrected estimators (with the homoscedastic version in orange doing somewhat better) but also confirms their ability to reduce bias significantly.

## D.4 Residual plots

The following residual plots use an approach pioneered by Card et al. (2013) to test for matching effects. I use the implementation by Fenizia (2022).

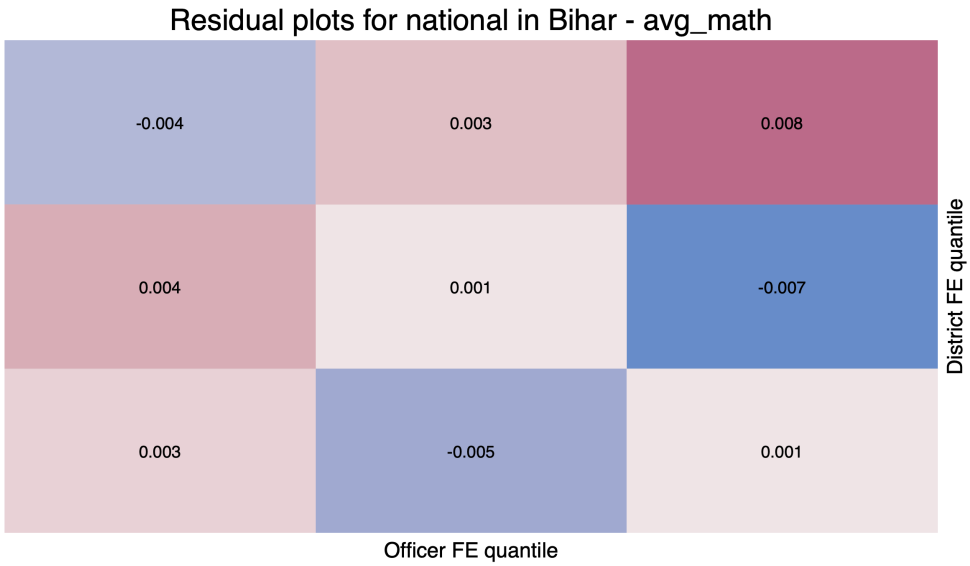


**Figure 9:** Residual plots for national bureaucrats for math in Bihar. The quantiles of the officer fixed-effects are on the x-axis and the quantiles of the district fixed-effects are on the y-axis. Residuals are provided as numbers and highlighted by colours. They are small throughout, and no clear pattern emerges.

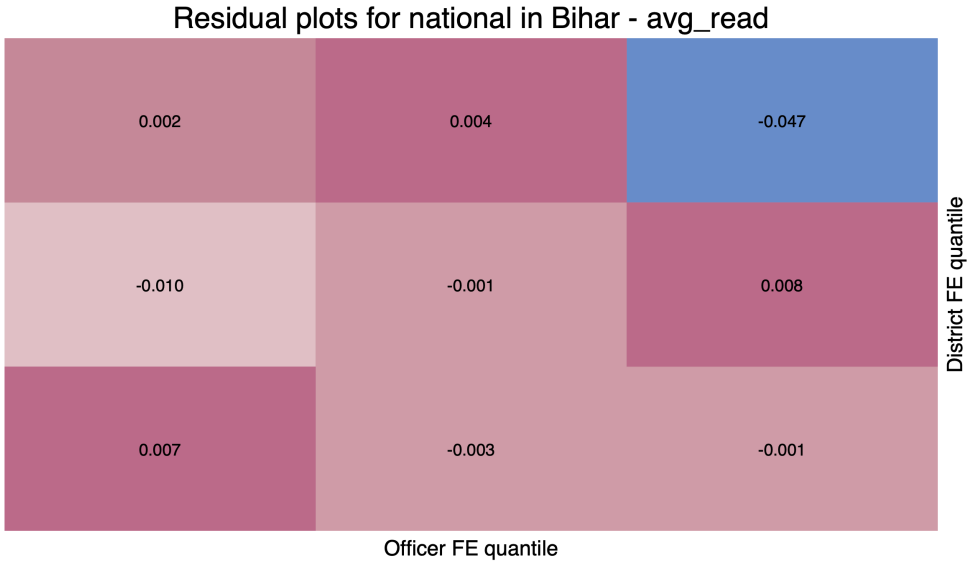


**Figure 10:** Residual plots for national bureaucrats for reading in Bihar. The quantiles of the officer fixed-effects are on the x-axis and the quantiles of the district fixed-effects are on the y-axis. Residuals are provided as numbers and highlighted by colours. They are small throughout, and no clear pattern emerges.

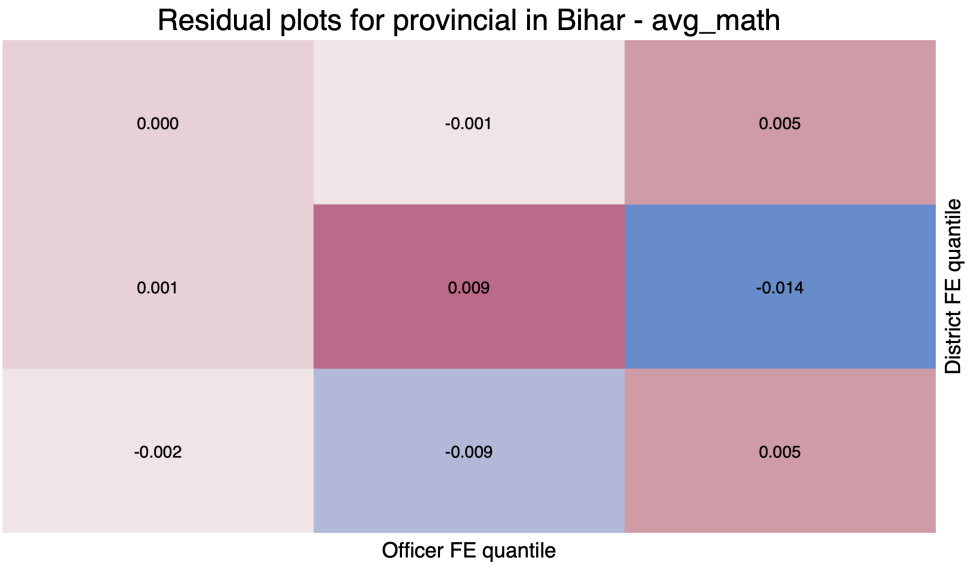


**Figure 11:** Residual plots for provincial bureaucrats for math in Bihar. The quantiles of the officer fixed-effects are on the x-axis and the quantiles of the district fixed-effects are on the y-axis. Residuals are provided as numbers and highlighted by colours. They are small throughout, and no clear pattern emerges.

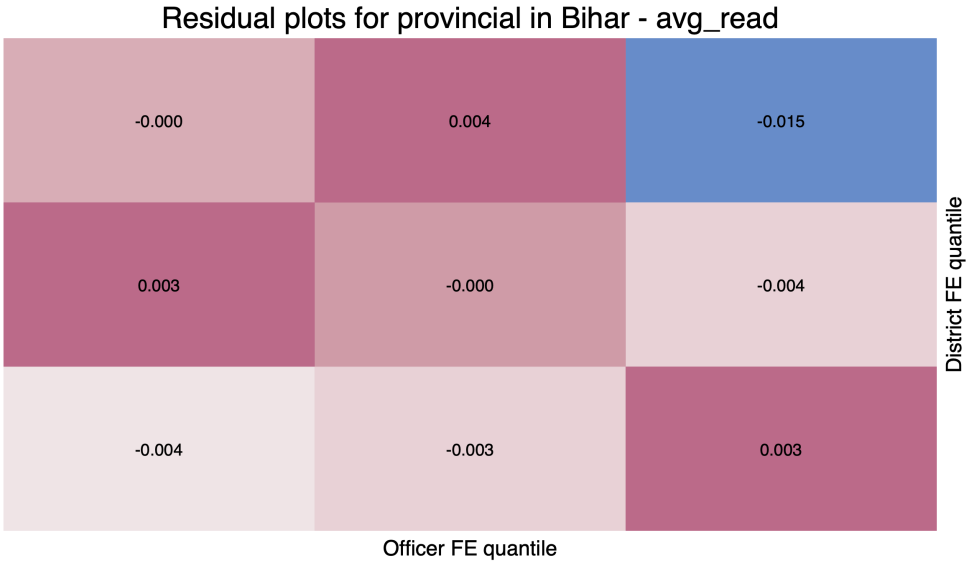


**Figure 12:** Residual plots for provincial bureaucrats for reading in Bihar. The quantiles of the officer fixed-effects are on the x-axis and the quantiles of the district fixed-effects are on the y-axis. Residuals are provided as numbers and highlighted by colours. They are small throughout, and no clear pattern emerges.

## D.5 Additional results

This section contains additional results. I first show that for the alternative outcome of reading in public schools, results are similar as for math. I then detail that for the placebo of private school

learning, neither the provincial nor the national bureuacrats can explain variation in learning.

### D.5.1 Variance decomposition public schools - reading

| | Estimator | National bureaucrats | Provincial bureaucrats |
|---|---|---|---|
| **Officer** | Homescedastic | -17% | 25% |
| | Heteroscedastic | 6% | 55% |
| **District** | Homescedastic | 71% | 50% |
| | Heteroscedastic | 77% | 57% |
| **Sorting** | Homescedastic | -48% | -63% |
| | Heteroscedastic | -58% | -72% |

**Table 12:** Estimates of the variance decomposition for reading in public schools. The shares are given in percentage by dividing estimates through the total variance of the outcome. The homoscedastic correction is by Andrews et al. (2008) and the heteroskedastic bias correction by Kline et al. (2020). All estimates are based on the leave-out-spell connected set. I use the implementation by Bonhomme et al. (2023)

### D.5.2 Variance decomposition private schools - math

| | Estimator | National bureaucrats | Provincial bureaucrats |
|---|---|---|---|
| **Officer** | Homescedastic | -5% | -41% |
| | Heteroscedastic | -43% | -19% |
| **District** | Homescedastic | 76% | 51% |
| | Heteroscedastic | 61% | 61% |
| **Sorting** | Homescedastic | -58% | 9% |
| | Heteroscedastic | 36% | -5% |

**Table 13:** Estimates of the variance decomposition for math in private schools. The shares are given in percentage by dividing estimates through the total variance of the outcome. The homoscedastic correction is by Andrews et al. (2008) and the heteroskedastic bias correction by Kline et al. (2020). All estimates are based on the leave-out-spell connected set. I use the implementation by Bonhomme et al. (2023)

### D.5.3 Variance decomposition private schools - reading

| | Estimator | National bureaucrats | Provincial bureaucrats |
|---|---|---|---|
| **Officer** | Homescedastic | -33% | -36% |
| | Heteroscedastic | -93% | -39% |
| **District** | Homescedastic | 31% | 77% |
| | Heteroscedastic | 13% | 13% |
| **Sorting** | Homescedastic | -7% | -40% |
| | Heteroscedastic | 17% | -65% |

**Table 14:** Estimates of the variance decomposition for reading in private schools. The shares are given in percentage by dividing estimates through the total variance of the outcome. The homoscedastic correction is by Andrews et al. (2008) and the heteroskedastic bias correction by Kline et al. (2020). All estimates are based on the leave-out-spell connected set. I use the implementation by Bonhomme et al. (2023)