Management and Performance in Mid-Level Bureaucracies: Evidence from Ghanaian Education Districts

DeliverEd Initiative Working Paper

Michael Boakye-Yiadom, Clare Leaver, Zahra Mansoor, Maria Pia Iocco

March 2023

the Education Commission

UK aid from the British people

BLAVATNIK SCHOOL OF GOVERNMENT

UNIVERSITY OF OXFORD

IEPA Institute for Educational Planning and Administration at University of Cape Coast, Ghana.

University of Cape Coast Ghana

unesco Centre Under the auspices of UNESCO

DeliverEd: Building knowledge on how to use delivery approaches to advance education reforms

The DeliverEd Initiative was launched in 2019 to strengthen the evidence base for how governments can achieve their policy priorities through delivery units and other delivery approaches. Globally, more than 80 countries have used such approaches to achieve better outcomes for policy reform and implementation. Forty-seven percent of those include an education focus, either as a single focus sector or as part of a multisector approachi. But there was little empirical evidence, especially from developing countries, on the effectiveness of delivery approaches in delivering education outcomes or on the design choices, contextual features, and enabling factors that contribute to their performance.

DeliverEd has helped to fill this evidence gap and create a better understanding of the practices leaders can adopt to improve their policy delivery and reform efforts. It has conducted research within and across countries on the effectiveness of delivery approaches in improving reform implementation, with the key findings included in this final report. It has facilitated knowledge and experience sharing among countries—for example, through the Africa Policy Forum—to equip policymakers with a deeper understanding of delivery challenges and solutions to make informed decisions. It continues to increase awareness and the uptake of research to improve schooling and learning in low-income countries.

The Education Commission leads DeliverEd with Oxford University's Blavatnik School of Government and funding from the UK Foreign, Commonwealth, and Development Office (FCDO). Other partners include the University of Toronto, the Institute for Educational Planning and Administration (under the Auspices of UNESCO), University of Cape Coast, Ghana, Institute of Development and Economic Alternatives (IDEAS) in Pakistan, World Bank, and Georgetown University in the U.S. For more information about DeliverEd, and to view the country studies and other related research and policy engagement materials, please visit www.educationcommission.org/delivered-initiative.

We are very grateful to the Blavatnik School of Government and all our research partners for their in-depth research, especially during the COVID-19 pandemic. This DeliverEd Final report is the Education Commission's interpretation of the research. For the detailed research papers themselves, please see the next page.

DeliverEd Research Products

Anderson, K., Ibarra, A., & Javaid, N. (December 2022). The Education Commission. A Case Study of the Sierra Leone Delivery Unit. DeliverEd Initiative Policy Note.

Bell, S., Asim, M., Mundy, K., Pius Nuzdor, H., Boakye-Yiadom, M., & Mensah Adosi, C. (May 2023). How do regions, districts and schools respond to the introduction of a delivery approach: Evidence from Ghana. DeliverEd Initiative Working Paper.

Bell, S., Leaver, C., Mansoor, Z, Mundy, K., Qarout, D., & Williams, M. (March 2023). The Role of Delivery Approaches in Education Systems Reform: Evidence from a Multi-Country Study. DeliverEd Initiative Working Paper.

Boakye-Yiadom, M., Leaver C., Mansoor, Z., & Iocco, MP. (March 2023). Management and performance in mid-level bureaucracies: Evidence from Ghanaian education districts. DeliverEd Initiative Working Paper.

Gulzar, S., Ladino, JF., Mehmood, MZ., & Rogger, D. (March 2023). Command and Can't Control: An Evaluation of Centralized Accountability in the Public Sector. DeliverEd Initiative Working Paper.

Malik, R. & Bari, F. (May 2023). Improving Service Delivery via Top-Down Data-Driven Accountability: Reform Enactment of the Education Road Map in Pakistan. DeliverEd Initiative Working Paper.

Mansoor, Z., Qarout, D., Anderson, K., Carano, C., Yecalo-Tecle, L., Dvorakova, V., & Williams, M. (July 2021). A Global Mapping of Delivery Approaches. DeliverEd Initiative Working Paper.

Qarout, D. (November 2022). The Accountability Paradox: Delivery Units in Jordan's Education Sector 2010–2019. DeliverEd Initiative Working Paper.

Williams, M., Leaver, C., Mundy, K., Mansoor, Z., Qarout, D., Asim, M., Bell, S., & Bilous, A. (April 2021). Delivery Approaches to Improving Policy Implementation: A Conceptual Framework. DeliverEd Initiative Working Paper.

DeliverEd Policy Products

Anderson, K., & Bergmann, J. (2022). The Education Commission. Design Choices for Delivery Approaches in Education, DeliverEd Initiative Policy Brief.

Anderson, K., & Carano, C. (2021). The Education Commission. The Challenge of Delivering for Learning DeliverEd Initiative Policy Brief.

Villaseñor, P. The Education Commission. (2021) Delivery Approaches in Crisis or Conflict Situations, DeliverEd Initiative Policy Brief.

Management and Performance in Mid-Level Bureaucracies: Evidence from Ghanaian Education Districts[*]

Michael Boakye-Yiadom[†]
University of Cape Coast

Clare Leaver[‡]
University of Oxford

Zahra Mansoor[§]
University of Bristol

Maria Pia Iocco[¶]
University of Sussex

May 2023

Abstract

This paper studies mid-level bureaucracies: organizations operating between national and street levels that are commonly tasked with implementation of sector plans. We examine how these bureaucracies approach policy implementation, how they perform, and the relationship between the two. Our data come from a nationally representative survey of 174 Ghanaian district education offices (and six schools within each of these districts for triangulation), together with a more detailed survey of 1,261 schools in the north of Ghana. We document substantial variation in how district education offices approach policy implementation and their performance (conceptualized in terms of effort by office staff, direct office outputs, and indirect outputs at the frontline). We then show that our management index—a measure of organizational approach to implementation across four delivery functions—is positively associated with performance. Not all of the delivery functions matter equally. Our descriptive evidence points toward benefits from problem-solving practices over top-down accountability, raising policy considerations for Ghana's current delivery approach and beyond.

[†]Institute for Educational Planning and Administration (IEPA) (mboakye-yiadom@ucc.edu.gh)

[‡]Blavatnik School of Government, Toulouse School of Economics, and Center for Economic and Policy Research (CEPR) (clare.leaver@bsg.ox.ac.uk)

[§]School for Policy Studies (zahra.mansoor@bristol.ac.uk)

[¶]School of Economics (mi275@sussex.ac.uk)

1   Introduction

Governments around the world regularly commit to improving service delivery, often in the form of sector plans laying out specific policy reforms. Sadly, for the many who rely on public services, these plans can sometimes fail to translate into tangible improvements. Part of the problem may lie in the design of the chosen policies. However, as politicians are increasingly recognizing, and responding to with institutions such as delivery units, the quality of *implementation* by bureaucrats may be an important contributing factor.

The idea that implementation matters is not limited to policymakers; the last decade has also seen increased study of bureaucracies by academics. Inspired by work documenting a management–performance relationship in private firms (Bloom and Van Reenen, 2007), economists have been exploring whether a similar relationship exists between the management practices of bureaucracies and various dimensions of performance. Early contributions focus on street-level bureaucracies. Bloom et al. (2015), for example, adapt the World Management Survey (WMS) to a school setting and study the relationship between the use of management practices by school principals and student learning outcomes in six high-income countries.[1] More recent contributions focus on national-level bureaucracies. Rasul and Rogger (2018) and Rasul et al. (2021) also adapt the WMS and study the relationship between the use of management practices in civil service organizations and project/task completion rates, the former using a sample of organizations in the Nigerian Federal Civil Service and the latter all organizations in the Ghanaian Civil Service.

Understanding how bureaucracies function at both the national and street levels is important. However, in most countries there is also a sizeable middle tier of bureaucrats working in regional and district offices who are closer to, but not at, the frontline. Sector plans often rely heavily on this middle tier for implementation. For example, sector plans frequently mandate that new procedures (e.g., public health checks, pedagogy) be adopted at the street level. Regional and district bureaucrats are typically charged with delivering the associated training and subsequent monitoring and support of uptake. Likewise, when a sector plan provides for new equipment (e.g., bednets, teaching and learning materials) at the frontline, it is frequently these mid-level bureaucrats who are responsible for identifying need and then delivery of the assets.

Despite the pivotal role played by mid-level bureaucracies in implementation, both policymakers and academics appear to know relatively little about them. This is particularly true in the education sector of low and middle-income countries. Asim et al. (2022) conduct a systematic review of gray and academic literature in all social science disciplines, excluding high-income settings, and conclude that research on mid-level actors (staff at regional, district, and county levels) is limited. This paper addresses this gap. Our objective is to better understand mid-level bureaucracies—how they implement policy and shape service delivery at the frontline—focusing on the specific example of the Ghanaian education sector.

---

[1]Lemos et al. (2021) further refine the WMS for low-income settings. They use their Development World Management Survey (D-WMS) to explore the relationship between the use of structured management practices by school principals and student value-added (as well as other aspects of school performance such as teaching practice) in India.

District education offices are the main administrative layer between Ghana's 16 regions and more than 20,000 schools. These 260 offices lie at the heart of our analysis. We begin by describing their management, asking: what management practices do the directors in charge of district education offices use, and what are the correlates of this usage?

The data that we use to answer these questions come from a nationally representative survey of 174 district education offices undertaken in the summer of 2022. We build on prior work on national and street-level bureaucracies by adopting the methodology of the WMS. However, in contrast with previous papers, our survey instrument is framed around *delivery functions*. These are: prioritization and target setting, monitoring and use of data, accountability and incentives, and problem-solving and adaptation. We purposefully make this change because our goal is to study organizational approaches to implementation.[2]

We also study performance and its relationship with district office management practices. The constructs that previous papers consider under the heading of "performance" (sometimes referred to as "productivity") reflect the nature of the work at hand. Since national-level bureaucracies spend much of their time providing infrastructure, papers studying these organizations have emphasized project or task completion (Rasul and Rogger, 2018; Rasul et al., 2021)—did a borehole, building, or dam, say, get constructed? In contrast, street-level bureaucracies, such as schools, provide services directly at the frontline, prompting authors studying these organizations to focus on delivery outcomes, including student learning levels (Bloom et al., 2015; Leaver et al., 2022; Di Liberto et al., 2015) and learning gains (Crawfurd, 2017; Lemos et al., 2021).

Like many mid-level bureaucracies, the work of district education offices lies somewhere between: they are charged with a distinct set of tasks, but are close enough to the frontline to conceivably influence service delivery outcomes. With this in mind, we study performance at both district *and* school level. For the former, we include three fairly standard constructs: the intensity with which district directors and staff undertake activities (effort), completion of tasks related to the education sector plan (direct office output), and staff attitudes and satisfaction (potential drivers of office output).[3] We supplement these with a more novel construct: staff understanding of priorities and roles, which is also a plausible driver of office output (Adelman and Lemos, 2021). In focusing on these four dimensions, we respond to calls for "thick" (or at least thicker) description when assessing performance, (Honig and Pritchett, 2019). We measure all four dimensions using data from our nationally representative survey of district offices, taking care to independently verify task completion.

We also investigate whether district education offices influence what schools do. Our constructs here relate to head teachers and teachers: the intensity with which head teachers undertake activities in the classroom, head teacher attitudes and satisfaction, teacher absenteeism, and teacher classroom practice. We think of these as indirect district office outputs, although of course they are inputs from the perspective of a school production function. We

---

[2]As we discuss in Section 3, in our survey, district directors were asked a series of questions that were designed to elicit their use of 16 management practices, organized under the four delivery functions. Several of these practices are similar to constructs in the WMS, but the majority are different.

[3]See, e.g., Dasgupta and Kapur (2020) and White (2004) on activity intensity, Rasul et al. (2021) on task completion and staff attitudes, and Bloom et al. (2014) on satisfaction.

measure these dimensions of performance using data from a survey of 1,261 basic schools in the north of Ghana, also undertaken in the summer of 2022. The measures for head teachers are based on survey responses, while those for teachers are based on enumerator observation.[4]

The main takeaways from our analysis can be summarized as follows. First, there is substantial variation in how district education offices approach implementation. There are low scorers making almost no use of the delivery functions and high scorers where the delivery functions are well-established in formal, structured practices. This cross-district variation is positively associated with external factors (support from NGOs and donors, and engagement from politicians) but, interestingly, not with characteristics of directors themselves or office resources. There is also a striking geographic pattern in the style of approach. District offices in south western regions focus more on top-down accountability (practices such as identification of "target owners," accountability meetings, and use of rewards and sanctions) than on problem-solving (practices such as feedback loops, deep dives, and the facilitation of autonomy and innovation). The reverse is true in the rest of Ghana.

Second, performance is distinctly mixed. On the plus side, district staff are, on average, visiting schools frequently, and spend time in the classroom as well as on "compliance" activities. Average completion rates for tasks related to training are also reasonably high. Against this, completion rates are much lower for tasks that require delivery of equipment. District staff also show limited understanding of their district's priorities and their/others' roles, and over 20 percent of those surveyed said they were likely/very likely to leave in the next year. Turning to our sample of schools, there are positives: on average, teacher absenteeism is moderate by regional standards, and aspects of classroom practice meet international benchmarks. But there are also concerns: on average, head teacher effort and satisfaction are low, and, in a minority of schools, teacher absenteeism and classroom practice are a serious problem.

Third, our management index—measuring organizational approach to implementation—is positively associated with performance. Partial effects are statistically significant (correcting for multiple hypothesis tests) for the intensity of activities during field visits to schools by district staff, and classroom practice. The magnitudes are easiest to interpret for the latter: a one standard deviation increase in our management index is associated with just under a 3 percentage point increase in active instruction (against a mean of 74 percent) and a 5 percentage point decrease in students being off task (mean of 21 percent).

Finally, not all of the underlying delivery functions in our index matter equally. We find positive associations between district office use of problem-solving and district staff job satisfaction, teacher presence at school, and classroom practice. To illustrate, a one standard deviation increase in the problem-solving sub-index is associated with a 3 percentage point reduction in teacher absenteeism (against a mean of 13 percent). In contrast, we find *negative* associations between district office use of top-down accountability and teacher presence at

---

[4]We are awaiting access to data from learning assessments administered to all students in primary grades two and four by Ghana's National Council for Curriculum and Assessment (NaCCA) in December 2022. We plan to look at this service delivery outcome in future work.

school, and classroom practice.

These four takeaways provide substantive new evidence on mid-level bureaucracies: how these organizations are approaching policy implementation, how they are performing, and the relationship between the two. We are aware of only two comparable papers, those of Walter (2018) and Cilliers et al. (2022), who study management in district education offices and its relationship to student learning in Zambia and Tanzania, respectively.[5] The contribution of this paper is to consider delivery functions, rather than more general aspects of management, and to provide a broader assessment of performance, particularly as it relates to the quality of implementation. In doing so, we also contribute to wider academic debates on the relative merits of different management styles—so-called "Route X" versus "Route Y" (Honig, 2022; McGregor, 1960). The work of Rasul and Rogger (2018) and Rasul et al. (2021) suggests that high-stakes accountability practices may not be the way to drive task completion by national bureaucracies. Our results extend this conclusion to the middle tier and to performance more broadly defined.

To understand why sector plans so often fail to translate into tangible improvements, it is necessary to examine the nuts and bolts of implementation, and to do so at the middle tier. We took on this challenge and developed a new survey instrument to measure the relevant aspects of management. We also thought through how to achieve a "thick" description of performance, spanning the effort, direct and indirect outputs of mid-level bureaucrats. Our survey and conceptualization of performance can be used to study education systems in other countries, and readily adapted for other sectors, perhaps most obviously health. We see this as a further, albeit modest, methodological contribution of the paper.

2   Context: The Ghanaian education system

The Ghanaian education system is divided into basic, secondary, and higher education. Our focus in this paper is on basic education, which includes kindergarten, primary school, and junior high school and serves nearly 5.8 million students.[6] Basic education ends at grade nine, at which point students sit for the Basic Education Certificate Examination (BECE). The BECE is currently the main administrative source of information on learning among basic education students in Ghana.[7] In the paragraphs below we describe how the system is structured, first at national level and then at district level. Our aim is not to be exhaustive but to make it clear where district offices sit within the hierarchy of the country's education system, how they are organized and resourced, and their main tasks relating to basic education.

---

[5]Cilliers et al. (2022) focus on management practices under the same headings as the original WMS instrument (Bloom and Van Reenen, 2007) but, for logistical reasons, adopt a close-ended question methodology. Walter (2018) uses an instrument that is a close analogue of the D-WMS for schools and is focused primarily on how district offices manage schools.

[6]This enrollment figure was taken from the 2019–2020 Education Management Information System (EMIS) National Education Profile.

[7]The BECE includes English, any Ghanaian language offered at the exam candidate's school, social studies, science, mathematics, design and technology, information and communication technologies (ICT), French (optional), and religious and moral education.

National level The Ministry of Education holds the mandate for education service delivery across the 16 regions and 260 districts of Ghana. The most senior officer (bureaucrat) in the Ministry of Education is the chief director (see Figure 1). A total of 17 national agencies report to the chief director. The highest profile of these agencies is the Ghana Education Service. The GES, as the service is commonly known, is the largest employer in the education system and works through a subnational structure, with a network of regional and district education offices. Of the 16 remaining agencies, the National Teaching Council (NTC), National Council for Curriculum and Assessment (NaCCA), and National Inspectorate Board (NIB) interact with districts and schools to carry out functions such as licensing of teachers, curriculum development and assessment, and independent verification, respectively. Figure 1 also shows a multilayered political structure that sits adjacent to the bureaucracy, headed at national level by the minister of education.

The National Education Reform Secretariat (NERS) is a delivery unit that was established in 2019 with donor funding from the UK Foreign, Commonwealth & Development Office to support the 2018–2030 Education Sector Plan (ESP), an ambitious set of reforms focused on improving student learning outcomes. The ESP emphasizes three priorities: *access* to education, *quality* of education, and *management* across the education system. The NERS was set up to facilitate the achievement of these priorities and to support coordination among the 17 national agencies. During its early years (2019–2020), the NERS focused primarily on national-level targets. In August 2021, later than planned due to the impact of the COVID-19 pandemic, GES began to cascade aspects of the national delivery approach to the subnational level, mainly through the use of performance contracts with regions, districts, and schools (Bell et al., 2023).
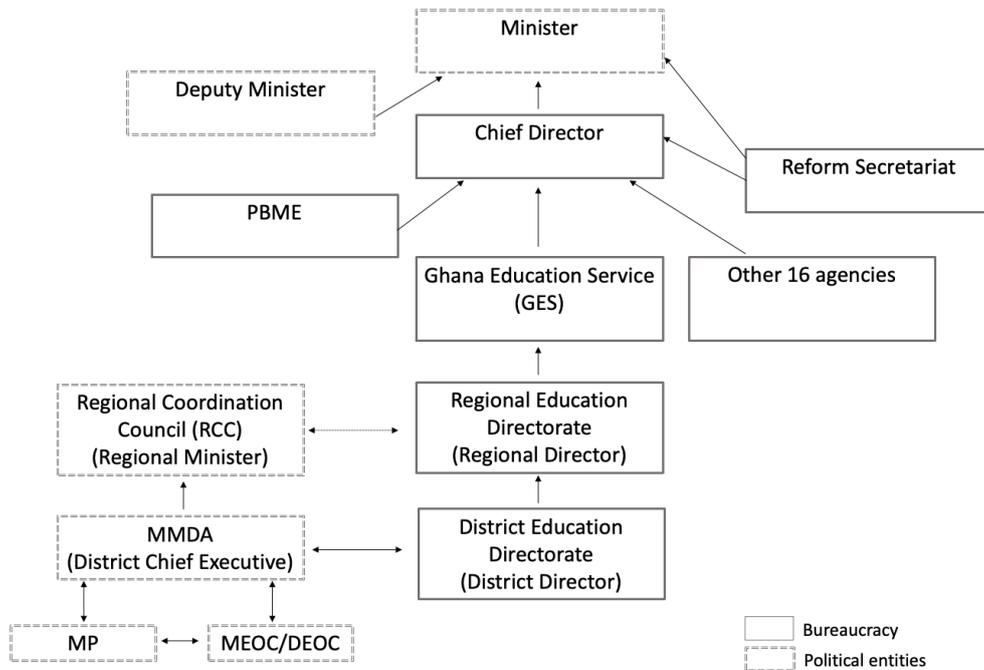
Figure 1: Structure of the Ghanaian Education System

District level District education offices are the main administrative layer between regions and schools. As Figure 2 shows, these offices are managed by a district director and four deputy directors (for finance and administration, planning and statistics, HR and training, as well as monitoring and supervision). Of the four deputy directors, those for HR and training and monitoring and supervision have primary responsibility for conducting school-facing tasks, and are supported by staff such as basic education officers, school health and education program (SHEP) coordinators, and school information support officers (SISOs). Each SISO is assigned a set of schools in a "circuit" and is tasked with monitoring and supporting these schools.[8] District offices report to their GES regional office but are also accountable to municipal/metropolitan district assemblies. Education oversight takes place via the municipal/district education oversight committee (MEOC/DEOC), which includes district assembly officials and members of Parliament (Ampratwum et al., 2019).

The main source of funding for district offices is their operational budget, a quarterly disbursement from GES at national level from funds allocated to the Ministry of Education from the Ministry of Finance. District assemblies have a mandate to provide school infrastructure and operate their own budget through the Common Fund (a central government transfer), the Ghana Education Trust Fund (GETFund), and a small amount of internally-generated funds (Government of Ghana, 2008).

District offices are responsible for carrying out a range of tasks, many of which involve direct engagement and interaction with schools. The list includes, but is not limited to, organizing and delivering in-service training (focused on leadership, literacy and numeracy, and health and hygiene), conducting monitoring visits to schools, supporting schools in developing school performance improvement plans (SPIPs), and conducting school performance appraisal meetings (SPAMs). Many of these tasks are associated with the Ghana Accountability for Learning Outcomes Project (GALOP), a US$ 218.7 million World Bank program running from 2019 to 2025 that supports a set of policies intended to improve access to, and the quality of, basic education. As part of the GALOP, the ministry is currently introducing the Ghana Education Outcomes Project (GEOP) in the north of Ghana. Under the GEOP, nonstate providers have been awarded contracts across six distinct geographic "lots" to provide support to basic schools and increase the enrollment of out-of-school children. These providers will start work in 2023. The GES historically has encouraged its district offices to manage their tasks using an Annual District Education Operational Plan (ADEOP), a planning document that (if completed) would set out key performance indicators and targets for the year.[9] The GALOP, GEOP, and nascent subnational cascade of the delivery approach is likely to impact how district offices are run, making it an interesting and important time to study management practices at this middle level of Ghana's education bureaucracy.

---

[8]SISOs were previously known as circuit supervisors, and are referenced as such in Figure 2.

[9]A scoping study for this project in May 2021 found substantial variation in the use of ADEOPs across five districts (Bell et al., 2023).
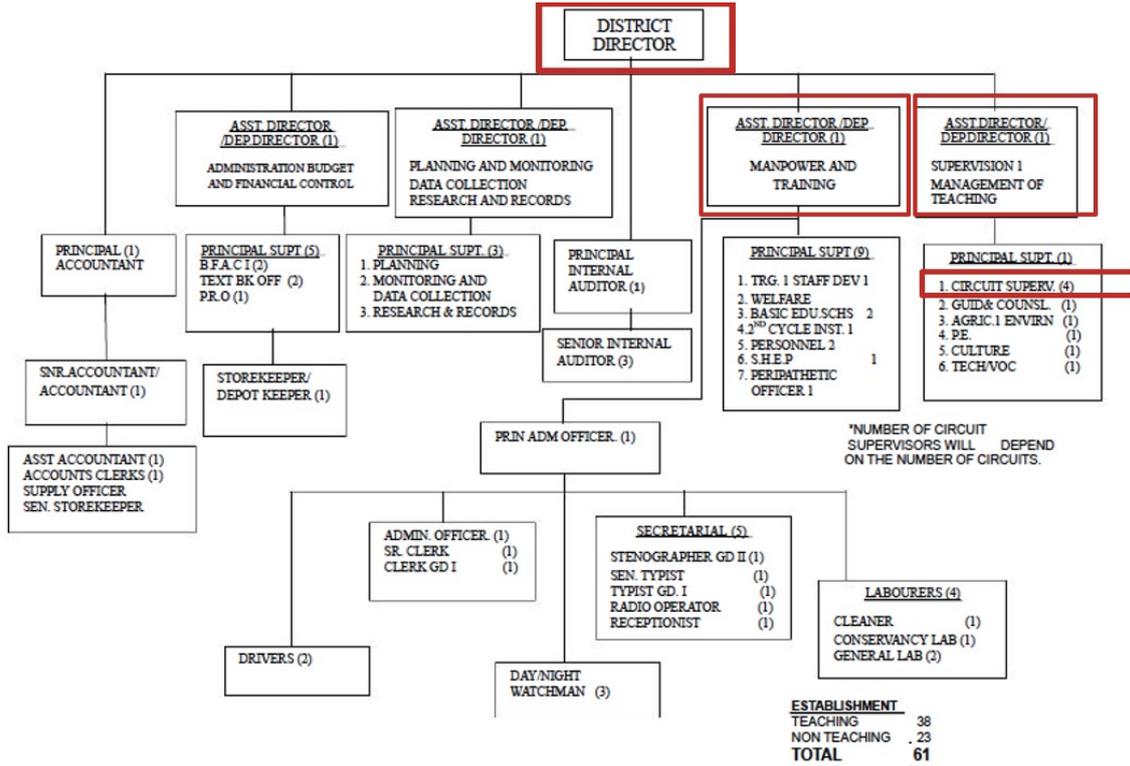
Figure 2: Organizational Chart for District Education Offices

3 Research questions and measurement approach

Our objective in this paper is to develop tools and conduct analysis to better understand mid-level bureaucracies, focusing on the specific context of the Ghanaian education sector. The six research questions that informed our instrument development and formed the basis of our empirical analysis are listed below.[10]

1. What management practices do district education offices use?

2. What factors explain (account for) variation in district office management practices?

3. How is performance at district level?

4. What is the relationship between district-level performance and district office management practices?

5. How is performance at school level?

6. What is the relationship between school-level performance and district office management practices?

Attempting to measure the management practices and performance of mid-level bureaucracies involves many choices and judgement calls—there are no universally agreed upon

---

[10]Figure A.1 illustrates these six question on a stylized "theory of change" diagram. The color coding follows the theory of change diagram in Williams et al. (2020).

metrics. In the remainder of this section we explain, drawing on relevant literature, *why* we chose our constructs, postponing a detailed discussion of *how* we operationalized and measured these constructs to Sections 5 to 7.

Measuring management   Our goal is to examine organizational approaches to implementation. To do so, we draw on the conceptual framework of Williams et al. (2020) who study delivery approaches, defined as: "an institutionalized unit or structured process within a government bureaucracy that aims to rapidly improve bureaucratic functioning and policy delivery by combining a set of managerial functions in a novel way to shift attention from inputs and processes to outputs and outcomes" (pp. 12–13). These authors review the literature and identify five functions commonly associated with delivery approaches.

The first function is *prioritization and target setting*. This refers to the establishment of a set of key priorities and objectives, measurable indicators to characterize progress against these objectives, and benchmark levels of performance to be achieved in a specified time period. The second function is *monitoring and use of data*, which refers to the establishment and execution of mechanisms to collect, report, and analyze information about the performance of divisions, districts, teams, schools, and/or individuals across the organization or sector. The third function concerns *accountability and incentives*, the establishment and execution of rewards and/or sanctions linked to performance—the "carrots and sticks" associated with delivery approaches. This could include a range of types of incentives: monetary incentives, the threat of firing or other formal career incentives, reporting through high-stakes meetings that create strong reputational concerns, "naming and shaming," or negative social perceptions. The fourth function is *problem-solving and adaptation*, which refers to the routinization of mechanisms of dialogue, coordination, and problem-solving across multiple individuals, divisions, or organizations that can improve performance through better sharing of information, performance data, and ideas. This could include horizontal collaboration and convening across teams, sectors, or actors, as well as the facilitation of "bottom-up" approaches that catalyze organizational learning across bureaucratic levels, through local problem-solving, adaptation, issue escalation, and policy feedback across the delivery chain. The fifth and final function is *leveraging political sponsorship*, which concerns the utilization and communication of high-level political backing for delivery. The audience for this signaling of political sponsorship can be the bureaucracy itself (to add pressure or legitimacy) or external stakeholders (to increase external pressure on the bureaucracy or to serve as a commitment device for government to hold itself accountable).

We built our measurement exercise around the first four of these functions. With the exception of leveraging political sponsorship, these functions describe the core of what most mid-level bureaucracies do (or could do) on a day-to-day basis to drive implementation.[11] Different organizations may place greater emphasis on some functions than others, but we feel that this list captures the key possibilities for delivery-related management at subnational level. Having chosen these four functions, we consulted a panel of experts—former presidents,

---

[11]We measure the level of engagement by national/subnational politicians with district offices but think of this as a contextual control, rather than a delivery function.

ministers, and bureaucrats, as well as current practitioners and consultants, with extensive first-hand experience of delivery approaches at national and/or subnational levels, to identify which management practices to measure under each heading.[12] The final list of 16 practices, chosen after a series of workshops with these experts, is shown in Figure A.2, together with a statement of what our survey aimed to "test" for each one.

In addition to measuring the level (of use) of these 16 management practices, we were also interested in the *style* of management approach in district offices. There are long-standing debates across disciplines concerning bureaucratic motivation and the appropriate management response. Some view bureaucracies as a sea of principal–agent problems and advocate practices under the accountability and incentives function.[13] Others take a more positive view, emphasizing bureaucrats' intrinsic motivation and professionalism and advocating practices under the problem-solving and adaptation function, in particular those relating to autonomy.[14] To contribute to these academic debates, and to provide evidence to inform the direction of the Ghanaian delivery approach cascade, we purposefully explore whether district offices make greater use of practices under the accountability function—what Honig (2022), appealing to McGregor (1960), might call "Route X," or *top-down* accountability—or the problem-solving function, which these authors refer to as "Route Y."

Measuring performance at district and school level   Our approach was to look for a range of metrics, from objective, verifiable measures of inputs to more subjective assessments of how the district office/school was functioning. The aim was to span both "thin and thick description," as advocated by Honig and Pritchett (2019).

We measured four dimensions of district office performance. The first was *district director and staff activity intensity*. The construct we were after here is "effort." We identified a set of six broad categories (and activities within them), using focus groups in five district offices in May 2021. The second dimension was *district task completion*, a direct office output. The construct was whether the district had completed school-facing delivery tasks over a specified time frame. We wanted to find a circumscribed set of tasks which reflect effort and speak to school-facing delivery, and which could be independently verified. We elected to focus on tasks prescribed under the government's Education Sector Plan. Working closely with the Ministry of Education and NERS, we chose a list of 10 tasks for which head teachers (or in two cases SISOs) could comment on completion. The third dimension was *staff understanding*. Here we pivoted to thick(er) description and attempted to measure the potential drivers of office output. Since prioritization is one of the delivery functions and formed part of the delivery cascade, the construct we sought to assess was whether district staff share the same understanding of district priorities as their director. An office in which

---

[12]For a list of panel members see https://educationcommission.org/delivered-high-level-advisory-group.

[13]See, for example, classic contract theory texts in economics such as Laffont and Martimort (2002). Also relevant here is the sizeable literature in economics evaluating (without necessarily advocating) accountability and incentive interventions in education and beyond (Finan et al., 2017; Glewwe and Muralidharan, 2016).

[14]See, for example, the arguments in Andrews et al. (2017) and Honig (2022), as well as a large body of literature in education on distributed leadership, organizational learning, and improvement science. Rather than attempt to summarize this large body of work here, we point the interested reader to Williams et al. (2020).

this is not the case may find it hard to perform well. Relatedly, to foster performance it seems likely that directors and their staff need to share a common understanding of who is responsible for what. Building on work by Adelman and Lemos (2021), we identified an additional construct: whether there is a common understanding of roles in the education system. We operationalized this by identifying seven roles following discussions with education stakeholders in May 2021. The final dimension was *staff attitudes and satisfaction*. Here, we wanted to get a sense of how staff felt about the office culture and their overall satisfaction with their jobs. We followed the literature and chose validated item banks used in prior work on bureaucracies and schools (Leaver et al., 2021; Rasul et al., 2021).

To investigate whether district education offices influence what schools do, we also measured four dimensions at school level. Two of these, *head teacher activity intensity* and *head teacher attitudes and satisfaction*, were chosen to be analogs of our district-level measures. The remaining two dimensions are *teacher absenteeism* and *classroom practice*. Both are commonly used measures in the economics literature on schools (Bold et al., 2017; Chaudhury et al., 2006). For reasons noted in the introduction, we do not include a measure of student learning.

4   Sampling

With the exception of a small number of variables obtained from the Ministry of Education, all of our data came from a survey conducted in collaboration with the Institute of Statistical, Social and Economic Research (ISSER) at the University of Ghana between May and July 2022. This short section describes the sampling frames, sampling protocols, and sample sizes of the resulting data sets.

The survey was designed to generate two data sets: one that is nationally representative and another that provides a baseline for the GEOP. The sampling frame for the former was all 260 districts in Ghana, from which we randomly selected 174 districts using region, BECE score (district above or below median), and basic education student enrollment (district above below or median) as strata.[15] These sampled districts are highlighted in gray in Figure 3 Panel (a).[16] Within each district, we surveyed the district director, the two deputy directors that carry out the majority of school-facing tasks (the deputy director of HR and training and deputy director of monitoring and supervision), two SISOs, and six head teachers. Our sampling frame for head teachers was a national list of GALOP schools, from which we randomly selected six schools per sampled district using GALOP score (above or below median) as strata. We asked directors to tell us the name of the circuit associated with each sampled school in their district, which gave us a sampling frame for SISOs.[17] From this list

---

[15]We were given access to Ministry of Education administrative data on BECE scores and student enrollment (2019–2020), aggregated at the district level, for the purposes of sampling.

[16]The different shades of gray indicate different geographic zones, which we used as controls in our subsequent regression analysis.

[17]Unfortunately, at the time of the survey, the Ministry of Education could not provide a list of circuits mapped onto both schools and districts. We used this protocol to ensure that we interviewed the SISOs working in the same circuits as our sampled schools.

Table 1: Number of Survey Respondents

| | Sample | |
| | Rep. | GEOP |
	(1)	(2)
Director	174	27
Deputy director	341	54
SISO	348	54
Head teacher	1035	1261

Notes: The two samples are not mutually exclusive, as some of the 27 GEOP districts are also included in the representative sample.

of circuits, we then randomly selected two SISOs per district. The number of respondents in this nationally representative data set is shown in Column (1) of Table 1.

The sampling for the second data set was dictated by the GEOP. The sampling frame was the list of GALOP schools in six geographic "lots," mainly in the north of Ghana. For each GEOP lot, we randomly sampled an agreed number of schools (sufficient to give the required number of treatment schools, plus an equal number of control schools, and some replacements) using an exclusion restriction based on the GALOP score to maximize cross-lot comparability. This gave us a sample of 1,261 schools located across 27 districts. These districts are highlighted in gray in Figure 3 Panel (b). Within each district, we surveyed the same categories of respondent in district offices (the director, two deputies, and two SISOs) and the head teachers of all selected schools.[18] The number of respondents in this GEOP sample is shown in Column (2) of Table 1.

---

[18]We used the same protocol to identify circuits: we chose six schools at random from the list of selected schools for the district (typically many more than six) and asked directors to tell us the name of the circuit associated with each of these six schools. We used this list of circuits to select two SISOs per district.

☐ No (86)	☐ No (233)
☐ Yes, North zone (42)	☐ Yes, North zone (20)
☐ Yes, Middle zone (80)	☐ Yes, Middle zone (7)
☐ Yes, South zone (52)	

(a) Representative sample   (b) GEOP sample
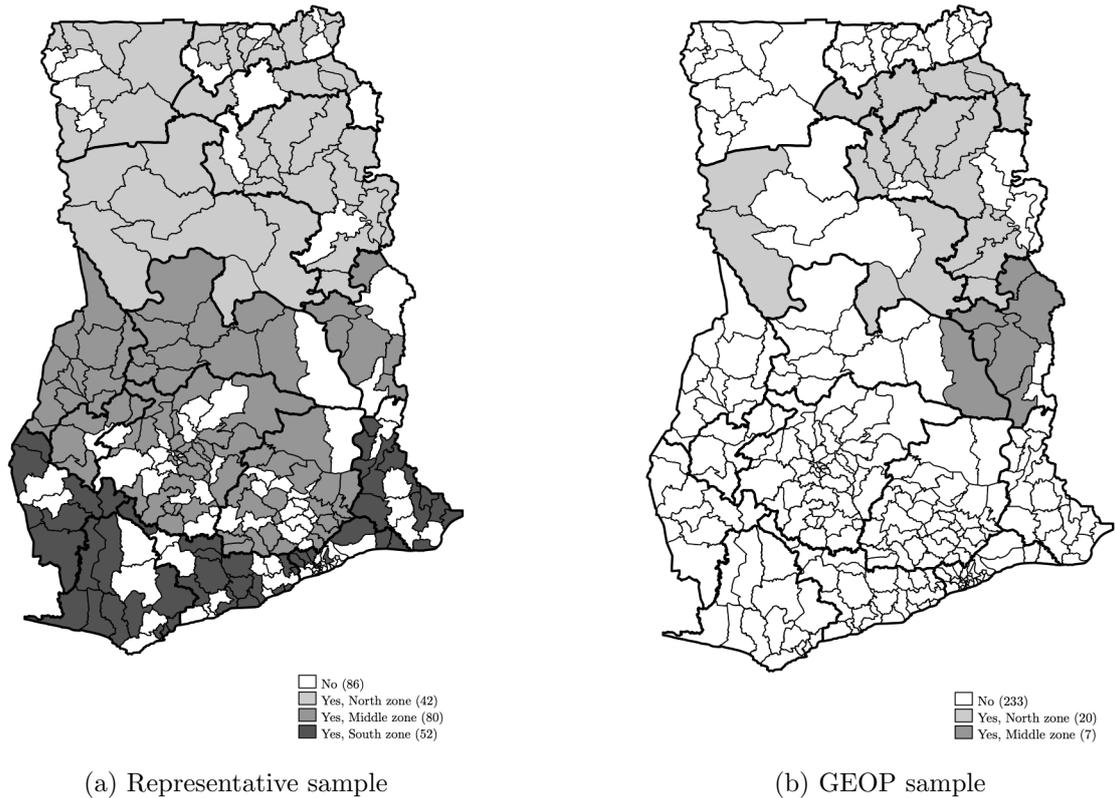
Figure 3: Map of Ghana Showing the Districts Sampled

5 District office management practices

5.1 What management practices do district offices use?

This subsection addresses our first research question: what management practices do district offices use? We describe how we measure management and then present descriptive statistics.

Our objective was to measure the use of 16 different management practices identified by a panel of experts as underpinning four functions commonly found in delivery approaches. To do this, we adopted the methodology developed by the World Management Survey team (Bloom and Van Reenen, 2007). Specifically, we asked each district director a series of open-ended questions that were designed to elicit the thoroughness with which his/her district office uses each management practice. The responses to these open-ended questions were post-coded on a four-point scale using a detailed scoring rubric. The general principle underlying this rubric (which purposefully echoes the WMS) is that a score of 1 is given when the practice does not exist, 2 is given when a practice exists but in an *ad hoc* or very informal form, 3 when the practice happens somewhat systematically, but adoption could be more thorough and with more structure, and the maximum score of 4 when a practice is carried out systematically.[19] This post-coding exercise gives us a set of raw scores. We use

---

[19]The open-ended questions and post-coding scoring rubric for all 16 practices is provided in Supplementary Appendix A.1. We used a four-point scale based on advice from the WMS team, who noted that their five-point scale can create a tendency for respondents to "hang in the middle" with a score of 3. It is worth reiterating that although we used the WMS methodology, most of our questions are different from those in

Table 2: District Office Management Practices

	Mean	SD	P25	P50	P75
Overall management index	2.64	0.43	2.28	2.66	2.96
Prioritization and targets sub-index	2.55	0.61	2.10	2.60	3.00
1. Priority setting	3.16	0.74	2.50	3.50	4.00
2. KPIs and targets	2.49	0.88	2.00	2.50	3.00
3. Strategic planning	1.93	0.94	1.00	2.00	2.50
4. Target interconnection	2.34	0.87	1.50	2.50	3.00
5. Priorities, KPIs, and target comms	2.83	0.89	2.00	3.00	3.50
Monitoring and use of data sub-index	2.50	0.50	2.12	2.50	2.88
6. Data collection	2.60	0.77	2.00	2.50	3.00
7. Data management and usage	2.32	0.50	2.00	2.50	2.50
8. Performance monitoring (district-level)	2.63	0.73	2.00	2.50	3.00
9. Performance monitoring (staff-level)	2.44	0.69	2.00	2.50	3.00
Accountability and incentives sub-index	2.67	0.60	2.33	2.67	3.17
10. Accountability structure	2.32	0.96	1.50	2.50	3.00
11. Accountability meetings	2.70	0.82	2.00	2.50	3.50
12. Rewards and sanctions	3.00	0.69	2.50	3.00	3.50
Problem-solving and adaptation sub-index	2.82	0.56	2.38	2.88	3.25
13. Feedback loops	2.91	0.85	2.00	3.00	3.50
14. Deep dives	2.75	1.02	2.00	3.00	4.00
15. Autonomy	2.81	0.90	2.00	3.00	3.50
16. Innovation	2.81	0.78	2.00	2.50	3.50

Notes: For each of the 16 management practices, the mean is the average response for the 174 districts and SD is the standard deviation in this district mean. Each sub-index mean is the average for the listed management practices. The overall index mean is the average of the four sub-indices. The values are weighted using design weights.

these scores for the descriptive statistics presented in this section but switch to standardized indices in subsequent regression analysis. We refer to the overall index aggregating the use of these 16 management practices as our *management index*, and we refer to the sub-indices aggregating use of the practices under each function by the name of that function (e.g. the *prioritization sub-index*, the *monitoring sub-index*, etc.).[20]

Table 2 reports summary statistics by district. The mean of the management index is 2.64, indicating only moderate use of these management practices in district offices. To give a sense of variation across practices, Table 2 also presents summary statistics by sub-index. We discuss each in turn.

The mean of the prioritization sub-index is 2.55. The score of 3.16 for *priority setting* is high and indicates that directors are typically able to articulate district priorities, albeit with some shortcomings.[21] Figure 4 shows what directors prioritize. Student learning is by far the most common category, listed by just under a quarter of directors. Student access, school

---

the WMS survey instruments, reflecting our distinct focus on organizational approaches to implementation.

[20]We compute the raw sub-indices by averaging raw scores across the relevant practices, and we compute the raw overall index by averaging the four sub-indices. We compute the standardized sub-indices by averaging the raw scores and then converting into a normalized z-score and the standardized overall index by averaging the four z-scores and then standardizing again.

[21]The main reasons the enumerators gave a score of 3 instead of the maximum of 4 were excessive length of the list of priorities or the use of broad statements lacking a connection to an education outcome or policy.

facilities, and teacher HR are the next three most common categories, with each mentioned by around 12 percent of directors. The scores for the other management practices under the prioritization function are lower than those for priority setting, especially for *strategic planning*, at just 1.93. This low score indicates that, although planning documents may exist, they are not typically linked to district priorities. The interquartile range of the prioritization sub-index is 0.9, revealing substantial variation across districts. In Figure 5 Panel (a), we rank all 174 districts from lowest to highest score on the prioritization sub-index and plot the corresponding sub-index score on the y-axis. Districts are distributed across all the possible scores, from just above the minimum of 1 all the way up to just below the maximum of 4.
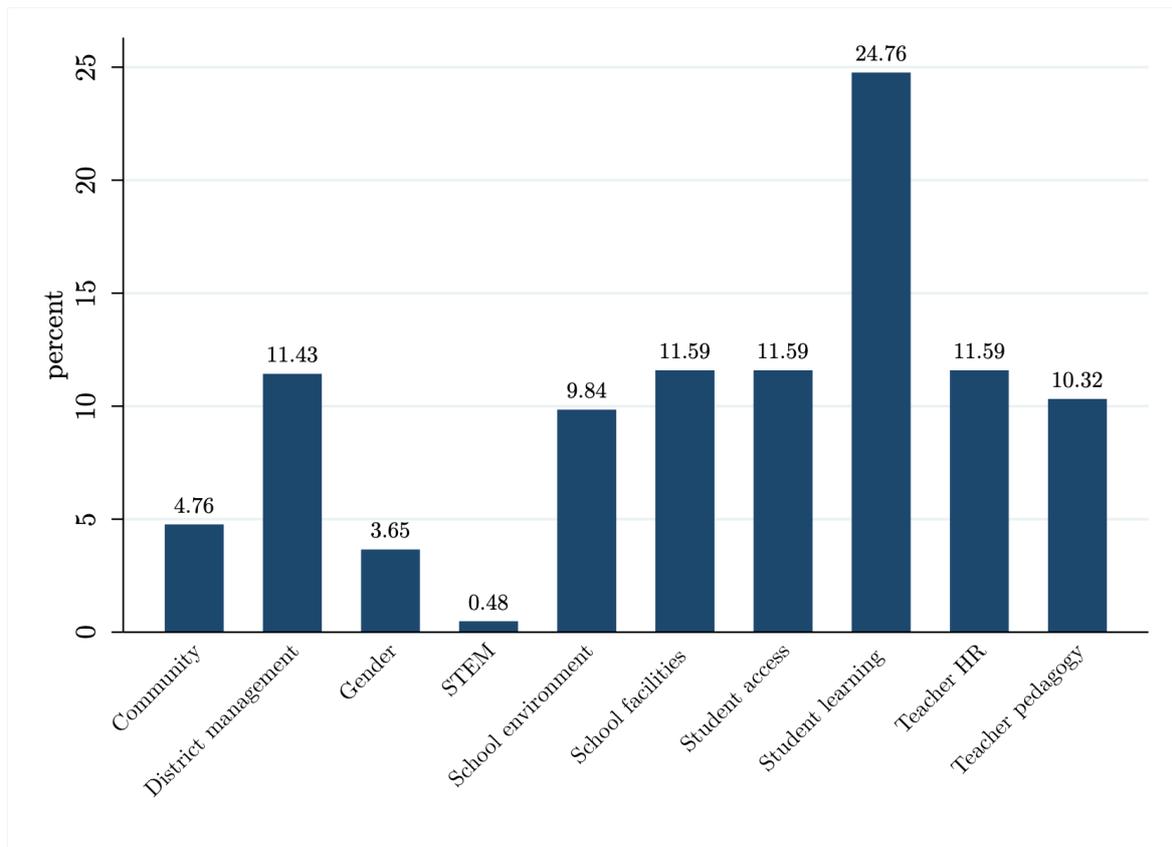
Figure 4: District Priorities, as Stated by the District Director

(a) Prioritization and target-setting

(b) Monitoring and use of data

(c) Accountability and incentives

(d) Problem-solving and adaptation
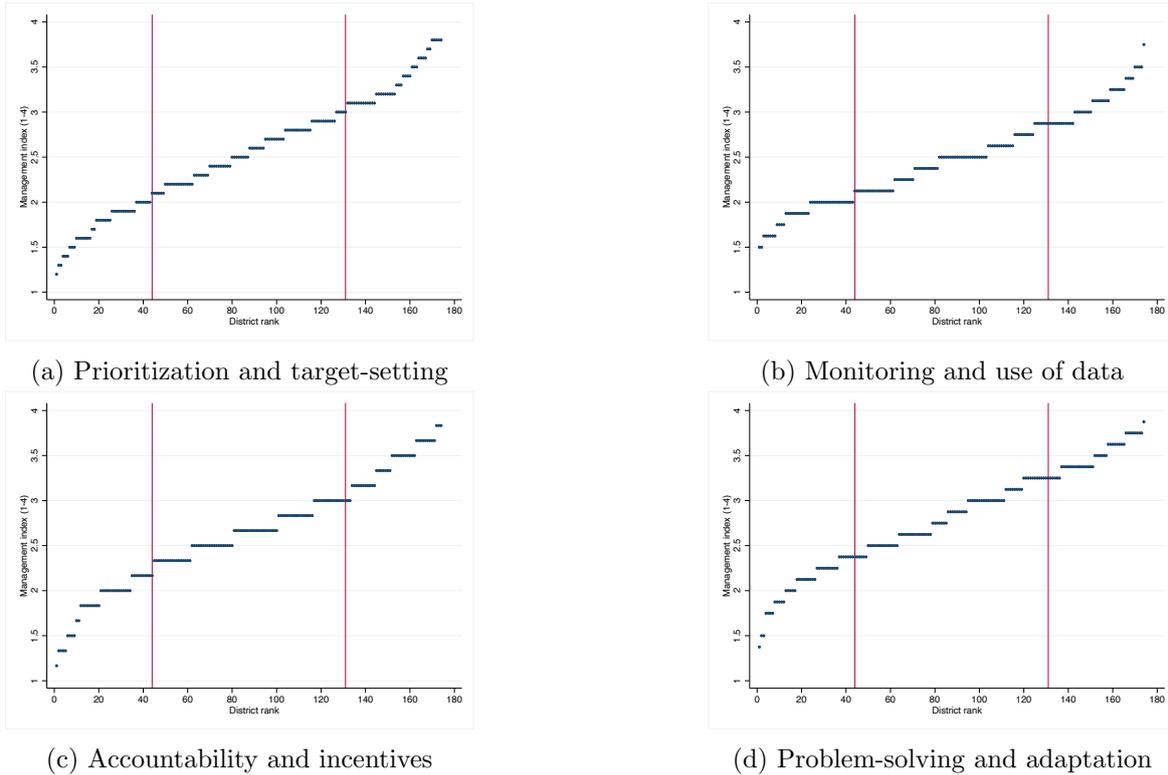
Figure 5: District Management Sub-Indices, by District Rank

The mean of the monitoring sub-index is 2.50, the lowest of the four sub-indices. In contrast to prioritization, there is less variation across management practices; the means of all four practices are between 2 and 3. It is notable that *data collection* appears to be more systematic than *data management and usage*. Figure 5 Panel (b) shows that there is also less variation in the monitoring function across districts. Scores range from 1.5 to 3.5 (with one exception) and are generally more compressed.

The mean of the accountability sub-index is 2.67. The score of 3.00 for *rewards and sanctions* is high and implies that on average, incentive systems are more formalized than practices establishing lines of accountability (*accountability structure* has a mean of 2.32) or indeed practices to monitor staff performance (*performance monitoring at staff level* has a mean of 2.44). Figure 5 Panel (c) shows that there is substantial variation across districts, with scores from close to 1 up to almost 4.

The mean of the problem-solving index is the highest of the four sub-indices at 2.82. There is limited variation across the underlying management practices and all four scores indicate some structure. Figure 5 Panel (d) shows that there is variation across districts, but this is largely at the top end; just 12 districts have a problem-solving score below 2.

15
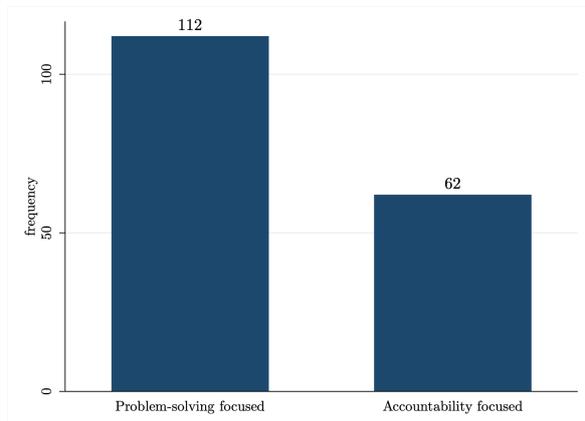
(a) Count of accountability-focused districts   (b) Variation in regional mean
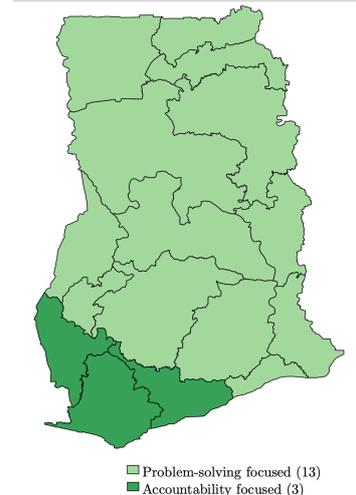
Figure 6: Accountability Focus of District Office Management Practices

Although the mean problem-solving sub-index is higher than the mean accountability sub-index, district management practices relating to problem-solving are not always more structured and developed than practices relating to accountability. We construct an indicator variable, *accountability focus*, that is coded to 1 if a district has an accountability-focused management style, in the sense that the accountability sub-index is greater than the problem-solving sub-index, and 0 otherwise. Figure 6 Panel (a) shows that there are 62 districts (close to 36 percent) where the accountability sub-index is higher than the problem-solving sub-index (seven districts have the same values for both sub-indices and are counted as "problem-solving focused"). In Panel (b), we show how this aspect of management style varies across the 16 regions of Ghana. For each region, we compute the average of the district-level accountability focus indicator. Dark green shading indicates a regional average above 0.5 and light green shading the reverse. There is a clear geographic pattern: in three southwestern regions (close to Accra but not the Central Accra region itself), there is a greater focus on management practices related to accountability than on those related to problem-solving, while in the rest of Ghana it is the reverse.

5.2 What explains district office use of management practices?

Having documented substantial variation in the use of management practices across district offices, we turn to our second research question and explore the contextual factors that might be driving these differences. We focus on three categories: characteristics of the district directors; characteristics of district offices; and "external" factors—the extent of financial, material, and/or technical support received by offices from nongovernmental and/or international organizations in the past year, and the level of political sponsorship from national, regional, and/or district-level politicians.

Table 3 provides summary statistics and highlights two important points. First, the average district office faces significant resource constraints. Almost all directors (98 percent)

reported delays of at least a month in the disbursement of their quarterly budget from the Ministry of Education. On average, offices receive their quarterly budget slightly over six months late. Not all directors were willing/able to tell us the proportion of their allocated budget for the last financial year that they actually received. However, among the directors who did answer this question, the average proportion received was only just over 50 percent. Although offices appear to have regular access to a telephone and internet connection (for 4.6 out of 5 working days on average), there is much more limited access to vehicles and fuel. On average, out of every 10 officers, fewer than four have regular access to these important resources.[22] Second, external engagement in districts is quite high. On average, just over 70 percent of districts had received some form of external support from a nongovernmental or international organization in the past year. Although visits from the minister himself may have been infrequent, three-quarters of directors agree with the statement that MPs and local politicians visit their office regularly.

To explore which (if any) contextual factors explain variation in the use of district office management practices, we run a series of bivariate and multivariate OLS regressions. In these specifications, our dependent variable is always the (standardized) district office management index, and our independent variables are selected from the list in Table 3.[23]

Table 4 presents results. In Panel A, the bivariate and multivariate regressions tell a similar story: a higher management index is strongly associated with external factors. In the final column, the partial effects of external support and regular engagement by MPs are 0.54 and 0.51 standard deviations, respectively.[24] Panel B repeats this analysis for the sub-sample of 129 districts where we have budget information. The coefficient on the proportion of budget received is positive but statistically insignificant in both the bivariate and multivariate specifications. Some of the coefficients in Column (13) change from Panel A to Panel B, but the main finding that the district office management index is positively associated with external factors persists.

---

[22]As Table 3 shows, schools are predominantly in rural locations, implying that access to a vehicle and adequate fuel is necessary for staff to conduct field visits to schools.

[23]Our choice of survey questions on district context was informed by a literature review. We selected variables for inclusion in the regression analysis via machine learning.

[24]The partial effect of access to a vehicle/fuel is negative. Since the coefficient is small in magnitude and only significant at 10 percent, we refrain from interpretation

Table 3: District Context

	Mean	SD	P25	P50	P75	Obs
Director						
1[Male]	0.61					174
Experience in GES (years)	29.16	4.72	25.75	28.00	32.58	174
Office						
Number of staff	46.91	12.53	40.00	46.00	52.00	174
Number of circuits	7.38	2.47	6.00	7.00	9.00	174
Number of schools	85.49	42.79	56.00	79.00	103.00	174
1[Quarterly budget delayed]	0.98					174
Quarterly budget delay (months)	6.08	2.96	5.00	6.00	6.00	170
Proportion of annual budget released	0.53	0.28	0.32	0.50	0.76	129
Proportion schools in rural locations	0.64	0.32	0.50	0.67	0.83	174
Access phone/internet (days, out of 5)	4.64	0.55	4.50	5.00	5.00	174
Access to vehicle/fuel (officers, out of 10)	3.72	2.28	2.00	3.50	5.00	174
Staff skills index	6.04	1.69	5.00	6.00	7.00	174
External						
1[External support]	0.71					174
1[Minister visits regularly]	0.09					174
1[Regional minister visits regularly]	0.19					174
1[MP visits regularly]	0.76					174
1[District assembly member visits regularly]	0.74					174
1[Aligned district assembly priorities]	0.68					174

Notes: For each variable, the mean is the average response for the 174 districts and SD is the corresponding standard deviation. The budget delay (months) variable is missing for four directors who indicated no delay. The "budget released" variable is missing for 45 directors who were not willing/able to answer this question. The skills index is the average of four questions asking how many officers (out of 10) have basic ICT skills (e.g. can use a laptop, MS Office software). External support is a binary variable coded 1 if the district received either financial, material, or technical support from nongovernmental or international organisations over the past year, and 0 otherwise. District directors were asked how strongly they agree with the statement that the minister (regional minister, MP, or district assembly member) regularly visits their office, with answers on a standard five-point Likert scale. The "visit" indicator variables are coded 1 if a respondent chose "agree" or "strongly agree," and 0 otherwise. The alignment indicator variable is coded 1 if the director believes that the district assembly has district-level priorities and these priorities are aligned with the director's priorities. The values are weighted using design weights.

Table 4: District Context and District Office Management Practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: Management index													
1[Male]	-0.267*												-0.220
	(0.152)												(0.151)
Director experience (yrs)		-0.012											-0.008
		(0.016)											(0.015)
Access to vehicle/fuel (officers)			-0.055*										-0.059*
			(0.033)										(0.032)
Staff skills index				-0.017									-0.036
				(0.045)									(0.043)
Proportion schools in rural areas						-0.129							-0.063
						(0.234)							(0.233)
1[External support]							0.498***						0.543***
							(0.161)						(0.163)
1[Minister visits regularly]								-0.158					-0.275
								(0.256)					(0.258)
1[Regional minister visits regularly]									0.055				-0.041
									(0.191)				(0.193)
1[MP visits regularly]										0.628***			0.510***
										(0.168)			(0.177)
1[District assembly member visits regularly]											0.031		-0.066
											(0.171)		(0.172)
1[Aligned district assembly priorities]												0.262	0.183
												(0.159)	(0.169)
Observations	174	174	174	174	129	174	174	174	174	174	174	174	174
Panel B: Management index (subsample)													
1[Male]	-0.215												-0.182
	(0.171)												(0.174)
Director experience (yrs)		-0.001											0.001
		(0.018)											(0.017)
Access to vehicle/fuel (officers)			-0.067*										-0.083**
			(0.040)										(0.040)
Staff skills index				0.027									-0.032
				(0.051)									(0.052)
Proportion of budget received					0.341								0.424
					(0.302)								(0.298)
Proportion schools in rural areas						-0.348							-0.346
						(0.254)							(0.262)
1[External support]							0.374**						0.460**
							(0.184)						(0.189)
1[Minister visits regularly]								-0.386					-0.485*
								(0.274)					(0.286)
1[Regional minister visits regularly]									-0.105				-0.178
									(0.233)				(0.245)
1[MP visits regularly]										0.432**			0.351
										(0.199)			(0.216)
1[District assembly member visits regularly]											-0.001		-0.060
											(0.201)		(0.201)
1[Aligned district assembly priorities]												0.257	0.254
												(0.184)	(0.201)
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129

Notes: In Panel A, the dependent variable is the standardized management index. Panel B restricts the sample to the sub-sample of 129 districts for which we have data on the proportion of budget received. The explanatory variables are described in the notes to Table 3. The estimates are significant at the *10%, **5%, and ***1% levels.

6 Performance at district level

In this section, we explore performance at district level and the relationship between this performance and district office management practices. Rather than presenting analysis by question, over the next four subsections we discuss different dimensions of performance in turn. Within each subsection, we summarize how we measure the dimension of performance, present descriptive statistics (giving a sense of cross-district variation), and finally report regression results. We postpone discussion of these results, including multiple hypothesis test corrections, to Section 8.

6.1 Director and staff activity intensity

We set out to measure the intensity with which directors and their staff (deputy directors and SISOs) undertake activities in six broad categories: interacting with superiors, meeting with subordinates, field visits to schools, interacting with political actors, engaging with community members, and office administration.[25] These categories were selected after scoping fieldwork in May 2021 and are intended to capture the main ways in which directors and staff could be spending their time. For each category, we asked respondents a series of questions about how often they undertook activities (within that category) over the past three months. Responses to each question were given on a five-point scale: 1 never, 2 once or twice, 3 at least monthly, 4 at least weekly, and 5 daily. The statistics presented below are based on the average of these responses within categories. We also wanted to gain a deeper understanding of field visits to schools. After an initial question about how often respondents undertook these visits over the past three months, we asked a series of 17 follow-up questions about the activities undertaken *during* these visits—for example, how often did the respondent sit in on a class, observe the school logbook, and so on—with answers on the same five-point scale.[26] In the analysis below, we compute averages for six teaching-related activities and the 11 remaining "other" activities separately.

Table 5 shows that the most frequent activity category for directors and their deputies is internal office meetings, which happen more often than once a week on average. For these job levels, school visits occur with more moderate intensity, between monthly and weekly. The reverse is true for SISOs. On average, SISOs visit schools more than once a week and spend less time in internal meetings.

A natural question is whether more time spent on one activity "crowds out" time for other activities. Table 6 suggests that this is not the case. For directors, there is no evidence of a correlation between the intensity of meetings with superiors/subordinates and field visits to schools; in fact there is a *positive* correlation between engagement with political actors and community members and school visits. Intensity of (strategic) office administration is also positively correlated with visits to schools.

---

[25]Office administration refers to strategic planning and budgeting, rather than "day-to-day" paperwork.

[26]The full list of questions is provided in Supplementary Appendix A.2.

Table 5: Activity Intensity, by Job Level

	Director		Dep. Director		SISO	
	Mean	SD	Mean	SD	Mean	SD
Interacting with superiors	2.89	0.67	4.76	0.52	3.86	0.74
Meeting with subordinates	4.13	0.71	4.16	0.89	.	.
Field visits to schools	3.69	0.72	3.10	1.06	4.18	0.69
Interacting with political actors	2.48	0.77	1.64	0.52	1.50	0.54
Engaging with community members	2.67	1.00	2.23	0.84	2.57	0.83
Office administration	2.74	0.73	2.17	0.73	2.12	0.59
Observations	156		300		329	

Notes: In the first cell, we first obtain a district-level variable by averaging the director's responses to two questions about interactions with superiors over the past three months (both measured on a five-point scale: 1 never, 2 once or twice, 3 at least monthly, 4 at least weekly, and 5 daily) and we then compute the average over 174 districts to obtain the district mean. SD is the standard deviation in this district mean. In the third cell, we obtain a district-level variable by averaging the two deputy directors' responses to one question about superiors and then computing the average over 174 districts. SD is again the standard deviation in this district mean. In the fifth cell, we do the same for the two SISOs but use two questions about superiors. We follow similar steps in the remaining rows. However, we have one question each for meeting with subordinates (excluding SISOs who do not answer this question), field visits to schools, and engaging with community members. We have two questions for interacting with political actors, and we have three questions for office administration. The values are weighted using design weights.

Table 6: Correlation Between Activity Categories, Director

	Superiors	Subordinates	Visiting schools	Political actors	Community members	Office admin
Superiors	1.000					
Subordinates	-0.003	1.000				
Visiting schools	-0.077	0.041	1.000			
Political actors	0.085	-0.082	0.161*	1.000		
Community	0.191*	0.126	0.217**	0.417***	1.000	
Office admin	0.103	-0.267***	0.224**	0.336***	0.152	1.000

Notes: This talbe presents the bivariate correlations between the district-level variables for the director from Table 5. The estimates are significant at the *10%, **5%, and ***1% levels.
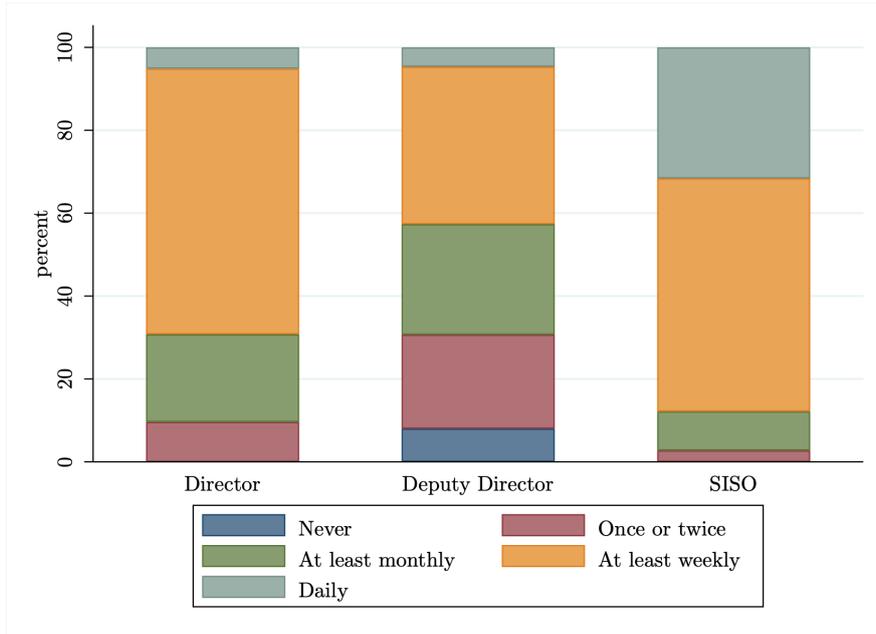
Figure 7: Field Visits to Schools, by Job Level

Figure 7 focuses on these field visits to schools, giving a sense of variation in the intensity of this activity. Most SISOs undertake school visits at least weekly, but a sizeable minority are more active (32 percent) and some are less active (13 percent). The existence of the latter group is worrying, since school visits are the primary aspect of the SISO role. A surprisingly high proportion of directors undertake school visits at least weekly (64 percent), but there is also a sizeable group of directors who do so only once or twice a quarter. There is most heterogeneity among deputy directors, with almost the same number of responses for at least weekly, at least monthly, and only once or twice per quarter, and eight percent of deputy directors *never* visit schools.

Table 7 shows the frequency with which respondents undertake different activities during these visits. Of the three job levels, SISOs undertake all activities with the highest frequency (in every row, the mean is highest in the SISO column). Notably, the two most frequent activities are the same for all job levels: observing the school logbook, and conducting spot checks. For SISOs, observation of school logbooks occurs weekly, on average, with spot checks undertaken between monthly and weekly (means of 4.03 and 3.81, respectively). For directors, observation of school logbooks and spot checks both occur between monthly and weekly on average (means of 3.49 and and 3.46, respectively). The frequencies for deputy directors are similar but a little lower.

We now explore whether district office management practices are associated with the intensity with which directors and their staff undertake activities, focusing on field visits to schools. We estimate the following respondent-level specifications

$$y_{idz} = \beta_0 + \beta_1 Mgmt_{dz} + X'_{dz}\nu + R'_{idz}\gamma + \alpha_z + \epsilon_{idz} \qquad (1)$$

Table 7: Activity Intensity During Field Visits to Schools, by Job Level

	Director		Dep. Director		SISO	
	Mean	SD	Mean	SD	Mean	SD
Panel A: Teaching-focused activities						
Classroom observation	2.56	0.98	2.45	1.09	3.35	0.91
Looking at sample of student work	2.62	1.08	2.50	1.09	3.32	0.89
Reviewing lessons plans	2.74	1.03	2.64	1.13	3.46	0.79
Coaching or mentoring session	2.33	1.04	2.38	0.94	2.90	0.95
Providing in-service training	1.88	0.86	2.06	0.85	2.38	0.89
Panel B: Other activities						
Collecting data	2.49	1.18	2.87	0.99	3.43	0.95
Observing school logbook	3.49	0.84	3.07	1.02	4.03	0.74
Spot checks	3.46	0.82	2.94	1.00	3.81	0.85
Discussing teacher career development	2.31	0.96	2.40	0.94	2.72	1.01
Supporting teachers in filling out their forms	1.34	0.67	1.74	0.75	1.92	0.78
Discussing school progress with PTA/SMC	1.96	0.65	1.93	0.61	2.27	0.57
Support, supervision or implementation of SPIP	1.81	0.83	1.81	0.77	2.27	0.79
Monitoring and audit of the school capitation grant	1.54	0.77	1.43	0.66	1.80	0.69
Related to COVID-19	2.65	1.17	2.44	1.24	3.07	1.36
Managing teacher disputes	1.90	0.93	2.08	0.83	1.98	0.88
Observations	156		276		329	

Notes: The means and SDs are computed as in Table 5. All rows are based on responses to a single question, except reviewing lesson plans and discussing school progress with the PTA/SMC which are based on two questions each. The respondents who joined the office in the three months immediately prior to the survey are excluded. This means that we drop 18 directors (10 percent), 41 deputy directors (12 percent), and 19 SISOs (5 percent). A further 24 deputy directors are excluded because they never visit schools. The values are weighted using design weights.

$$y_{idz} = \beta_0 + \beta_1 \, Targets_{dz} + \beta_2 Mon_{dz} + \beta_3 Acc_{dz} + \beta_4 PS_{dz} + X'_{dz}\nu + R'_{iz}\gamma + \alpha_z + \epsilon_{idz} \quad (2)$$

$$y_{idz} = \beta_0 + \beta_1 Acc \, focus_{dz} + X'_{dz}\nu + R'_{iz}\gamma + \alpha_z + \epsilon_{idz}, \quad (3)$$

where y_{idz} is a measure of activity intensity for respondent i in district d located in geographic zone $z \in \{$north, middle, or south$\}$. We use three different dependent variables: intensity of field visits to schools (a single question); intensity of teaching-focused activities during field visits to schools (an average of six questions); and other activities during field visits to schools (an average of 11 questions).[27] The management variables are all measured at district level. The coefficient of interest in specification (1) is β_1, the partial effect of the overall management index on activity intensity. Specification (2) decomposes this effect down into the separate functions, with $\beta_j, j = 1, ...4$, giving the partial effect of function j on activity intensity, holding the other three functions constant. Specification (3) replaces the management index with our indicator for accountability focus. Here, β_1 gives the partial effect of a more accountability-focused than problem-solving-focused management style on

---

[27]Responses to the underlying questions are measured on an ordered five-point scale. When we average over multiple questions, our dependent variables become continuous. For ease of comparison, we use a linear model rather than an ordinal nonlinear model in all specifications.

activity intensity.

The vector X_d contains district-level controls.[28] To capture district context, we include the district average GALOP score—a multidimensional measure constructed by the World Bank (at school level) to select disadvantaged primary schools into the GALOP. The key dimensions are BECE score and a measure of social deprivation. Following Rasul et al. (2021), we also include four "organizational" controls, namely the number of office staff, a staff skill index, and two types of district facilities indices that capture staff access to phone/internet and a vehicle/fuel. In line with Bloom and Van Reenen (2007) and Rasul et al. (2021), we also include an extensive list of "noise" controls: characteristics of the district director (gender, education, tenure, public service motivation, and Big Five personality index), the time of day that the interview was conducted, and the reliability of the survey as coded by the interviewer. The vector R_i contains two respondent-level controls: gender and age. Finally, α_z are zone fixed effects. The results that follow therefore exploit variation in district office management practices *within* geographic zones.

Columns (1)–(3) in Table 8 present results for field visits to schools. We find no evidence of a relationship between the management index and the intensity of school visits in the full sample, although we do see a positive association between the problem-solving sub-index and the intensity of this activity. Panel B shows that this partial effect is statistically significant for directors: a one standard deviation (SD) in the problem-solving sub-index is associated with just under a fifth of a point rise (on the five point scale) in the intensity of school visits. This represents a five percent increase on the director mean intensity for field visits (see Table 5), which is a moderately-sized effect. For the director sub-sample, the partial effect of an accountability-focused management style is negative and significant. A shift from problem-solving focused to an accountability focused approach is associated with a just over a quarter of a point in the intensity of school visits, which represents a seven percent decrease on the director mean intensity for field visits to schools.

Columns (4)–(6) present results for teaching-focused activities (listed in Table 7) during these field visits to schools. The coefficient in the first row of Panel A Column(4) is positive and significant at 1 percent: a one SD increase in the management index is associated with a tenth of a point rise in teaching-focused activities. Panels B to D show that the partial effect is positive for all three job levels, although with varying levels of statistical significance. Interestingly, the effect is largest for deputy directors who visit schools least often on average. The coefficient of 0.126 represents a four percent increase on the deputy director mean intensity for teaching-focused activities (see Table 7 Panel A). There are no discernible patterns in terms of the sub-index, or style of management approach, that matters most.

Results for the "other" activity category are in Columns (7)–(9) and are similar to those for teaching-focused activities. The coefficient in the first row of Panel A Column (7) is

---

[28]See Table A.1 in Appendix A for a full list of controls, together with descriptive statistics. We chose controls "by hand," following the prior literature, but also ran robustness checks using specifications with controls chosen by machine learning. The results presented here (and in all subsequent sections) are robust to this alternative methodology. See Supplementary Appendix B for details.

Table 8: District Office Management Practices and Activity Intensity During School Visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	School visits			Teaching focused			Other activities		
Panel A: Full sample									
Management index	0.044			0.099***			0.089***		
	(0.028)			(0.028)			(0.019)		
Prioritization and targets sub-index		0.050			0.068*			0.036	
		(0.036)			(0.037)			(0.026)	
Monitoring and use of data sub-index		-0.018			0.004			0.023	
		(0.031)			(0.035)			(0.024)	
Accountability and incentives sub-index		-0.033			0.031			0.044*	
		(0.039)			(0.039)			(0.024)	
Problem-solving and adaptation sub-index		0.065**			0.029			0.012	
		(0.032)			(0.034)			(0.024)	
1[Accountability focus]			-0.061			-0.053			0.017
			(0.057)			(0.064)			(0.045)
Observations	785	785	785	761	761	761	761	761	761
Panel B: Directors									
Management index	-0.010			0.059			0.094**		
	(0.063)			(0.065)			(0.040)		
Prioritization and targets sub-index		0.026			0.075			0.043	
		(0.080)			(0.084)			(0.052)	
Monitoring and use of data sub-index		-0.069			0.073			-0.018	
		(0.082)			(0.086)			(0.054)	
Accountability and incentives sub-index		-0.127			-0.063			0.080	
		(0.081)			(0.085)			(0.053)	
Problem-solving and adaptation sub-index		0.175**			-0.022			0.024	
		(0.071)			(0.075)			(0.047)	
1[Accountability focus]			-0.268**			-0.058			0.020
			(0.133)			(0.139)			(0.088)
Observations	156	156	156	156	156	156	156	156	156
Panel C: Deputy directors									
Management index	0.055			0.126***			0.076**		
	(0.051)			(0.048)			(0.030)		
Prioritization and targets sub-index		0.120*			0.105			0.038	
		(0.066)			(0.072)			(0.038)	
Monitoring and use of data sub-index		-0.034			-0.051			0.028	
		(0.056)			(0.063)			(0.035)	
Accountability and incentives sub-index		-0.082			0.136**			0.012	
		(0.073)			(0.066)			(0.041)	
Problem-solving and adaptation sub-index		0.071			-0.015			0.019	
		(0.047)			(0.051)			(0.032)	
1[Accountability focus]			-0.023			0.018			-0.001
			(0.099)			(0.105)			(0.059)
Observations	300	300	300	276	276	276	276	276	276
Panel D: SISOs									
Management index	0.010			0.067*			0.075***		
	(0.040)			(0.039)			(0.024)		
Prioritization and targets sub-index		-0.019			0.044			0.032	
		(0.048)			(0.046)			(0.032)	
Monitoring and use of data sub-index		-0.011			-0.006			0.022	
		(0.047)			(0.040)			(0.029)	
Accountability and incentives sub-index		0.066			0.001			0.051	
		(0.053)			(0.051)			(0.032)	
Problem-solving and adaptation sub-index		-0.023			0.055			-0.010	
		(0.041)			(0.045)			(0.029)	
1[Accountability focus]			0.033			-0.043			0.059
			(0.089)			(0.082)			(0.055)
Observations	329	329	329	329	329	329	329	329	329
District controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Respondent controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zone dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All specifications are at the individual level. The dependent variables are constructed from responses to questions about the frequency of school visits or activities conducted during field visits to schools over the past three months, measured on a five-point scale: 1 never, 2 once or twice, 3 at least monthly, 4 at least weekly, and 5 daily. The management index and sub-indices are standardized, meaning the unit of change is one standard deviation. Columns (1)–(3): all panels use a single question that asks about the frequency of visits to schools. Columns (4)–(6): all panels use questions on teaching-focused activities (see Table 7 Panel A for the list of activities). Columns (7)–(9): all panels use questions on general management activities (see Table 7 Panel B for the list of activities). Panel A: full sample, including directors, deputy directors, and SISOs. Panel B: directors only. Panel C: deputy directors only. Panel C: SISOs only. The respondents who joined the office in the three months immediately prior to the survey are excluded (this includes 18 directors (10 percent of the director sample), 41 deputy directors (12 percent of the deputy director sample), and 19 SISOs (5 percent of the SISO sample). Note that the number of observations for deputy directors drops by 24 in columns (4)–(9) because 24 deputy directors never visit schools. All specifications include district and respondent controls (as listed in Table A.1) as well as zone dummies. The estimates are significant at the *10%, **5%, and ***1% levels.

positive and significant at one percent. Panels B to D show that this partial effect is positive, significant, and similarly sized for all three job levels. The coefficient of 0.094 for directors represents a four percent increase on the director mean intensity for other activities (see Table 7 Panel B). Again, there are no discernible patterns in terms of the sub-index, or style of management approach, that matters most.

6.2   Task completion

In subsection 6.1, we explored how often directors and their staff were undertaking activities within the district office, in schools, and beyond, in the wider community and education system. We now explore whether specific, school-facing delivery tasks (prescribed under the Education Sector Plan) are completed by district offices.

We set out to measure district task completion across the three ESP priorities of access, quality, and management. We asked each director to report on tasks completed over the last full school year (January to December 2021) but also the recipients of these efforts, head teachers and SISOs, to describe their experience of delivery over the same time frame. In this paper, we focus on the latter "experienced" measures.[29]

Our list of tasks was chosen in collaboration with representatives from the Ministry of Education. For access, we asked head teachers about two aspects of delivery: whether the school's furniture request had been met (Task 1); and whether the School Health Education Programme (SHEP) coordinators had received training (Task 2). For each school in a district and each task, we first construct a dummy variable that takes the value 1 if the head teacher tells us that the school had its request/training need met and 0 otherwise. We then compute the average of these dummy variables across the six schools to obtain a district–task level measure for Tasks 1 and 2, and an access index by averaging across these two tasks. For quality, we asked head teachers about three aspects of delivery: whether the school had received (standard) teaching and learning materials (Task 3); in-service training for primary teachers (Task 4); and in-service training for instructional leadership (Task 5). We compute measures using the same method as for access. For management, we asked questions to both SISOs and head teachers. For SISOs, we asked whether they had received training on: school performance improvement plans (SPIPs) (Task 6) and school performance appraisal meetings (SPAMs) (Task 7). For head teachers, we asked whether the school had received training by its SISO relating to SPIPs (Task 8); delivery of training by its SISO relating to SPAMs (Task 9); and whether the school had received a "monitoring and supervision" visit from its SISO (Task 10). We compute measures using the same method as for access, except that for Tasks 6 and 7 we average over the two SISOs. In what follows, we use the access, quality, and management indices separately, as well as an overall task completion index, which is the simple average of the three.

Table 9 shows that the mean rate of task completion (averaging over the 10 tasks and 174 districts) is 73 percent. This figure is high compared to the findings in Rasul et al. (2021), although this is likely due to the more basic nature of the tasks considered in our

---

[29]We plan to examine the relationship between reported and experienced task completion in future work.

survey. The high mean masks considerable variation across tasks. Reassuringly, every school had received at least one "monitoring and supervision" visit from its SISO in the past year. However, on average, just 45 percent schools had their furniture requests met and only 49 percent had received the standard teaching and learning materials (TLMs). While training of, and delivery by, SISOs for SPIPS was high, the completion rate for SPAMS was much lower. Just 52 percent of schools had received training from their SISO on this governance function in the past year.

Table 9: District Task Completion, by Task

Task	Mean	SD
Overall task completion index	0.73	0.30
Access		
Task 1: furniture request met	0.45	0.32
Task 2: SHEP training	0.87	0.18
Quality		
Task 3: delivery of TLMs	0.49	0.25
Task 4: training for primary teachers	0.81	0.19
Task 5: training for instructional leadership	0.80	0.19
Management		
Task 6: training on SPIPs for SISOs	0.83	0.28
Task 7: training on SPAMs for SISOs	0.64	0.35
Task 8: support for SPIPs by SISOs	0.92	0.11
Task 9: support for SPAMs by SISOs	0.52	0.25
Task 10: School visits by SISOs	1.00	0.02
Observations	1740	

Notes: For all tasks except Tasks 6 and 7, we first calculate a district-level proportion by averaging the indicator variable (coded 1 if the school received delivery and 0 otherwise) for six schools and then compute the mean proportion across 174 districts, together with the SD of this district proportion. For Tasks 6 and 7, we do likewise, but we calculate the district-level proportion by averaging the indicator variable for two SISOs instead.

(a) Task completion (overall)

(b) Task completion (access)

(c) Task completion (quality)

(d) Task completion (management)
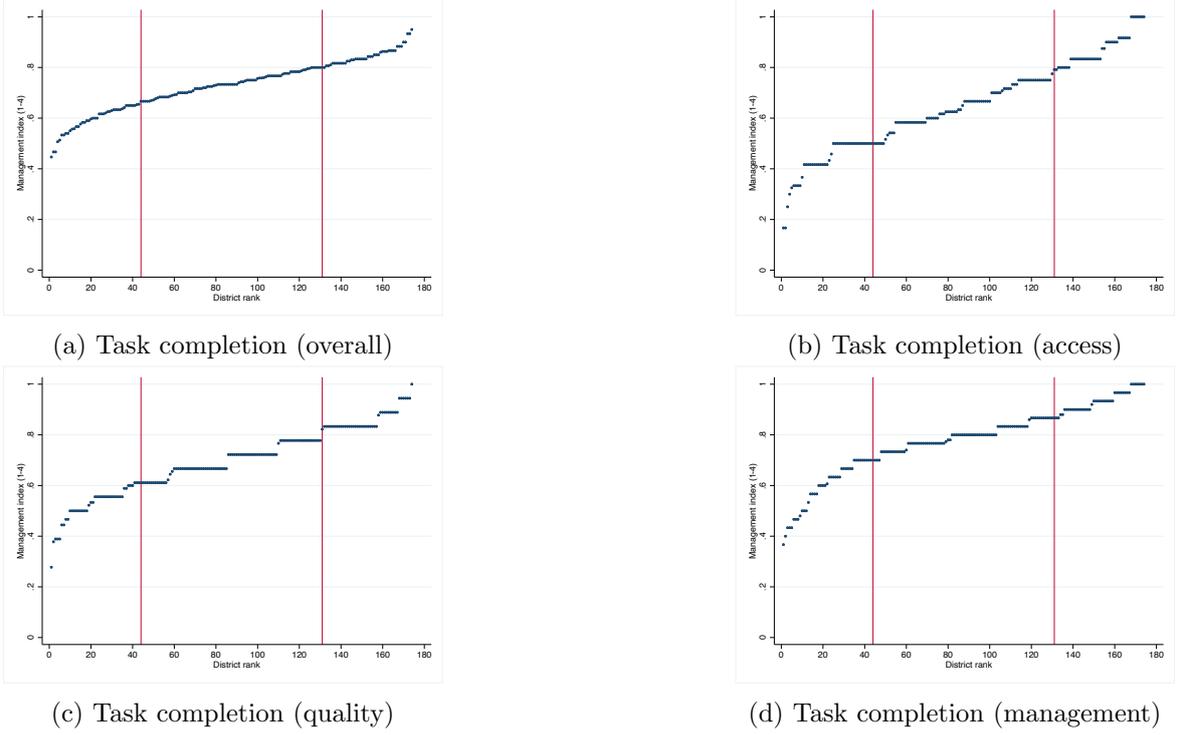
Figure 8: Task Completion (Overall and Sub-Indices), by District Rank

Figure 8 shows that there is also considerable variation in task completion across districts. In each plot, the dots show the proportion of tasks completed for each of the 174 districts (y-axis), ranked from lowest to highest task completion (x-axis). The red lines indicate the districts at the 25th and 75th percentiles. In Panel (a), showing overall task completion, we see moderate dispersion. The district at the 25th percentile has a score of 0.67, while the district at the 75th percentile has a score of 0.80. There is more dispersion for access, however, with an interquartile range of 0.29. Districts below the 25th percentile have an access score of less than 0.50, whereas districts above the 75th percentile have an access score exceeding 0.80. Clearly, far fewer tasks are getting done in some districts than in others.

To study whether district office management practices are associated with district task completion, we estimate the following *task-level* specifications

$$y_{kdz} = \beta_0 + \beta_1 Mgmt_{dz} + X'_{dz}\nu + \alpha_z + \epsilon_{kdz} \tag{4}$$

$$y_{kdz} = \beta_0 + \beta_1 Targets_{dz} + \beta_2 Mon_{dz} + \beta_3 Acc_{dz} + \beta_4 PS_{dz} + \nu X_{dz} + \alpha_z + \epsilon_{kdz} \tag{5}$$

$$y_{kdz} = \beta_0 + \beta_1 Acc\ focus_{dz} + X'_{dz}\nu + \alpha_z + \epsilon_{kdz}, \tag{6}$$

where y_{kdz} is a binary measure of completion for task k in district d in geographic zone z. The management variables are at district level; X_{dz} is a vector of district-level controls (described in Section 6.1), and α_z are zone fixed effects.

28

Table 10: District Office Management Practices and District Task Completion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Overall			Access			Quality			Management		
Management index	0.008			0.009			0.007			0.008		
	(0.009)			(0.014)			(0.012)			(0.013)		
Prioritization and targets sub-index		0.012			-0.037**			0.004			0.037**	
		(0.010)			(0.015)			(0.014)			(0.015)	
Monitoring and use of data sub-index		-0.003			0.012			-0.008			-0.006	
		(0.010)			(0.019)			(0.014)			(0.015)	
Accountability and incentives sub-index		0.002			0.033**			0.008			-0.014	
		(0.011)			(0.017)			(0.015)			(0.016)	
Problem-solving and adaptation sub-index		-0.000			0.003			0.007			-0.006	
		(0.010)			(0.014)			(0.012)			(0.013)	
1[Accountability focus]			0.012			0.007			0.007			0.016
			(0.018)			(0.026)			(0.022)			(0.026)
Observations	1740	1740	1740	348	348	348	522	522	522	870	870	870
District controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zone dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is computed at the district–task level as the proportion of sampled head teachers (or SISOs) who report a task as delivered. Columns (1)–(3) are for all tasks in Table 8 (10 tasks for 174 districts, giving a sample size of 1,740). Columns (4)–(6) are for the two tasks under access, Columns (7)–(9) are for the three tasks under quality, and Columns (10)–(12) are for the five tasks under management. All specifications include district controls (as listed in Table A.1) and zone dummies. The estimates are significant at the *10%, **5%, and ***1% levels.

Table 10 presents our results. We find no evidence that the management index is associated with overall task completion, or completion for access, quality, or management tasks measured separately. The coefficients in the first row are small and imprecisely estimated. There is some evidence of associations for the different sub-indices, although we are cautious about overinterpreting the numerous hypothesis tests. The main point to note here is that in contrast to Rasul et al. (2021), who study task completion across the central bureaucracy in Ghana, we find no evidence of a *negative* relationship between our accountability sub-index and district task completion. As we see below, however, this is not the case for other dimensions of performance.

6.3   Staff understanding of district priorities and roles

We have so far explored how our district office management index—which increases with a more structured use of the management practices that underpin delivery functions—is associated with "getting things done" (intensity of activities and task completion). These metrics are somewhat narrow, however, and unlikely to capture all aspects of performance. In this subsection, we examine a broader metric: district staff *understanding*. Do staff know what their director's priorities are for the district? Do they understand their role within the office and the wider education system? Such understanding may be important in driving district office outputs, and, ultimately, better education outcomes.

We set out to measure whether deputy directors and SISOs understand the priorities of the district as reported by their director. We asked each director to list the top five priorities for his/her office. We then asked the two deputy directors and two SISOs to tell us what they understood to be the top three priorities for the office. We post-code these responses into 10 categories: student learning, teacher pedagogy, student access, district management, school facilities, teacher's human resources, school environment and extracurricular activities, gender issues, STEM initiatives, and community involvement. For each response by a deputy or SISO, we create a binary indicator that is coded 1 if the category is listed among the director's priorities, and 0 otherwise. Taking the average of these dummy variables gives us the proportion of matches for the given staff member, which we use as our measure of understanding of priorities.

For an education system to be effective, it seems likely that staff need to share a common understanding of who is responsible for what. To measure the extent that this is true, we build on work by Adelman and Lemos (2021). These authors select a set of core education roles and then compare de facto understanding of roles with de jure roles (as captured in policy documents) and coherence in how head teachers understand roles, as compared to officials at higher levels in the system. We take a similar approach, but with a focus on staff at the district level. Based on discussions with stakeholders, we first identified seven roles deemed to be of central importance in the Ghanaian context: identifying need for constructing new schools; identifying need for teacher recruitment; managing newly recruited teachers; teacher sanctions; coaching and mentoring for head teachers; in-service teacher training; and monitoring and supervision. Next, we constructed a series of dummy variables

Table 11: District Staff Understanding, by Job Level

	Dep. Director		SISO	
	Mean	SD	Mean	SD
Panel A: District director's priorities				
Overall	0.47	0.33	0.46	0.36
Observations	334		340	
Panel B: Roles in the education system				
Overall	0.33	0.19	0.31	0.19
1. Identifying a need to construct new schools	0.25	0.43	0.26	0.44
2. Identifying teacher recruitment needs	0.18	0.39	0.23	0.42
3. Managing newly recruited teachers	0.49	0.50	0.42	0.49
4. Teacher sanctions	0.24	0.43	0.25	0.44
5. Coaching and mentoring head teachers	0.30	0.46	0.27	0.44
6. In-service teaching training	0.26	0.44	0.31	0.46
7. Monitoring and supervision of schools	0.59	0.49	0.43	0.50
Observations	341		348	

Notes: The first cell in Panel A is constructed as follows: the directors were asked to list their top five priorities for the district, and the deputy directors and SISOs were asked to list three district priorities. For each deputy director, we ascertain the proportion of the listed priorities that agree with their director's list. We then compute the average of these proportions across the two deputy directors to get a district-level variable. The first cell is the average of these variables across 174 districts. The SD is the standard deviation in this district mean. The SISO cell is computed in the same way. The means in Panel B are constructed as follows: directors and staff were asked to pick the level/organization of the education system with responsibility for seven roles. For each deputy director and each role, we create an indicator coded 1 if the respondent's understanding of the roles is the same as the director's (and 0 otherwise). We then compute the average across the two deputy directors to get a district-level variable. The first cell in each row is the average across 174 districts, and SD is the standard deviation in these district means. The SISO cells are computed in the same way. The means in Panel A are unweighted. The means in Panel B are weighted using the design weights.

measuring the coherence in understanding of who (which level/organization) is responsible for each of these seven roles. We did this in pairs, focusing on the director and matching to each deputy and SISO for each of the seven roles. Each dummy variable takes the value 1 if, for that role, the staff member has the same understanding as the director and 0 otherwise. Taking the average of these dummy variables gives us the proportion of matches for the given staff member, which we use as our measure of understanding of roles.

Table 11 Panel A shows that district staff have only a moderate understanding of district priorities. On average, just under half of the priorities listed by deputy directors and SISOs match those of their director. In Panel B, we see that there is heterogeneity in understanding of roles in the education system. On average, close to 60 percent of deputy directors have the same understanding as their director as to who is responsible for monitoring and supervision of schools. However, this proportion drops to less than 20 percent of deputies when it comes to identifying teacher recruitment needs. Taking the mean across all seven roles, only a third of deputy directors have the same understanding of education system roles as their director. Among SISOs, coherence in understanding with the director follows a broadly similar pattern, except for monitoring and supervision of schools, which is much lower for SISOs than for deputy directors. This is surprising, since it would typically be SISOs who

have that responsibility.[30]

To study whether district office management practices are associated with understanding of priorities and/or roles, we estimate the same (respondent-level) specifications as in equations (1)–(3).[31] Table 12 presents results. In the specifications for the understanding of priorities, the partial effect of the prioritization function is always positive and is significant in Panel A for the full sample and in Panel B for deputy directors. This is intuitive and indeed reassuring from a measurement perspective. Based on the coefficient for the full sample, a one SD increase in prioritization is associated with a five percentage point rise in understanding of priorities. This is a sizeable effect, given that mean "understanding" is less than 50 percent (see Table 11 Panel A). The partial effect of the problem-solving function is also always positive, statistically significant in each panel, and of a similar, sizeable economic magnitude. We find no evidence of a substantive association between the management index/sub-indices and understanding of roles; coefficients are small and, typically, not statistically significant.

6.4 Staff attitudes and satisfaction

In this final subsection, we explore staff attitudes and satisfaction. In our survey, we administered a suite of question banks intended to capture attitudes, all based on internationally validated scales. These were: organizational culture in the district office (Azulai et al., 2020; Sexton et al., 2006), promotive and prohibitive dimensions of voice (Liang et al., 2012), and leadership support for staff autonomy (Deci et al., 2017), all measured on a standard five-point Likert scale. We measured satisfaction in two ways. First, we used the Maslach Burnout Inventory, following Bloom et al. (2014), who use the validated scale from Maslach and Jackson (1981). Exhaustion was measured using a zero (every day) to six (never) scale. In the original scale, all statements are negatively worded, hence a lower value indicates more exhaustion. We found it convenient to reverse the scale, meaning that higher exhaustion is indicated by a higher value. We also measured intention to leave. Here, we asked respondents a single question: how likely it is that you will leave your job in the coming year.[32] The answers were given on a five-point scale (1 very likely to 5 very unlikely). We present descriptive statistics using the five-point scale, but follow the literature and use a dummy variable, coded 1 for likely/very likely and 0 otherwise, in our regressions.

---

[30]This echoes Adelman and Lemos (2021), who find that respondents often claim responsibility for roles other than their own.

[31]The number of observations differ because we do not include directors in the full sample and because for activity intensity we drop respondents who had not been in post during the previous three months.

[32]This question is commonly used in teacher surveys; see, for example, Leaver et al. (2021) and the references therein.

Table 12: District Office Management Practices and Understanding of Priorities and Roles

	(1)	(2)	(3)	(4)	(5)	(6)
	Understanding of priorities			Understanding of roles		
Panel A: Full sample						
Management index	0.026			0.002		
	(0.017)			(0.009)		
Prioritization and targets sub-index		0.048**			0.012	
		(0.024)			(0.012)	
Monitoring and use of data sub-index		-0.033			-0.001	
		(0.023)			(0.012)	
Accountability and incentives sub-index		-0.020			-0.013	
		(0.023)			(0.013)	
Problem-solving and adaptation sub-index		0.046**			0.004	
		(0.022)			(0.011)	
1[Accountability focus]			-0.023			-0.010
			(0.039)			(0.020)
Observations	674	674	674	689	689	689
Panel B: Deputy directors						
Management index	0.022			0.015		
	(0.020)			(0.014)		
Prioritization and targets sub-index		0.061**			0.013	
		(0.029)			(0.016)	
Monitoring and use of data sub-index		-0.026			0.005	
		(0.026)			(0.017)	
Accountability and incentives sub-index		-0.048*			-0.025	
		(0.027)			(0.017)	
Problem-solving and adaptation sub-index		0.047**			0.028*	
		(0.023)			(0.015)	
1[Accountability focus]			-0.050			-0.040
			(0.045)			(0.028)
Observations	334	334	334	341	341	341
Panel C: SISOs						
Management index	0.028			-0.012		
	(0.025)			(0.011)		
Prioritization and targets sub-index		0.034			0.011	
		(0.032)			(0.018)	
Monitoring and use of data sub-index		-0.041			-0.005	
		(0.032)			(0.016)	
Accountability and incentives sub-index		0.002			-0.004	
		(0.030)			(0.017)	
Problem-solving and adaptation sub-index		0.051*			-0.018	
		(0.029)			(0.014)	
1[Accountability focus]			-0.002			0.019
			(0.049)			(0.026)
Observations	340	340	340	348	348	348
District controls	Yes	Yes	Yes	Yes	Yes	Yes
Respondent controls	Yes	Yes	Yes	Yes	Yes	Yes
Zone dummies	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All specifications are at the individual level. In Columns (1)–(3), the dependent variable is a continuous variable between 0 and 1 showing the proportion of matched priorities with the director (see notes to Table 11). The management index and sub-indices are standardized, meaning the unit of change is one standard deviation. In Columns (4)–(6), the dependent variable is a continuous variable between 0 and 1, showing the proportion of the seven roles with coherence in understanding with the director (see notes to Table 11). Panel A: full sample, including deputy directors and SISOs. Panel B: individual-level, just deputy directors. Columns (1)–(3) have 14 missing values for deputy directors. Panel C: individual-level, just SISOs. Columns (1)–(3) have eight missing values for SISOs. All specifications include district and respondent controls (as listed in Table A.1) as well as zone dummies. The estimates are significant at the *10%, **5%, and ***1% levels.

(a) Five-point scales   (b) Exhaustion, seven-point scale
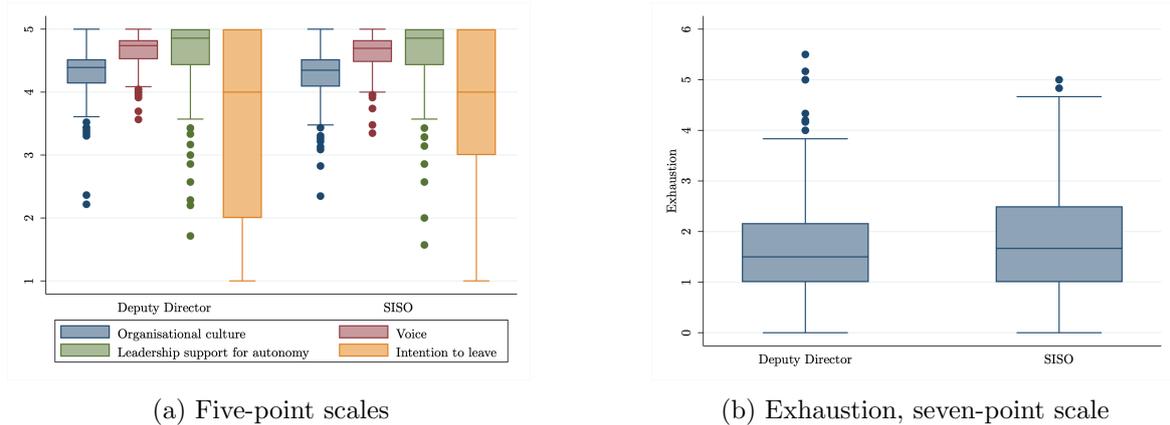
Figure 9: Attitudes and Satisfaction of District Staff, by Job Level

Figure 9 Panel (a) shows that there is limited variation in the majority of these scales; the vast majority of respondents agreed (or strongly agreed) with *all* of the statements in the organizational culture, voice, and leadership support question banks. We suspect that this reflects social desirability bias, and refrain from interpreting these results. We do, however, see meaningful variation in the questions relating to job satisfaction. For both deputy directors and SISOs, the median respondent appears to be content with his/her job and is unlikely (or very unlikely) to leave in the next year. A minority of respondents are less satisfied. Among deputy directors, we see *negative* responses (likely or very likely to leave) at the 25th percentile and below. Perhaps reflecting this broad satisfaction, we also find limited evidence of exhaustion among either deputy directors or SISOs. Figure 9 Panel (b) shows that the median score (and indeed the entire interquartile range) on our exhaustion index is well below three, indicating that it is uncommon for most respondents to experience feelings of exhaustion and burnout.

Given the lack of variation in most of the scales, we focus on just two dependent variables in our regression analysis, again using the (respondent-level) specifications in equations (1)–(3). Table 13 Columns (1)–(3) present results for our (standardized) exhaustion index. We find no evidence of an association between the management index/sub-indices and district staff exhaustion; coefficients are small and not statistically significant. Columns (4)–(6) present results for an indicator variable of intention to leave. We find no evidence of an association between the management index and this measure of job satisfaction. However, we do find evidence that the problem-solving sub-index is associated with a statistically significant lower intention to leave. The economic magnitude is meaningful; a one SD increase in the problem-solving sub-index is associated with a reduction of intention to leave of just under five percentage points. Since, on average, 22 percent of district staff indicated that they are likely or very likely to leave in the next year, this is a sizeable (partial) effect. The coefficient on the accountability focus indicator is negative and also large in absolute magnitude. A shift from a problem-solving to accountability-focused approach is associated with close to a nine percentage point increase in intention to leave in the next year.

34

Table 13: District Office Management Practices and District Staff Satisfaction

| | (1) | (2) | (3) | (4) | (5) | (6) |
	Exhaustion			Intention to leave		
Panel A: Full sample						
Management index	-0.007			0.010		
	(0.045)			(0.022)		
Prioritization and targets sub-index		-0.028			0.001	
		(0.072)			(0.028)	
Monitoring and use of data sub-index		0.054			0.042*	
		(0.072)			(0.022)	
Accountability and incentives sub-index		0.002			0.008	
		(0.061)			(0.027)	
Problem-solving and adaptation sub-index		-0.047			-0.047**	
		(0.061)			(0.020)	
1[Accountability focus]			0.083			0.086**
			(0.103)			(0.039)
Observations	689	689	689	689	689	689
Panel B: Deputy directors						
Management index	-0.017			0.018		
	(0.066)			(0.031)		
Prioritization and targets sub-index		-0.063			0.008	
		(0.091)			(0.031)	
Monitoring and use of data sub-index		0.092			0.045	
		(0.094)			(0.030)	
Accountability and incentives sub-index		-0.035			0.000	
		(0.087)			(0.035)	
Problem-solving and adaptation sub-index		-0.028			-0.037	
		(0.081)			(0.028)	
1[Accountability focus]			0.110			0.106**
			(0.139)			(0.051)
Observations	341	341	341	341	341	341
Panel C: SISOs						
Management index	0.000			-0.001		
	(0.064)			(0.023)		
Prioritization and targets sub-index		0.006			-0.002	
		(0.092)			(0.036)	
Monitoring and use of data sub-index		0.018			0.035	
		(0.101)			(0.029)	
Accountability and incentives sub-index		0.032			0.013	
		(0.077)			(0.035)	
Problem-solving and adaptation sub-index		-0.063			-0.056**	
		(0.069)			(0.024)	
1[Accountability focus]			0.051			0.066
			(0.129)			(0.053)
Observations	348	348	348	348	348	348
District controls	Yes	Yes	Yes	Yes	Yes	Yes
Respondent controls	Yes	Yes	Yes	Yes	Yes	Yes
Zone dummies	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All specifications are at the individual level. In Columns (1)–(3), the dependent variable is a standardized index based on answers to six questions about how frequently the respondents experience feelings of exhaustion/burnout measured on a seven-point scale (0 never to 6 every day). In Columns (4)–(6), the dependent variable is a dummy variable coded 1 if respondents state that they are 1 very likely or 2 likely to leave the office over the next year, and coded 0 otherwise. The management index and sub-indices are standardized, meaning the unit of change is one standard deviation. Panel A: full sample, including deputy directors and SISOs. Panel B: individual-level, just deputy directors. Panel C: individual-level, just SISOs. All specifications include district and respondent controls (as listed in Table A.1) as well as zone dummies. The estimates are significant at the *10%, **5%, and ***1% levels.

7 Performance at school level

In this section, we explore performance at school level and the relationship between this performance and district office management practices. As in Section 6, we organize the analysis by different dimensions of performance. Within each subsection, we summarize measurement, present descriptive statistics (this time giving a sense of cross-school variation), and then report regression results. Again, we postpone discussion of these results to Section 8. We use the GEOP sample since this gives us much richer data. As Figure 3 illustrates, these districts are mostly located in the north of Ghana. Although the findings in this section cannot be viewed as nationally representative, they are policy-relevant since the north of Ghana is the geographic area where schools have, historically, faced the most challenges.

7.1 Head teacher activity intensity

We measure the intensity of head teachers' activities in school using broadly the same methodology as in Section 6.1. We focus on identical teaching-focused activities but exclude the three governance-related activities (collecting data, observing log books, and spot checks) and add an activity on meeting with district staff. Table 14 shows that the most frequent activities undertaken by head teachers are teaching-focused and take place *in the classroom*: observing lessons and reviewing student work. It is somewhat surprising that on average, even these activities occur on a less than weekly basis. In contrast to SISOs, no mean intensity is greater than four, indicating that the frequency of activity is between weekly and daily.

To explore whether district office management practices are associated with the intensity with which head teachers undertake these activities, we estimate the following *school-level* specifications

$$y_{sd} = \alpha + \beta_1 Mgmt_d + \nu X_d + \gamma Z_s + \epsilon_{sd} \tag{7}$$

$$y_{sd} = \alpha + \beta_1 Targets_d + \beta_2 Mon_d + \beta_3 Acc_d + \beta_4 PS_d + \nu X_d + \gamma Z_s + \epsilon_{sd} \tag{8}$$

$$y_{sd} = \alpha + \beta_1 Acc\ focus_d + \nu X_d + \gamma Z_s + \epsilon_{sd}, \tag{9}$$

where y_{sd} is a measure of activity intensity for the head teacher in school s in district d. We use two different dependent variables: intensity of teaching-focused activities (an average of six questions) and other activities (an average of nine questions). The management variables are at the district level: the vector X_d contains district-level controls, and the vector Z_s contains school-level controls (head teacher gender and age, rural-urban dummies, a school facilities index, the school's GALOP score, and the number of permanent teachers). The full list of controls is provided in Table A.2 in Appendix A. Since the entire GEOP sample is drawn from the northern and middle zones, we include a single indicator that is coded 1 if

Table 14: Activity Intensity by Head Teachers (GEOP Sample)

	Mean	SD
Panel A: Teaching-focused activities		
Classroom observation	3.70	0.96
Looking at sample of student work	3.67	0.87
Reviewing lesson plans	3.55	0.85
Coaching or mentoring session	2.54	1.12
Providing in-service training	2.05	0.70
Panel B: Other activities		
Discussing teacher career development	2.25	0.96
Supporting teachers in filling out their forms	1.76	0.79
Discussing school progress with PTA/SMC	2.15	0.43
Support, supervision, or implementation of SPIP	1.96	0.58
Monitoring and audit of the school capitation grant	1.80	0.60
Related to COVID-19	3.25	1.51
Managing teacher disputes	1.56	0.90
Meeting with SISO/DTST	2.80	0.74
Observations	1048	

Notes: In the first cell, we take the head teacher's response to a single question about the frequency with which he/she undertook classroom observation over the past three months (measured on a five-point scale: 1 never, 2 once or twice, 3 at least monthly, 4 at least weekly, and 5 daily) and then compute the average over 1,048 schools. All rows are based on responses to a single question, except for the rows on reviewing lesson plans and discussing school progress with the PTA/SMC, which are based on two questions each. The respondents who joined the school in the three months immediately prior to the survey are excluded (213 head teachers, 16 percent of the EOF sub-sample). The values are weighted using design weights.

the school is located in the northern zone, and 0 otherwise.[33]

Table 15 presents results. We find no evidence of a relationship between the management index and the intensity of activities undertaken by head teachers; the coefficients in the first row are small and imprecisely estimated. When we look at the sub-indices, we see some competing partial effects. The problem-solving sub-index has a positive, but only weakly significant, partial effect: for both activity categories, a one SD increase in this sub-index is associated with just under a twentieth of a point rise (on a five point scale) in activities by the head teacher. This is a very small effect. To illustrate, the coefficient in Column (2) represents only a 1.5 percent increase on the head teacher mean intensity for teaching-focused activities (see Table 14 Panel A). In contrast, the prioritization index has a negative, but similarly small, partial effect.

7.2 Head teacher attitudes and satisfaction

We administered the full bank of attitudinal questions to head teachers, albeit with edits to reflect the school rather than district office context. As was the case for district staff, there was limited variation in responses, except for the questions relating to job satisfaction (see Figure 10). The median head teacher is unlikely or very unlikely to leave his/her job in the next year, although there are some negative responses below the 25th percentile. Exhaustion

---

[33]See Figure 3 Panel (b). Seven districts are located in the middle zone and 20 are in the northern zone.

Table 15: District Office Management Practices, Activity Intensity by HTs (GEOP Sample)

	(1)	(2)	(3)	(4)	(5)	(6)
	Teaching focused			Other activities		
Management index	0.023			0.010		
	(0.020)			(0.023)		
Prioritization and targets sub-index		-0.093**			-0.074**	
		(0.035)			(0.027)	
Monitoring and use of data sub-index		0.036			0.023	
		(0.022)			(0.026)	
Accountability and incentives sub-index		0.032			0.008	
		(0.025)			(0.017)	
Problem-solving and adaptation sub-index		0.046*			0.047*	
		(0.024)			(0.025)	
1[Accountability focus]			0.021			-0.008
			(0.060)			(0.044)
Observations	1047	1047	1047	1047	1047	1047
District controls	Yes	Yes	Yes	Yes	Yes	Yes
School controls	Yes	Yes	Yes	Yes	Yes	Yes
Location dummy	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All specifications are at the individual level. The respondents who joined the office in the three months immediately prior to the survey are excluded. The dependent variables are constructed from responses to questions about activities conducted over the past three months, measured on a five-point scale: 1 never, 2 once or twice, 3 at least monthly, 4 at least weekly, and 5 daily. In Columns (1)–(3) these questions are about teaching-focused activities (see Table 14 Panel A for the list of activities). In Columns (4)–(6) these questions are about general management activities (see Table 14 Panel B for the list of activities). All specifications include district and school controls (as listed in Table A.2). The estimates are significant at the *10%, **5%, and ***1% levels. The errors are clustered at the district level.

is also low, on average, although there are a small number of negative responses (coded here as high scores), indicating a high degree of burnout.
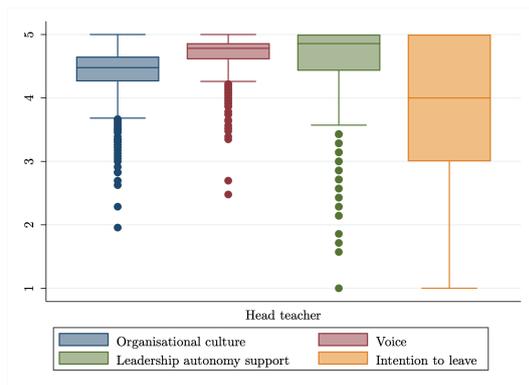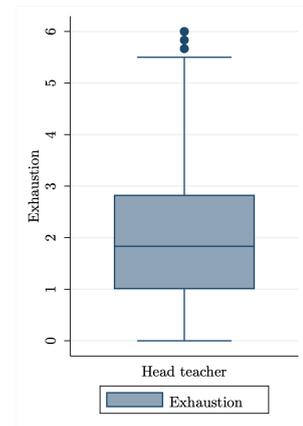
(a) Five-point scales   (b) Seven-point scale

Figure 10: Attitudes of Head Teachers (GEOP Sample)

To explore whether district office management practices are associated with head teacher satisfaction, we repeat the regression analysis from Section 7.1 using the (standardized) exhaustion index and the binary "intention to leave" indicator as dependent variables. Table 16 reports results. We find no evidence of associations between our management index and these two measures of job satisfaction for head teachers. We do, however, see evidence of

competing partial effects for the different sub-indices. Echoing our results for district staff, the monitoring sub-index is positively associated with intention to leave. A one SD increase in the monitoring sub-index is associated with a 4 percentage point *increase* in head teacher intention to leave. Since, on average, 23 percent of head teachers indicated that they are likely or very likely to leave in the next year, this is a meaningfully sized partial effect. The coefficient for the accountability sub-index is also positive (albeit smaller and only weakly significant).

Table 16: District Office Management Practices, Head Teacher Satisfaction (GEOP Sample)

| | (1) | (2) | (3) | (4) | (5) | (6) |
		Exhaustion			Intention to leave	
Management index	-0.046			0.013		
	(0.037)			(0.010)		
Prioritization and targets sub-index		0.011			-0.038***	
		(0.076)			(0.013)	
Monitoring and use of data sub-index		-0.019			0.039***	
		(0.055)			(0.010)	
Accountability and incentives sub-index		-0.058			0.022*	
		(0.053)			(0.012)	
Problem-solving and adaptation sub-index		-0.003			-0.009	
		(0.063)			(0.011)	
1[Accountability focus]			-0.087			0.010
			(0.127)			(0.034)
Observations	1260	1260	1260	1211	1211	1211
District controls	Yes	Yes	Yes	Yes	Yes	Yes
School controls	Yes	Yes	Yes	Yes	Yes	Yes
Location dummy	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All specifications are at the individual level. In Columns (1)–(3), the dependent variable is a standardized index that captures head teacher exhaustion. In Columns (4)–(6), the dependent variable is a dummy variable coded 1 if the head teacher reports that his/her intention to leave the school in the next year is 1 very likely or 2 likely, and coded 0 otherwise. All specifications include district and school controls (as listed in Table A.2). The estimates are significant at the *10%, **5%, and ***1% levels. The errors are clustered at the district level.

7.3 Teacher absenteeism and classroom practice

We measure teacher absenteeism using the same methodology as in Bold et al. (2017), which is an extended version of the strategy used in Chaudhury et al. (2006). In our survey, an enumerator first made an announced visit to a school, during which time he/she constructed a teacher roster, collected information on timetables, and randomly selected four teachers for a spot-check. The enumerator then conducted an *unannounced* visit two to four days after the initial visit in which he/she checked whether the identified teachers were in school. If the teachers were in school, the enumerator checked whether each teacher was in the classroom teaching during scheduled class time. Our measure of teacher absence is the share of teachers (out of the four selected teachers) who were not in the school during the unannounced visit.

We measure classroom practice using the Stallings Classroom Observation tool. Under this rubric, data is collected by an external observer who takes 10 visual snapshots during a class period, and fills out a questionnaire based on their observations. Our enumerators observed two class periods in each school: a fourth-grade math class, and another fourth-grade class of any subject. We construct three variables at the school level: "active instructional time," "students off-task," and "teacher off-task." Active instructional time is the propor-

Table 17: Teacher Absenteeism and Time on Task (GEOP Sample)

	Mean	SD	P25	P50	P75
Teacher absence from school	0.13	0.21	0.00	0.00	0.25
Active instructional time	0.74	0.18	0.60	0.75	0.90
Students off-task	0.21	0.23	0.00	0.15	0.35
Teacher off-task	0.03	0.06	0.00	0.00	0.05
Observations	1259	1259	1259	1259	1259

Notes: Teacher absence from school is measured using the SDI methodology. To compute the proportion in the first row, we first calculate a school-level proportion of the four sampled teachers who are absent, and then we average this proportion over the 1,259 schools (we have missing data for two schools). SD is the standard deviation in this school-level proportion. The "time on task" variables are measured using the Stallings methodology and are based on two observations per school: a fourth-grade math class and another fourth-grade class of any subject on the same day. To compute the proportion in the second row, we first calculate a school-level proportion of the 20 snapshots (10 per class) categorized as active instruction. We then average this proportion over the 1,259 schools. SD is again the standard deviation in this school-level proportion. The third and fourth rows are computed in the same way.

tion of observed lesson time (based on the 20 snapshots, 10 per class) in which the teacher is undertaking one of five activities: reading aloud, demonstration/lecture, Q&A/discussion, practice/drilling, or assignment/classwork.[34] "Students off-task" refers to a situation where a large group of students are distracted, being disciplined, socially interacting, or not engaged in the lesson. "Teacher off-task" includes social interaction between the teacher and students, social interaction between the teacher and another adult, or the teacher not being present in the classroom. Both variables are constructed as a proportion of observed lesson time.

Table 17 presents summary statistics. Across 1,259 schools in the 27 GEOP districts, the average rate of teacher absence from school is 13 percent. This figure is low compared to other estimates for sub-Saharan Africa (cf. Chaudhury et al. (2006), and more recent Service Delivery Indicator (SDI) estimates in Bold et al. (2017)). Figure 11 Panel (a) gives a sense of the variation across schools. In the vast majority of schools, none of the four randomly selected teachers were absent from school on the day of the unannounced visit, and at the 75th percentile, only one of the four sampled teachers was absent. Multiple absences (two or more of the four randomly selected teachers) occurred in just a small number of schools.

---

[34]There is only one instructional activity not included in the *active* instruction category.

(a) Teacher absence from school

(b) Active instructional time

(c) Students off-task

(d) Teacher off-task
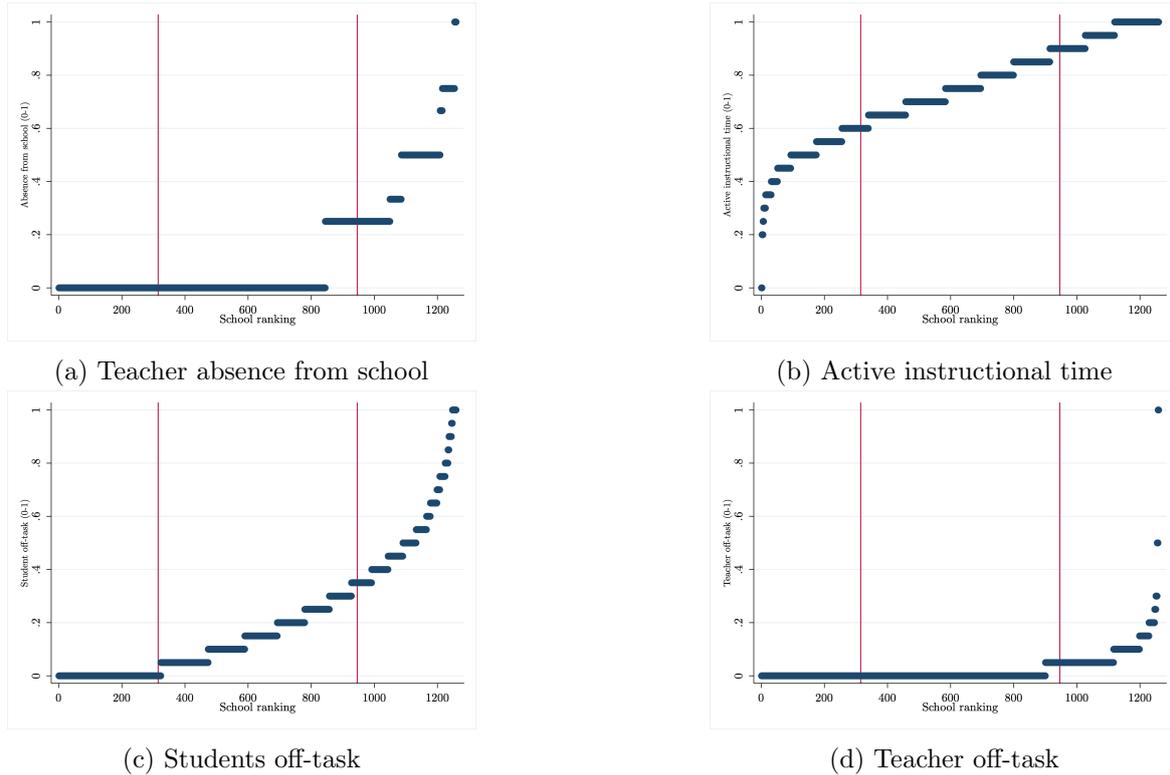
Figure 11: Teacher Absenteeism and Time on Task, by School Rank (GEOP Sample)

Turning to classroom practice, on average, across the two classes in the 1,259 schools for which we have data, the proportion of observed lesson time considered to be "active instruction" was 0.74. This figure is higher than the 60 percent reported in Stallings et al. (2014) for Ghana. On average, the proportion of observed lesson time where students were "off task" (whether being disciplined, interacting socially, or simply not engaged) was 0.21. Finally, on average, the proportion of observed lesson time where the teacher was "off task" was 0.03. The latter two findings are comparable to earlier estimates for Ghana (Stallings et al., 2014), although low compared to global best practice.[35] Although these averages are moderately encouraging, Panels (b) and (c) in Figure 11 show that there is substantial variation across schools. In some schools, active instructional time is well below 50 percent, and students are off task more than 50 percent of the time. Reassuringly, there appear to be very few schools where the teacher is off task for the majority of the time.

To explore whether district office management practices are associated with teacher absenteeism and classroom practice, we use the specifications from equations (7)–(9), this time with the following dependent variables: the proportion of four randomly sampled teachers who were absent from school, and the proportion of observed lesson time coded as "active instruction," "students off task," and "teacher off task."

Table 18 presents results for teacher absenteeism. We find no evidence of an association between the management index and teacher absenteeism, but we do see evidence of competing partial effects for the different sub-indices. The partial effect of the problem-solving index is

---

[35]Stallings good practice benchmarks for effective time use are 50 percent or more for active instruction, 0 percent for teacher off task, and 6 percent or less for students off task, per Bruns and Luque (2015).

Table 18: District Office Management Practices and Teacher Absence (GEOP Sample)

	(1)	(2)	(3)
Management index	-0.002		
	(0.010)		
Prioritization and targets sub-index		-0.025	
		(0.017)	
Monitoring and use of data sub-index		0.029***	
		(0.010)	
Accountability and incentives sub-index		0.027**	
		(0.011)	
Problem-solving and adaptation sub-index		-0.033***	
		(0.011)	
1[Accountability focus]			0.072**
			(0.029)
Observations	1259	1259	1259
District controls	Yes	Yes	Yes
School controls	Yes	Yes	Yes
Location dummy	Yes	Yes	Yes

Notes: The dependent variable is the school-level proportion of four randomly sampled teachers who were absent from school at the start of the day of an unannounced visit. All specifications include district and school controls (as listed in Table A.2). The estimates are significant at the *10%, **5%, and ***1% levels. The errors are clustered at the district level.

negative and strongly significant: a one SD increase in the problem solving index is associated with a 3.3 percentage point reduction in teacher absenteeism (from school). By contrast, the partial effect of the accountability index is positive: a one SD increase in the accountability index is associated with a 2.7 percentage point increase in teacher absenteeism. Consistent with this, the coefficient on the accountability focus indicator in Column (3) is positive: a shift from a problem-solving focused to an accountability-focused approach is associated with a 7.2 percentage point increase in teacher absenteeism. This is a substantial effect, given the mean absence rate reported in Table 17 of 13 percent.

Table 19: District Office Management Practices and Time on Task (GEOP Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Active instructional time			Students off-task time			Teacher off-task time		
Management index	0.027**			-0.050***			0.001		
	(0.010)			(0.007)			(0.001)		
Prioritization and targets sub-index		0.017**			-0.028**			-0.003**	
		(0.008)			(0.011)			(0.001)	
Monitoring and use of data sub-index		0.040***			-0.017*			0.001	
		(0.007)			(0.009)			(0.001)	
Accountability and incentives sub-index		-0.030***			-0.005			0.006***	
		(0.010)			(0.007)			(0.001)	
Problem-solving and adaptation sub-index		0.009			-0.022**			-0.003**	
		(0.006)			(0.008)			(0.001)	
1[Accountability focus]			-0.038			0.007			0.013***
			(0.033)			(0.034)			(0.003)
Observations	1259	1259	1259	1259	1259	1259	1259	1259	1259
District controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is school-level active instruction in the classroom in Columns (1)–(3), school-level students off-task in Columns (4)–(6), and school-level teacher off-task in Columns (7)–(9). All specifications include district and school controls (as listed in Table A.2). The estimates are significant at the *10%, **5%, and ***1% levels. The errors are clustered at the district level.

Table 19 presents results for classroom practice. For active instructional time, the partial effect of the management index is positive and significant: a one standard deviation increase in the management index is associated with 2.7 percentage point increase in active instruction. This is a small effect, given the mean proportion of observed lesson time coded as active instruction was 74 percent. Column (2) shows that this is driven by a mix of the functions, but notably *not* the accountability sub-index which has a negative partial effect. For students off task, the partial effect of the management index is negative and strongly significant: a one standard deviation increase in the management index is associated with a five percentage point decrease in students being off task. The mean proportion of observed lesson time coded as students off task was 21 percent, making this a much more meaningfully sized effect. Column (5) shows that this is driven by a mix of functions, although again not the accountability sub-index. The coefficients for "teacher off task" are small, and we refrain from interpretation except to note that qualitatively, the findings echo the previous results.

8 Discussion and concluding remarks

We now collect together the results reported in Sections 6 and 7 to provide a series of takeaways. We begin with descriptions of district office management practices and performance, and we then turn to the relationships between the two.

On average, across the 174 districts in the nationally representative sample, we found *only moderate use of the 16 practices in our management index.* The scores indicate that practices exist, but often in *ad hoc* or informal ways. The average district is strong on priority setting, but makes more limited use of strategic planning, and while data are collected, it appears that they are not always reviewed/used by district management. In terms of the style of approach, the average district office makes greater use of problem-solving practices than (top-down) accountability practices.

Behind these averages, there is *substantial variation in how district offices are run.* District offices are distributed across (almost) the full range of scores for all four management sub-indices. There is also a striking geographic split in the style of approach, with districts in southwestern regions focusing more on accountability than problem-solving and the reverse holding in the rest of Ghana. This cross-district variation in the use of management practices is more closely associated with "external factors" (support from non-governmental and international organizations, and engagement from politicians) than with characteristics of directors or office resources.

Performance at district level is mixed. On the plus side, district staff are, on average, visiting schools frequently. While most emphasis is on narrow, compliance activities during these visits, time is also spent in the classroom. For certain tasks, particularly around training, average completion rates are reasonably high. On the negative side, tasks that require delivery of physical goods (e.g., furniture, textbooks) are completed much less often, perhaps reflecting the budget delays and shortfalls documented here and elsewhere (Bell et al., 2023). For both activity intensity and task completion, there is a tail of poor performers;

some directors undertake field visits to schools only once a quarter, and the overall rate of task completion in the lowest-ranked district is just above 40 percent. Further concerns emerge from our exploration of staff understanding and job satisfaction. In the average district office, staff show relatively little understanding of what the district is trying to achieve and of their/others' roles within the education system. Just over a fifth of the 863 district staff that we surveyed said that they are likely/very likely to leave in the next year.

Aspects of performance at school level are concerning (in the GEOP sample). On average, head teachers are spending relatively little time in the classroom, and even less time on teacher support, parent/community engagement, and governance. Just over a fifth of the 1,261 head teachers surveyed said that they are likely/very likely to leave in the next year. On average, teacher absenteeism from school is moderate by regional standards (although high by global ones), but this belies considerable variation across schools. There is a small minority of schools where teacher absenteeism and poor classroom practice are serious issues.

We now turn to the relationship between district office management practices and performance. In Sections 6 and 7, we studied eight dimensions of performance using three different econometric specifications and, for some dimensions, different sub-samples. As a result, we ran many hypothesis tests. Since the analysis was purposefully exploratory, we did not correct $p-$values when reporting results in those sections. However, when emphasizing takeaways, we wish to be more conservative and therefore correct for multiple hypothesis tests (within dimensions of performance).[36]

There is *some evidence of positive associations between the management index and performance.* We find moderately-sized positive effects for the intensity of activities during field visits to school by district staff (Table 8), as well as positive effects (of varying magnitudes) for classroom practice (Table 19). Both of these findings are robust to corrections for multiple hypothesis tests. Given that we studied a total of eight dimensions of performance, this is fairly meager evidence. However, when we examine the management sub-indices individually, as suggested by the conceptual framework in Section 3, a richer picture emerges.

First, *there is no robust evidence of positive associations between district office use of accountability practices and performance, but robust evidence of negative associations between accountability and performance (at school level and in the GEOP sample).* The weak evidence that accountability may be beneficial for process-oriented measures at district level (specifically district task completion, Table 10) does not survive multiple hypothesis test correction. In contrast, some of the evidence against accountability does survive this correction, specifically for teacher absenteeism (Table 18) and classroom practice (Table 19).

Second, *there is robust evidence of positive associations between district office use of problem-solving practices and performance.* In Sections 6 and 7, we reported many positive associations between problem-solving practices and better performance. Not all of these relationships survive multiple hypothesis test correction, but three do. These are for: district staff job satisfaction (negative coefficient for intention to leave, Table 13), teacher presence at school (negative coefficient for absenteeism, Table 18), and classroom practice (Table 19).

---

[36]Details of how we perform these corrections, and the associated $q-$values, are provided in Appendix C.

The picture that emerges is that district staff working in offices where management practices that foster problem-solving are more prevalent (holding other management practices, a host of other district observables, and geographic zone constant) are less likely to want to leave in the next year. At the frontline, teachers working in schools under the jurisdiction of these district offices are (*ceteris paribus*) present at school more often, and teach classes in which they and their students are less often "off-task."

Although we have not emphasized prioritization and monitoring practices to the same extent, for completeness it is worth summarizing the MHT-corrected results for these functions. Prioritization is positively associated with head teacher job satisfaction and good classroom practice. Monitoring is positively associated with good classroom practice (particularly active instruction time) but negatively associated with head teacher job satisfaction, as well as teacher presence at school.[37]

There are, of course, limitations and caveats that should be noted alongside these takeaways. Most importantly, we are highlighting relationships but *not* making causal claims. We have done our utmost to control for observable characteristics of districts and schools, as well as factors such as geography, that could be correlated with both management and performance. But omitted variables may remain. We have highlighted what we feel are the consistent, robust relationships. The next step is to investigate whether the practices that seem promising, such as problem-solving, show impacts in a causal research design. It is also worth emphasizing that although our results for district performance are nationally representative, our school-level analysis is for the GEOP sub-sample concentrated in the north of Ghana. Future research is also needed to explore whether these findings extend to the rest of Ghana.

We conclude by noting considerations for policy. To be clear, our paper is not presenting an impact evaluation of the cascade of aspects of the national delivery approach to the subnational level. Upheavals due to the COVID-19 pandemic affected the rollout of the cascade and made such an evaluation impossible. Our results do, however, speak to this policy agenda. The extent of variation in district office management practices, as well as in performance at district and school level warrants attention. To date, GES has emphasized strengthening of the prioritization and accountability functions in the subnational cascade. In light of the results presented in this paper, it seems important to better understand the potential of the problem-solving function.

---

[37]Note that this management practice refers to performance monitoring of district-level data and staff, and not monitoring visits to schools, which fall under activity intensity and task completion.

References

Adelman, Melissa and Renata Lemos, *Managing for Learning: Measuring and Strengthening Education Management in Latin America and the Caribbean* International Development in Focus, Washington, DC: World Bank, 2021.

Ampratwum, Edward, Mohammed Awal, and Franklin Oduro, "Decentralization and teacher accountability: The political settlement and sub-national governance in Ghana's education sector," in "The Politics of Education in Developing Countries: From Schooling to Learning," Oxford University Press, 2019.

Anderson, Michael L., "Multiple inference and gender differences in the effects of early intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Andrews, Matt, Lant Pritchett, and Michael Woolcock, *Building State Capability: Evidence, Analysis, Action*, Oxford University Press, 2017.

Asim, Minahil, Karen Mundy, Carly Manion, and Izza Tahir, "The missing middle of education service delivery in low- and middle-income countries," Working Paper 2022.

Azulai, Michel, Imran Rasul, Daniel Rogger, and Martin J. Williams, "Can training improve organizational culture? Experimental evidence from Ghana's civil service," 2020.

Bell, Sheena, Minahil Asim, Karen Mundy, Hope Pius Nudzor, Michael Boakye Yiadom, and Christopher Mensah Adosi, "Understanding the delivery approach cascade to regions, districts and schools: Evidence from a deep dive in five districts in Ghana," Working Paper 2023.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *The Quarterly Journal of Economics*, 2007, *CXXII* (4), 1351–1408.

Bloom, Nick, Renata Lemos, Raffaella Sadun, and John Van Reenen, "Does management matter in schools?," *The Economic journal*, 2015, *125* (584), 647–674.

_ , _ , _ , Daniela Scur, and John Van Reenen, "The new empirical economics of management," NBER Working Paper no. 20102 2014.

Bold, Tessa, Deon Filmer, Gayle Martin, Ezequiel Molina, Brian Stacy, Christophe Rockmore, Jakob Svensson, and Waly Wane, "Enrollment without learning: Teacher effort, knowledge, and skill in primary schools in Africa," *The Journal of Economic Perspectives*, 2017, *31* (4), 185–204.

Bruns, Barbara and Javier Luque, *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*, Washington, DC: World Bank, 2015.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers, "Missing in action: Teacher and health worker absence in developing countries," *The Journal of Economic Perspectives*, 2006, *20* (1), 91–116.

Cilliers, Jacobus, Eric Dunford, and James Habyarimana, "What do local government education managers do to boost learning outcomes?," *World Bank Economic Review*, 2022, *36* (3), 629–645.

Crawfurd, Lee, "School management and public-private partnerships in Uganda," *Journal of African Economies*, 2017, *26* (5), 539–60.

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

Deci, Edward L, Anja H. Olafsen, and Richard M. Ryan, "Self-determination theory in work organizations: The state of a science," *Annual Review of Organizational Psychology and Organizational Behavior*, 2017, *4* (1), 19–43.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande, "The personnel economics of the state," in "Handbook of Economic Field Experiments," Vol. 2, Elsevier, 2017, pp. 467–514.

Glewwe, Paul and Karthik Muralidharan, "Improving education otcomes in developing countries," in "Handbook of the Economics of Education," Vol. 5, Elsevier B.V, 2016, pp. 653–743.

Government of Ghana, "Education Act," 2008.

Honig, Dan and Lant Pritchett, "The limits of accounting-based accountability in education (and far beyond): Why more accounting will rarely solve accountability problems," RISE Working Paper 2019.

Honig, Daniel, "Managing for motivation as public performance improvement strategy in education and far beyond," CID Faculty Working Paper No. 409 2022.

Kim, Sangmook, Wouter Vandenabeele, Bradley E. Wright, Lotte Bøgh Andersen, Francesco Paolo Cerase, Robert K. Christensen, Céline Desmarais, Maria Koumenta, Peter Leisink, Bangcheng Liu, Jolanta Palidauskaite, Lene Holm Pedersen, James L. Perry, Adrian Ritz, Jeannette Taylor, and Paola De Vivo, "Investigating the structure and meaning of public service motivation across populations: Developing an international instrument and addressing issues of measurement invariance," *Journal of Public Administration Research and Theory*, 2013, *23* (1), 79–102.

Laffont, Jean-Jacques and David Martimort, *The Theory of Incentives*, Blackwell, 2002.

Leaver, Clare, Owen Ozier, Pieter Serneels, and Andrew Zeitlin, "Recruitment, effort, and retention effects of performance contracts for civil servants: Experimental evidence from Rwandan primary schools," *American Economic Review*, 2021, *111* (7), 2213–46.

_ , **Renata Lemos, and Daniela Scur**, "Understanding school management with public data: A new measurement approach and applications," 2022.

Lemos, Renata, Karthik Muralidharan, and Daniela Scur, "Personnel management and school productivity: Evidence from India," Working Paper 2021.

Liang, Jian, Crystal I. C. Farh, and Jiing-Lih Farh, "Psychological antecedents of promotive and prohibitive Voice: A two-wave examination," *Academy of Management Journal*, 2012, *55* (1), 71–92.

Liberto, Adriana Di, Fabiano Schivardi, and Giovanni Sulis, "Managerial practices and student performance," *Economic Policy*, 2015, *30* (84), 683–728.

Maslach, Christina and Susanna Jackson, "The measurement of experienced burnout," *Journal of Organizational Behavior*, 1981, *2*, 99–113.

McGregor, Douglas, *The Human Side of Enterprise*, McGraw-Hill, 1960.

Perry, James, "Measuring public service motivation: an assessment of construct reliability and validity," *Journal of Public Administration Research and Theory*, 1996, *6* (1), 5–24.

Perry, James L. and Wouter Vandenabeele, "Public service motivation research: Achievements, challenges, and future directions," *Public Administration Review*, 2015, *75* (5), 692–699.

Rammstedt, Beatrice and Oliver P. John, "Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German," *Journal of Research in Personality*, 2007, *41* (1), 203–212.

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

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

Sexton, John B, Robert L Helmreich, Torsten B Neilands, Kathy Rowan, Keryn Vella, James Boyden, Peter R Roberts, and Eric J Thomas, "The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research," *BMC Health Services Research*, 2006, *6* (1), 44–44.

Soto, Christopher J and Oliver P. John, "Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS," *Journal of Research in Personality*, 2017, *68* (C), 69–81.

Stallings, Jane A, Stephanie L Knight, and David Markham, *Using the Stallings Observation System to Investigate Time on Task in Four Countries*, World Bank, Washington, DC, 2014.

Walter, Torsten Figueiredo, "State management of education systems and educational performance: Evidence from a management survey in district education offices in Zambia," Working Paper 2018.

White, Howard Nial, *Books, Buildings and Learning Outcomes: An Impact Evaluation of World Bank Support to Basic Education in Ghana*, Washington, DC: World Bank, 2004.

Williams, Martin, Karen Mundy, Clare Leaver, Zahra Mansoor, Dana Qarout, Anna Billous, Minahil Asim, and Sheena Bell, "Delivery approaches to improving policy implementation: A conceptual framework," 2020.

Appendix A: Additional figures and tables
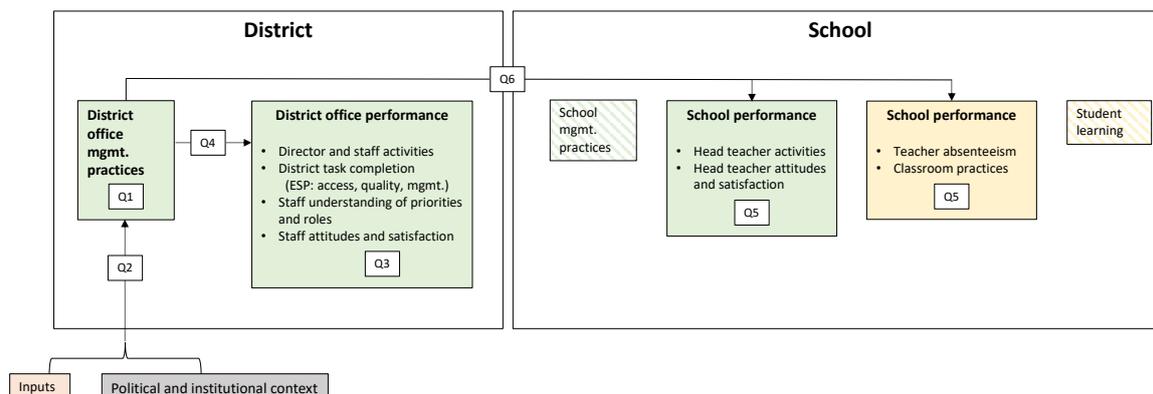
Figure A.1: Research Questions Illustrated on a "Theory of Change"

Prioritization & targets	Monitoring & use of data	Accountability & incentives	Problem solving & adaptation
1. Priority setting Tests whether there is a clear, focused set of priorities for the district 2. KPIs and targets Tests whether there are district KPIs and SMART ((Specific, Measurable, Achievable, Realistic, and Timely) targets against priorities 3. Strategic planning Tests whether there is a District Roadmap (or similar strategic plan) that is clearly linked to these priorities, KPIs, and targets with milestones and deadlines 4. Target interconnection Tests whether individual staff targets are aligned with districts targets (i.e., targets are cascaded and district-level targets are broken down into individual-level targets) 5. Priorities, KPI and target Communication Tests whether there is a clear communication plan for articulating district priorities, KPIs, and targets to district staff	6. Data collection Tests whether there are regularized routines specifying the data to be collected to monitor progress against each target on a monthly/quarterly basis 7. Data management and usage Tests whether there is a a district-office data management system, including a process to collate and visualize these monitoring data, and district management use these data 8. Performance monitoring (district-level) Tests whether district performance is monitored by district management 9. Performance monitoring (staff-level) Tests whether individual district staff performance is monitored by district management	10. Accountability structure Tests whether there is a clear line of accountability for each district target set against a KPI (e.g., via named 'target owners') 11. Accountability meetings Tests whether district staff (e.g., target owners) are asked to report in person on progress at monthly/quarterly meetings with the Director 12. Rewards and sanctions Tests whether staff receive any rewards and sanctions based on their performance through formal or informal means	13. Feedback loops Tests whether there are regular, structured opportunities for *two-way feedback* between actors across the delivery chain 14. Deep dives Tests whether problem-solving routines/deep dives are used to understand complex problems, and identify remedies, by engaging actors across the delivery chain (office, circuit, school) 15. Autonomy Tests whether district staff have discretion to choose activities to meet their targets against KPIs 16. Innovation Tests whether the district has channels and structures for its district staff to innovate and contribute with new ideas upwards

Figure A.2: Description of District Office Management Practices

Table A.1: List of Controls (Representative Sample)

	Mean	SD	Obs
District controls			
District-level GALOP score	-0.31	0.19	174
Organizational controls			
Number of staff	46.91	12.53	174
Staff skills index	6.04	1.69	174
Staff access to phone and internet (days, out of 5)	4.64	0.55	174
Staff access to vehicle and fuel (officers, out of 10)	3.72	2.28	174
Noise controls			
1[Male director]	0.61	0.49	174
1[Director has graduate degree]	0.97	0.17	174
Director's experience in GES, years	29.16	4.72	174
Director's Public Service Motivation index	0.01	1.01	174
Director's Big Five Personality index	0.02	1.00	174
1[Interview started before noon]	0.62	0.49	174
1[Director's interview went very well]	0.92	0.27	174
Respondent controls			
1[Male]	0.78	0.42	863
Age, years	49.84	5.89	862

Notes: The respondent controls are for directors, deputy directors, and SISOs. We have a missing value for the age variable because one director did not state their age. The values are weighted using design weights.

Table A.2: List of Controls (GEOP Sample)

	Mean	SD	Obs
District controls			
Organizational controls			
Number of staff	44.74	11.79	27
Staff skills index	5.03	1.46	27
Staff access to phone and internet (days, out of 5)	4.18	0.71	27
Staff access to vehicle and fuel (officers, out of 10)	3.97	2.47	27
Noise controls			
1[Male director]	0.67	0.48	27
Director's experience in GES, years	29.30	5.18	27
Director's Public Service Motivation index	-0.00	1.00	27
Director's Big Five Personality index	0.00	1.00	27
1[Interview started before noon]	0.63	0.49	27
1[Director's interview went very well]	0.93	0.27	27
School controls			
1[Male head teacher]	0.94	0.24	1,261
Head teacher age, years	41.27	7.44	1,261
1[Urban area]	0.06	0.25	1,261
1[Rural area]	0.81	0.39	1,261
1[Semi-urban area]	0.13	0.33	1,261
School facilities index	0.49	0.23	1,260
School GALOP score	-0.35	0.20	1,261
Permanent school teachers	6.78	3.83	1,261

Notes: We have a missing value for the school facilities index for one school. The values are weighted using design weights.

Appendix B: Measurement

The Supplementary Appendix contains the key parts of our survey instruments. Below, we explain how we use this data (and in two cases administrative data) to construct the control variables.

District controls

GALOP score   When selecting basic schools for the GALOP, the World Bank created an index based on BECE scores and a measure of social deprivation. Every basic school in Ghana was assigned a score, and the 10,000 most disadvantaged schools with the lowest scores were enrolled in the project. The Ministry of Education provided us with a list of these 10,000 schools together with their GALOP scores and Education Management Information System (EMIS) codes. We constructed a district-level GALOP score by computing the average for the six randomly selected schools per district in our sample.

Number of staff   This variable is constructed using a survey question administered to the deputy director of HR and supervision about the number of staff working in the district office.

Staff skills index   We asked the two deputy directors how many, out of every 10 officers in their office, can use a computer/laptop and software such as MS Word/Excel/PowerPoint. We computed the average of the four questions for each deputy, and we then averaged again to obtain a district-level variable.

Staff access to phone and internet   We asked the two deputy directors how many days, out of the five working days, the telephone worked for 50 percent or more of calls and internet access was good enough to check email. We computed the average of the two questions for each deputy, and we then averaged again to obtain a district-level variable.

Staff access to vehicle and fuel   We asked the two deputy directors how many, out of every 10 officers at the non-director level, had access to fuel and a vehicle (privately owned or otherwise) that could be used for work. We averaged each deputy's response to obtain a district-level variable.

Public service motivation (PSM)   We measured PSM using the framework proposed by (Perry, 1996); specifically, the 16-item scale proposed by Kim et al. (2013) and endorsed by Perry and Vandenabeele (2015).

Big Five personality traits   We used the short Big Five Inventory (BFI) validated by Rammstedt and John (2007) and Soto and John (2017)). The five traits are: agreeableness, openness, conscientiousness, extroversion, and neuroticism. We first averaged the raw scores

for each trait, then we normalized the averages score to get five normalized sub-indices, and finally we averaged all the normalized z-scores to obtain a BFI index.

School controls

School facilities   During the second, unannounced school visit, our enumerators completed a school observation module following the SDI methodology set out in Bold et al. (2017). The external observers noted whether the school had toilets (and if so whether they were clean) and/or water and soap. We also asked head teachers to self-report whether the school had access to electricity, computers, and/or the internet. We created a set of binary indicators coded to 1 if the facility was present and 0 otherwise. The school facilities index is the average of these six binary variables.

Permanent school teachers   This variable is constructed using a survey question administered to head teachers about the number of permanent teachers working in the school.

Appendix C: Multiple hypothesis test correction

This appendix describes our methodology for multiple hypothesis test correction and reports conventional $p-$values and corrected $q-$values for each regression table presented in Sections 6 and 7.

We studied eight dimensions of performance, four at the district level and four at the school level. To assess the relationship between district management practices and each outcome of interest, we used three different specifications (using the overall management index, the four sub-indices, and the accountability focus indicator). This gave us six different independent variables of interest. We ran these specifications for the overall sample, but sometimes also for sub-samples, by job level. Since our main takeaways in the Discussion section primarily rely on the results in the full sample, we carry out multiple hypothesis test correction for the full sample only. For each dimension of performance, we conduct corrections for the six independent variables of interest across all outcomes within that dimension. For example, the *district staff activity intensity* dimension includes three outcomes: school visits, teaching-focused activities, and other activities. Hence, we carry out corrections for 18 $p-$values. Specifically, we adjust conventional $p-$values for false detection rates and compute sharpened false discovery rate (FDR) "$q-$values" as outlined by Anderson (2008).[38] Below, we present eight tables showing the conventional $p-$values and the sharpened $q-$values corresponding to each dimension of performance.

---

[38]Anderson (2008) argues that controlling for the FDR, in comparison to conducting the standard Bonferroni correction, is generally more suited for exploratory analysis as it allows researchers to tolerate a small number of type I errors in exchange for more power.

Table C.1: Multiple Hypothesis Testing (Table 8: District Office Management Practices and Activity Intensity During School Visits)

	(1) Management index	(2) Prioritization sub-index	(3) Monitoring sub-index	(4) Accountability sub-index	(5) Problem-solving sub-index	(6) 1[Accountability focus]
School visits						
p-values	0.116	0.161	0.568	0.394	0.046**	0.282
q-values	0.337	0.370	0.610	0.584	0.289	0.584
q-values*	0.382	0.382	0.475	0.475	0.382	0.475
Teaching focused						
p-values	0.000***	0.066*	0.915	0.43	0.402	0.41
q-values	0.001***	0.289	1.00	0.584	0.584	0.584
q-values*	0.001***	0.198	1.00	0.525	0.525	0.525
Other activities						
p-values	0.000***	0.166	0.347	0.070*	0.606	0.703
q-values	0.001***	0.370	0.584	0.289	0.610	0.706
q-values*	0.001***	0.285	0.498	0.213	0.543	0.543

Notes: The sharpened q-values are based on the methodology in Anderson (2008). For each outcome, the first set of sharpened q-values show the correction based on all outcomes within this dimension of performance (note that this equals a total of 18 tests). The q-values* set denotes a less conservative correction, where we only correct for the six hypothesis tests for that outcome. The values are significant at the *10%, **5%, and ***1% levels.

Table C.2: Multiple Hypothesis Testing (Table 10: District Office Management Practices and District Task Completion)

	(1) Management index	(2) Prioritization sub-index	(3) Monitoring sub-index	(4) Accountability sub-index	(5) Problem-solving sub-index	(6) 1[Accountability focus]
Access						
p-values	0.512	0.016**	0.52	0.052*	0.842	0.787
q-values	1	0.169	1	0.364	1	1
q-values*	1	0.107	1	0.143	1	1
Quality						
p-values	0.560	0.784	0.568	0.606	0.536	0.756
q-values	1	1	1	1	1	1
q-values*	1	1	1	1	1	1
Management						
p-values	0.533	0.014**	0.682	0.371	0.664	0.534
q-values	0.834	0.144	0.834	0.834	0.834	0.834
q-values*	1	0.092*	1	1	1	1

Notes: The sharpened q-values are based on the methodology in Anderson (2008). For each outcome, the first set of sharpened q-values show the correction based on all outcomes within this dimension of performance (note that this equals a total of 18 tests). The q-values* set denotes a less conservative correction, where we only correct for the six hypothesis tests for that outcome. The values are significant at the *10%, **5%, and ***1% levels.

Table C.3: Multiple Hypothesis Testing (Table 12: District Office Management Practices and Understanding of Priorities and Roles)

	(1) Management index	(2) Prioritization sub-index	(3) Monitoring sub-index	(4) Accountability sub-index	(5) Problem-solving sub-index	(6) 1[Accountability focus]
Understanding of priorities						
p-values	0.138	0.048**	0.152	0.393	0.037**	0.567
q-values	0.613	0.405	0.613	0.839	0.405	1
q-values*	0.180	0.169	0.180	0.296	0.169	0.296
Understanding of roles						
p-values	0.84	0.337	0.956	0.334	0.699	0.637
q-values	1	0.839	1	0.839	1	1
q-values*	1	1	1	1	1	1

Notes: The sharpened q-values are based on the methodology in Anderson (2008). For each outcome, the first set of sharpened q-values show the correction based on all outcomes within this dimension of performance (note that this equals a total of 12 tests). The q-values* set denotes a less conservative correction, where we only correct for the six hypothesis tests for that outcome. The values are significant at the *10%, **5%, and ***1% levels.

Table C.4: Multiple Hypothesis Testing (Table 13: District Office Management Practices and District Staff Satisfaction)

	(1) Management index	(2) Prioritization sub-index	(3) Monitoring sub-index	(4) Accountability sub-index	(5) Problem-solving sub-index	(6) 1[Accountability focus]
Exhaustion						
p-values	0.875	0.694	0.456	0.970	0.440	0.422
q-values	1	1	1	1	1	1
q-values*	1	1	1	1	1	1
Intention to leave						
p-values	0.648	0.962	0.066*	0.761	0.021**	0.026**
q-values	1	1	0.283	1	0.185	0.185
q-values*	0.841	0.927	0.097*	0.841	0.085*	0.085*

Notes: The sharpened q-values are based on the methodology in Anderson (2008). For each outcome, the first set of sharpened q-values show the correction based on all outcomes within this dimension of performance (note that this equals a total of 12 tests). The q-values* set denotes a less conservative correction, where we only correct for the six hypothesis tests for that outcome. The values are significant at the *10%, **5%, and ***1% levels.

Table C.5: Multiple Hypothesis Testing [Table 15: District Office Management Practices and Head Teacher Satisfaction (GEOP Sample)]

	(1) Management index	(2) Prioritization sub-index	(3) Monitoring sub-index	(4) Accountability sub-index	(5) Problem-solving sub-index	(6) 1[Accountability focus]
Teaching focused						
p-values	0.257	0.013**	0.118	0.221	0.061*	0.727
q-values	0.396	0.085*	0.271	0.396	0.216	0.794
q-values*	0.259	0.085*	0.224	0.259	0.180	0.446
Other activities						
p-values	0.666	0.011**	0.387	0.622	0.071*	0.852
q-values	0.788	0.085*	0.513	0.788	0.216	0.788
q-values*	1	0.071*	1	1	0.216	1

Notes: The sharpened q-values are based on the methodology in Anderson (2008). For each outcome, the first set of sharpened q-values show the correction based on all outcomes within this dimension of performance (note that this equals a total of 12 tests). The q-values* set denotes a less conservative correction, where we only correct for the six hypothesis tests for that outcome. The values are significant at the *10%, **5%, and ***1% levels.

Table C.6: Multiple Hypothesis Testing [Table 16: District Office Management Practices and Head Teacher Satisfaction (GEOP Sample)]

	(1) Management index	(2) Prioritization sub-index	(3) Monitoring sub-index	(4) Accountability sub-index	(5) Problem-solving sub-index	(6) 1[Accountability focus]
Exhaustion						
p-values	0.223	0.89	0.733	0.280	0.964	0.498
q-values	0.686	1	1	0.725	1	1
q-values*	1	1	1	1	1	1
Intention to leave						
p-values	0.226	0.007***	0.001***	0.081*	0.446	0.783
q-values	0.686	0.041**	0.013**	0.370	1	1
q-values*	0.205	0.018**	0.007***	0.122	0.366	0.513

Notes: The sharpened q-values are based on the methodology in Anderson (2008). For each outcome, the first set of sharpened q-values show the correction based on all outcomes within this dimension of performance (note that this equals a total of 12 tests). The q-values* set denotes a less conservative correction, where we only correct for the six hypothesis tests for that outcome. The values are significant at the *10%, **5%, and ***1% levels.

Table C.7: Multiple Hypothesis Testing [Table 18: District Office Management Practices and Teacher Absence (GEOP Sample)]

	(1) Management index	(2) Prioritization sub-index	(3) Monitoring sub-index	(4) Accountability sub-index	(5) Problem-solving sub-index	(6) 1[Accountability focus]
p-values	0.839	0.141	0.006***	0.025**	0.007***	0.020**
q-values	0.204	0.060*	0.022**	0.026**	0.022**	0.026**

Notes: The sharpened q-values are based on the methodology in Anderson (2008). The values are significant at the *10%, **5%, and ***1% levels.

Table C.8: Multiple Hypothesis Testing [Table 19: District Office Management Practices and Time on Task (GEOP Sample)]

	(1) Management index	(2) Prioritization sub-index	(3) Monitoring sub-index	(4) Accountability sub-index	(5) Problem-solving sub-index	(6) 1[Accountability focus]
Active instructional time						
p-values	0.012**	0.029**	0.000***	0.007***	0.175	0.257
q-values	0.022**	0.030**	0.001***	0.020**	0.090*	0.124
q-values*	0.021**	0.025**	0.001***	0.018**	0.076*	0.094*
Students off-task						
p-values	0.000***	0.016**	0.055*	0.479	0.011**	0.841
q-values	0.001***	0.026**	0.044**	0.219	0.022**	0.32
q-values*	0.001***	0.028**	0.044**	0.238	0.028**	0.390
Teacher off-task						
p-values	0.647	0.039**	0.376	0.000***	0.012**	0.000***
q-values	0.296	0.034**	0.178	0.001***	0.012**	0.000***
q-values*	0.275	0.031**	0.178	0.001***	0.017**	0.001***

Notes: The sharpened q-values are based on the methodology in Anderson (2008). For each outcome, the first set of sharpened q-values show the correction based on all outcomes within this dimension of performance (note that this equals a total of 18 tests). The q-values* set denotes a less conservative correction, where we only correct for the six hypothesis tests for that outcome. The values are significant at the *10%, **5%, and ***1% levels.